



Εθνικό Μετσόβιο Πολυτεχνείο

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ

Ανάλυση Πολυμεσικού Περιεχομένου και
Κατηγοριοποίηση Χρηστών σε Περιβάλλοντα
Ηλεκτρονικής Μάθησης

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

της

ΠΑΡΑΣΚΕΥΗΣ Κ. ΤΖΟΥΒΕΛΗ

Διπλωματούχου Ηλεκτρολόγου Μηχανικού &
Μηχανικού Υπολογιστών Ε.Μ.Π.

Αθήνα, Ιούλιος 2009



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ

Ανάλυση Πολυμεσικού Περιεχομένου και Κατηγοριοποίηση Χρηστών σε Περιβάλλοντα Ηλεκτρονικής Μάθησης

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

της

ΠΑΡΑΣΚΕΥΗΣ Κ. ΤΖΟΥΒΕΛΗ

Διπλωματούχου Ηλεκτρολόγου Μηχανικού &
Μηχανικού Υπολογιστών Ε.Μ.Π.

Συμβουλευτική Επιτροπή: Στέφανος Κόλλιας (Επιβλέπων Καθηγητής)
Ανδρέας-Γεώργιος Σταφυλοπάτης
Παναγιώτης Τσανάκας

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή τη 2^α Ιουλίου 2009.

...
Σ.Κόλλιας
Καθηγητής Ε.Μ.Π.

...
Α. Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

...
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

...
Π. Μαραγκός
Καθηγητής Ε.Μ.Π.

...
Α. Ντελόπουλος
Επ. Καθηγητής Α.Π.Θ.

...
Ν. Παπασπύρου
Επ. Καθηγητής Ε.Μ.Π.

...
Γ. Στάμου
Λέκτορας Ε.Μ.Π.

Αθήνα, Ιούλιος 2009

...

Παρασκευή Κ. Τζούβελη

Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

© 2009 - Με επιφύλαξη παντός δικαιώματος - All rights reserved

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περιεχόμενα

Περιεχόμενα	i
Κατάλογος Σχημάτων	v
Κατάλογος Πινάκων	ix
Πρόλογος	xi
Περίληψη	xiii
1 Εισαγωγή	1
1.1 Συστήματα μαθησιακής τεχνολογίας και κατηγοριοποίηση περιεχομένου και χρηστών	1
1.2 Συνεισφορές και δομή διδακτορικής διατριβής	3
2 Συστήματα Μαθησιακής Τεχνολογίας	5
2.1 Εισαγωγή	5
2.2 Εξ απόστασεως εκπαίδευση	5
2.3 Προτυποποίηση μαθησιακών τεχνολογιών	7
2.4 Επισκόπηση αρχιτεκτονικής του συστήματος του IEEE για μαθησιακή τεχνολογία	10
2.5 Προτεινόμενη επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE	13
3 Εξόρυξη Γνώσης και Προσωποποιημένη Πρόσβαση	17
3.1 Εισαγωγή	17
3.2 Εξόρυξη γνώσης από δεδομένα	17
3.2.1 Εξέταση δεδομένων	18
3.2.2 Επιλογή μεθόδου μέτρησης ομοιότητας	19
3.2.3 Ταίριασμα μοντέλου στα δεδομένα	20
3.3 Προτεινόμενη τεχνική εξόρυξης γνώσης	26
3.3.1 Επιλογή χαρακτηριστικών και καθορισμός απόστασης	27
3.3.2 Βελτίωση συσταδοποίησης μέσω κατηγοριοποίησης	28
3.4 Προσωποποιημένη πρόσβαση και συστήματα μαθησιακής τεχνολογίας	29
3.4.1 Κατηγοριοποίηση χρηστών βασισμένη σε ερωτηματολόγια	29
3.4.2 Δυναμική εξαγωγή προφίλ χρηστών	33
3.4.3 Κατηγοριοποίηση χρηστών βασισμένη σε καταγραφή συμπεριφοράς	35

4	Ανάλυση και τεκμηρίωση πολυμεσικού περιεχομένου	51
4.1	Εισαγωγή	51
4.2	Ερευνητικό πλαίσιο	52
4.3	Ερμηνεία της Βυζαντινής τέχνης- Διονύσιος εκ Φουρνά.....	54
4.4	Σύστημα αυτόματης κατηγοριοποίησης Βυζαντινών εικόνων	59
4.4.1	Σύστημα αυτόματης κατηγοριοποίησης Βυζαντινών εικόνων με βάση εντοπισμό του φωτοστέφανου	60
4.4.2	Σύστημα αυτόματης κατηγοριοποίησης και τεκμηρίωσης Βυζαντινών εικόνων με βάση τη γνώση	66
4.4.3	Εξαγωγή χαρακτηριστικών από τα σημασιολογικά τμήματα και σημασιολογική ερμηνεία	85
5	Αναπαράσταση Γνώσης και Συστήματα Εξ Αποστάσεως Εκπαίδευσης	95
5.1	Εισαγωγή	95
5.2	Οντολογίες και Σημασιολογικός Ιστός	96
5.3	Τυπικές μέθοδοι αναπαράστασης γνώσης και συλλογιστική.....	97
5.3.1	Γλώσσες αναπαράστασης γνώσης	98
5.3.2	Υπηρεσίες και μηχανισμοί εξαγωγής συμπερασμάτων	102
5.4	Συστήματα γνώσης και συστήματα μαθησιακής τεχνολογίας	104
5.4.1	Αναπαράσταση πληροφορίας από ανάλυση πολυμεσικών δεδομένων με χρήση ασαφούς γνώσης	106
5.5	Πειραματικά αποτελέσματα	109
6	Ασφάλεια Πολυμεσικού Περιεχομένου	113
6.1	Εισαγωγή	113
6.2	Ανασκόπηση τεχνικών υδατογράφησης πολυμεσικών αρχείων.....	115
6.3	Μέθοδος σημασιολογικής υδατογράφησης πολυμεσικού περιεχομένου ..	117
6.3.1	Διακριτός μετασχηματισμός κυματιδίων	117
6.3.2	Δέντρα συντελεστών κυματιδίων	119
6.3.3	Μέθοδος εισαγωγή υδατογραφήματος σε πολυμεσικό υλικό	120
6.3.4	Μέθοδος εξαγωγής υδατογραφήματος από πολυμεσικό υλικό ...	122
6.4	Πειραματικά Αποτελέσματα Μεθόδου Υδατογράφησης	123
7	Εφαρμογές	127
7.1	Εισαγωγή	127
7.2	1ο σενάριο: Σύστημα εξ αποστάσεως εκπαίδευσης σε θέματα νέων τεχνολογιών	127
7.2.1	Στατιστική ανάλυση των ερωτηματολογίων	131
7.3	2ο σενάριο: Σύστημα εκπαίδευσης βασισμένη σε ανάλυση μη λεκτικών στοιχείων και με λειτουργία υποβοήθησης ανάγνωσης	133
7.3.1	Γενική περιγραφή ευφυούς συστήματος ανάγνωσης	135
7.3.2	Περιγραφή υποσυστημάτων ευφυούς συστήματος ανάγνωσης ...	136
7.3.3	Δοκιμαστικά σενάρια χρήσης	141
7.4	3ο σενάριο: Σύστημα εκπαίδευσης σε αγιογραφίες της Βυζαντινής τέχνης.....	142

8 Συμπεράσματα και κατευθύνσεις μελλοντικής έρευνας	147
Κατάλογος δημοσιεύσεων	149
Βιβλιογραφία	153
Παράρτημα Α: Σώμα ορολογίας για πρόσωπο Βυζαντινής μορφής	167
Παράρτημα Β: Ερωτηματολόγια συστήματος εξ αποστάσεως εκπαίδευσης σε θέματα νέων τεχνολογιών	??
Παράρτημα Γ: Στατιστική ανάλυση απαντήσεων χρηστών στα ερωτηματολόγια συστήματος εξ αποστάσεως εκπαίδευσης	199

Κατάλογος Σχημάτων

2.1	Επίπεδα αρχιτεκτονικής συστήματος μαθησιακής τεχνολογίας	10
2.2	Σύστημα Μαθησιακής Τεχνολογίας του IEEE	11
2.3	Αρχιτεκτονική προτεινόμενου ΣΜΤ.....	14
3.1	Ιεραρχικός αλγόριθμος: στοιχεία και δενδρόγραμμα	23
3.2	Συνάρτηση συμμετοχής: απόδοση ασάφεια σε πιθανό σύνολο απαντήσεων.....	30
3.3	Δομή αντιστοίχισης Προφίλ Χρηστών	33
3.4	Εμπλουτισμός της διαδικασίας Αξιολόγησης.....	40
3.5	Αποτελέσματα εκτίμησης στάσης κεφαλής.....	41
3.6	Κατανομή των τιμών που λαμβάνει το κλάσμα της τρέχουσας τιμής της απόστασης προς την τιμή που είχε αυτή η απόσταση στο καρέ αρχικοποίησης για σύνολο 250 διαδοχικών καρέ	44
3.7	Τρισδιάστατη απεικόνιση του μήκους του διανύσματος βλέμματος και στάσης της κεφαλής σε σχέση με τις καταστάσεις του χρήστη.....	45
3.8	Τρισδιάστατη απεικόνιση του μήκους του κλάσματος της απόστασης μεταξύ ματιών και στάσης της κεφαλής σε σχέση με τις καταστάσεις του χρήστη	45
4.1	Αναπαράσταση μέτρων για ιστάμενες Μορφές σύμφωνα με τα μέτρα του Διονυσίου εκ Φουρνά	57
4.2	Εντοπισμός κεφαλής με χρήση φωτοστέφανου: (α) Εξαγωγή κεφαλής από εικόνα (β) Εξαχθείσες κεφαλές από διάφορες μορφές	62
4.3	Αρχιτεκτονική του συστήματος τεκμηρίωσης κατηγοριοποίησης Βυζαντινών εικόνων. Η μονάδα <i>Ανάλυση Βυζαντινών εικόνων</i> εξάγει πληροφορίες για την εικόνα και η μονάδα <i>Αναπαράσταση Γνώσης και Συλλογιστική</i> χρησιμοποιεί αυτές τις πληροφορίες για την εξαγωγή της υπονοούμενης γνώσης, κατηγοριοποιώντας την εικόνα	66
4.4	Ανάλυση Βυζαντινής εικόνας. Κατ' αρχάς, ο αλγόριθμος ανιχνεύει την περιοχή του προσώπου καθώς και τα μάτια, τη μύτη του και τον άξονα αντικίνησης. Κατόπιν, εξάγει τις σημασιολογικές περιοχές των μαλλιών, του μετώπου, των παρειών, της γενειάδας και του μουστακιού καθώς και τα βασικά χρώματα που χρησιμοποιεί ο αγιογράφος για τα σαρκώματα του προσώπου. Η περαιτέρω ανάλυση των εξαγόμενων τμημάτων του προσώπου παρέχει πληροφορίες για τα χαρακτηριστικά του προσώπου. Τέλος, ο αλγόριθμος παράγει μια σημασιολογική ερμηνεία για κάθε ένα από αυτά τα χαρακτηριστικά γνωρίσματα.	68

4.5	Ένα παράδειγμα συλλογιστικής. Οι εξαγόμενες πληροφορίες από τη μονάδα ανάλυσης εικόνας αποτελούν το ασαφές σώμα ρόλων (ABox) της βάσης γνώσεων. Το ασαφές σώμα ορολογίας (TBox) καθορίζεται βάσει του εγχειριδίου του Διονυσίου εκ Φουρνά. Αυτά τα σώματα αποτελούν την είσοδο για τη μηχανή συλλογισμού, η οποία εξάγει πληροφορίες για την εικόνα.	69
4.6	Τέσσερις τύποι Haar-ομοειδών χαρακτηριστικών γνωρισμάτων. Αυτά τα χαρακτηριστικά γνωρίσματα εμφανίζονται σε διαφορετικές θέσεις και κλίμακες. Τα Haar-ομοειδή χαρακτηριστικά γνωρίσματα υπολογίζονται ως διαφορά των σκουρόχρωμων και ανοιχτόχρωμων περιοχών. Μπορούν να θεωρηθούν ως χαρακτηριστικά γνωρίσματα που συλλέγουν τις τοπικές πληροφορίες ακμών με διαφορετικό προσανατολισμό και κλίμακα. Το σύνολο των Haar-ομοειδών χαρακτηριστικών γνωρισμάτων είναι μεγάλο, και μόνο μια μικρή ποσότητα αυτών μαθαίνεται από τα θετικά και αρνητικά παραδείγματα για την ανίχνευση προσώπου	73
4.7	α) Περιοχή προσώπου, β) οριζόντιες συντεταγμένες διανυσματικού πεδίου, γ) κάθετες συντεταγμένες περιοχής προσώπου	75
4.8	Μετασχηματισμός Hough α) Αρχική εικόνα β) Εικόνα μετασχηματισμού Hough	77
4.9	Παράδειγμα κατάτμησης εικόνας 3x3 όπως εμφανίζεται στην εργασία [29].....	82
4.10	Εφαρμογή του αλγορίθμου με επιλογή διαφορετικού πλήθους κατωφλίων	84
4.11	Ασαφής Διαμέριση Σχήματος Μαλλιών	86
4.12	Κατηγορίες κεφαλής (1) Χριστός, (2) Παναγία και (3) Άλλοι Άγιοι ...	89
4.13	Εντοπισμός προσώπου, ματιών, μύτης και άξονα αντικίνησης.....	91
4.14	Εντοπισμός σημασιολογικών τμημάτων προσώπου	91
6.1	Πυραμιδική αποσύνθεση μιας εικόνας	117
6.2	Η πυραμιδική αποσύνθεση τριών στρώματος μιας εικόνας	118
6.3	Εικόνα-Υδατογράφημα	121
6.4	Στάδια αποσύνθεσης σημασιολογικής εικόνας. Αρχικά εξάγεται το σημασιολογικό τμήμα της εικόνας, το οποίο αποσυνθέτεται με χρήση του μετασχηματισμού κυματιδίου	121
6.5	Διάγραμμα μεθόδου εισαγωγής υδατογραφήματος	122
6.6	Διάγραμμα μεθόδου εξαγωγής υδατογραφήματος	123
6.7	Στάδια μεθόδου υδατογράφησης α) Αρχική εικόνα, β) Σημασιολογικό αντικείμενο, γ) Εικόνα υδατογράφημα, δ)Υδατογραφημένη εικόνα	123
6.8	Εξαγόμενα υδατογραφήματα για διαφορετικούς λόγους συμπίεσης(13.2,9.5,7.9,3.2)	124
6.9	Εξαγόμενα υδατογραφήματα για τρόπους επεξεργασίας εικόνας όπως αναφέρονται στον Πίνακα 6.2.....	125
6.10	Ευρωστία συστήματος υδατογράφησης σε βυζαντινές εικόνες. Στην αρχική εικόνα εφαρμόζεται η μέθοδος εξαγωγή μορφών Αγίων από Βυζαντινές εικόνες βασισμένο σε γνώση και στην εξαγόμενη μορφή ενσωματώνεται το υδατογράφημα	125

6.11 α. Επίθεση 1η: αποκοπή όλου του αντικειμένου και τμήματος του φόντου, β. Επίθεση 2η: αποκοπή μέρους του αντικειμένου, γ. Περίπτωση 1η: Εξαγωγή υδατογραφήματος και δ. Περίπτωση 2η: Εξαγωγή υδατογραφήματος	126
7.1 Ερωτηματολόγιο σχετικό με τη σχολική μονάδα	129
7.2 Ερωτηματολόγιο για τον χρήστη	129
7.3 Αποτελέσματα κατηγοριοποίησης όλων των χρηστών στα βασικά προφίλ και παρουσίαση των προφίλ του κάθε χρήστη ανάλογα με το κάθε ερωτηματολόγιο	130
7.4 Αποτελέσματα κατηγοριοποίησης όλων των χρηστών στα βασικά προφίλ και παρουσίαση των προφίλ του κάθε χρήστη ανάλογα με το κάθε ερωτηματολόγιο	132
7.5 Κατηγοριοποίηση απαντήσεων στο ερωτηματολόγιο Α σύμφωνα με τα προφίλ που ορίστηκαν ανά ερωτηματολόγιο Β	133
7.6 Δομικά υποσυστήματα του ευφυούς συστήματος ανάγνωσης	137
7.7 Ανθίβολα μορφής νεαρού Αγίου	144
7.8 Τμήμα σύνθεσης- Περιβάλλον διεπαφής - αρχική οθόνη	144
7.9 Τμήμα σύνθεσης- Περιβάλλον διεπαφής - τελική οθόνη	145
1 Στατιστική ανάλυση συνολικής συμμετοχής	199
2 Στατιστική ανάλυση συνολικής συμμετοχής σχολείων	200
3 Στατιστική ανάλυση συνολικής συμμετοχής χρηστών	200
4 Ερώτηση Β.1.1.Είναι η θέση σας	201
5 Ερώτηση Β.1.3.Προσόντα/ κατάρτιση σε Ειδικές Εκπαιδευτικές Ανάγκες	201
6 Ερώτηση Β.1.5.Η ομάδα ηλικίας σας	201
7 Ερώτηση Β.2.1. Έχετε Η/Υ στο σπίτι	202
8 Ερώτηση Β.2.1. Έχετε σύνδεση με το Διαδίκτυο (Internet) στο σπίτι	202
9 Ερώτηση Β.2.4.Για ποιο από τα παρακάτω χρησιμοποιήσατε τον Η/Υ τουλάχιστον μια φορά στον μήνα που πέρασε	202
10 Ερώτηση Β.2.5.Ποια από τις παρακάτω εργασίες εκτελέσατε τουλάχιστον μια φορά, χωρίς καμία βοήθεια	203
11 Ερώτηση Β.3.1.Υπάρχουν Η/Υ στο περιβάλλον της εργασίας σας	203
12 Ερώτηση Β.3.2.Πόσο συχνά χρησιμοποιήσατε Η/Υ την τελευταία εβδομάδα στο σχολείο	204
13 Ερώτηση Β.3.3.Έχετε πρόσβαση στο Διαδίκτυο ή σε εκπαιδευτικό λογισμικό στο εργασιακό σας περιβάλλον	204
14 Ερώτηση Β.3.5.Χρησιμοποιείτε το Διαδίκτυο για έρευνα και αναζήτηση πληροφοριών που αφορούν σε ανάγκες των μαθητών ΜΕΑ	205
15 Ερώτηση Β.3.6.Χρησιμοποιείτε το Διαδίκτυο από το σχολείο για να αναζητήσετε πρόσθετες πηγές εκπαιδευτικού υλικού	205
16 Ερώτηση Β.3.7.Σε ποιο χώρο γενικά αναζητούν οι μαθητές σας πρόσβαση στο Διαδίκτυο για εκπαιδευτικούς σκοπούς	206
17 Ερώτηση Β.3.10.Αποψη που αφορά στην τεχνολογική ποιότητα του εκπαιδευτικού λογισμικού που διατίθεται στην τάξη σας	206

Κατάλογος Πινάκων

3.1	Επιλογή χαρακτηριστικών γνωρισμάτων	31
3.2	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 3 συστάδες	34
3.3	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 5 συστάδες	35
3.4	Αποτελέσματα συσταδοποίησης για 100 χρήστες - 9 συστάδες	35
3.5	Οπτικές Ενδείξεις και Χαρακτηριστικά	47
3.6	Κατηγοριοποίηση Πιθανών Καταστάσεων Χρήστη	48
3.7	Επίδοση κατηγοριοποίησης πιθανών καταστάσεων χρήστη με χρήση νευροασαφούς δικτύου	48
3.8	Σύνδεση κατάστασης χρήστη με την αντίδραση του συστήματος	49
4.1	Κατηγοριοποίηση Μαλλιών	60
4.2	Κατηγοριοποίηση Μετώπου	60
4.3	Κατηγοριοποίηση Παρειάς	60
4.4	Κατηγοριοποίηση Μουσταχίου	61
4.5	Κατηγοριοποίηση Γενειάδας	61
4.6	Βάρη Ακμών Συνόλου E	83
4.7	Μέση τιμή και διακύμανση του συνόλου δεδομένων UD	90
4.8	Σώμα Ισχυρισμών όπως αυτό εξήχθη αυτόματα από τη μονάδα επεξεργασίας Βυζαντινής εικόνας	92
5.1	Σημασιολογία Εννοιών και Ρόλων	101
5.2	Ασαφή ερωτήματα για τη μηχανή συμπερασματολογίας	110
5.3	Αξιολόγηση συστήματος με τη χρήση των μέτρων ορθότητας και ανάκλησης	111
6.1	Εξαγωγή υδατογραφήματος από την περιοχή προσώπου και σώματος έπειτα από συμπίεση JPEG με απώλειες	124
6.2	Εξαγωγή υδατογραφήματος από την περιοχή προσώπου και σώματος έπειτα από επιθέσεις επεξεργασίας εικόνας	125

Πρόλογος

Η διατριβή αυτή είναι το αποτέλεσμα μιας ερευνητικής πορείας που ξεκίνησε στο τέλος του 2001 στο εργαστήριο Εικόνας, Βίντεο και Συστημάτων Πολυμέσων του Ε.Μ.Π. Κατά τη διάρκεια αυτής της πορείας, πέρα της προσωπικής μου προσπάθειας, υπήρξε ένα πλήθος ανθρώπων που συνέβαλαν στην ολοκλήρωσή της και για τους οποίους θα ήθελα να εκφράσω την εκτίμησή μου.

Πρωτίστως, θα ήθελα να αναφερθώ στο σημαντικό ρόλο που έπαιξε ο επιβλέπων καθηγητής μου κ. Στέφανος Κόλλιας στην εξέλιξη της έρευνάς μου. Τον ευχαριστώ θερμά για την εμπιστοσύνη και καθοδήγηση του και που χωρίς ποτέ να περιορίσει την επιστημονική μου ελευθερία, έδειξε την αμέριστη στήριξή του, κατά την ενασχόλησή μου με τόσο ενδιαφέροντα ερευνητικά πεδία. Θα ήθελα να ευχαριστήσω και τα άλλα δύο μέλη της τριμελούς μου επιτροπής, Ανδρέα Σταφυλοπάτη και Παναγιώτη Τσανάκα, Καθηγητές Ε.Μ.Π., οι οποίοι ήταν πάντα πρόθυμοι να με βοηθήσουν και να με στηρίξουν, όποτε τους ζητήθηκε. Ευχαριστώ επίσης τους Καθηγητές Πέτρο Μαραγκό, Τάσο Ντελόπουλο, Νικόλαο Παπασπύρου και Γεώργιο Στάμου, οι οποίοι μου έκαναν την τιμή να είναι μέλη της επταμελούς μου επιτροπής.

Τα ερεθίσματα που δέχθηκα όλα αυτά τα χρόνια από το περιβάλλον του εργαστηρίου ήταν καθοριστικά για την πρόοδο της διατριβής. Οι καινοτόμες ιδέες που προτείνονται στη διατριβή αυτή, προήλθαν από καρποφόρες συνεργασίες που είχα με τα μέλη καθώς και με συνεργάτες του εργαστηρίου, πάντα υπό τη δημιουργική επίβλεψη του κ. Κόλλια.

Στα πρώτα βήματά μου, καθοριστική ήταν η βοήθεια και ενθάρρυνση τόσο του Νικόλα Τσαπατσούλη, Καθηγητή Τεχνολογικού Πανεπιστημίου Κύπρου για την ενασχόλησή μου με συστήματα εξ αποστάσεως εκπαίδευσης στο πλαίσιο της επικοινωνίας ανθρώπου - μηχανής, όσο και του Κλήμη Νταλιάνη, Διδάκτορα Ε.Μ.Π. για την έρευνα της περιοχής της προστασίας πολυμεσικού υλικού. Στη μετέπειτα πορεία μου, με την προτροπή του Δρ. Γιάννη Αβρίθη, άρχισε η ενασχόλησή μου με την ανάλυση πολυμεσικού υλικού και εξαγωγή γνώσης από αυτό. Ο ρόλος του Γιάννη ήταν ουσιαστικός, ενθαρρύνοντάς την έρευνά μου στη συγκεκριμένη περιοχή, καθώς και την οργάνωση της συνολικής εργασίας μου και τον ευχαριστώ για αυτό. Άφθογες ήταν και οι ερευνητικές συνεργασίες που είχα με τον Φοίβο Μυλωνά, Διδάκτορα Ε.Μ.Π., τον Κώστα Καρπούζη, Ερευνητή ΕΠΙΣΕΥ και το Στυλιανό Αστεριάδη, Υποψήφιο Διδάκτορα Ε.Μ.Π. στο πλαίσιο της μοντελοποίησης των προφίλ χρηστών. Άριστη συνεργασία υπήρξε επίσης με τον Νίκο Σίμου, Υποψήφιο Διδάκτορα Ε.Μ.Π. και τον Γεώργιο Στάμου, Λέκτορα Ε.Μ.Π. με σκοπό τη μοντελοποίηση της γνώσης από το πεδίο της Βυζαντινής Αγιογραφίας και τη δημιουργία συστήματος αυτόματης κατηγοριοποίησης και τεκμηρίωσης των Βυζαντινών εικόνων. Ιδιαίτερα θα ήθελα να ευχαριστήσω την Αμαρυλλίδα Ραουζαίου, Διδάκτορα Ε.Μ.Π. για τη συνεχή ενθάρρυνση, συμπαράσταση και βοήθεια, οποιαδήποτε στιγμή και αν της ζητήθηκε,

καθ' όλη τη διάρκεια εκπόνησης της διατριβής, αλλά πάνω από όλα για τη φίλια της.

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω όλα τα μέλη του εργαστήριο Εικόνας, Βίντεο και Συστημάτων Πολυμέσων του Ε.Μ.Π. για την πάντα χαρούμενη και δημιουργική συνύπαρξη που είχαμε όλα αυτά τα χρόνια. Όλοι τους, παλιά και νέα μέλη, μηδενός εξαιρουμένου, ήταν πάντα πρόθυμοι να προσφέρουν τη βοήθειά τους, τόσο σε ερευνητικά όσο και σε τεχνικά θέματα που με απασχόλησαν, κατά τη διάρκεια της διατριβής αυτής. Τέλος, θα ήθελα να εκφράσω ιδιαίτερες ευχαριστίες στον κ. Δημήτρη Καλομοιράκη, Ερευνητή Βυζαντινής τέχνης και στην υπόλοιπη ομάδα από το Ίδρυμα του Όρους Σινά για την άψογη συνεργασία μας. Επίσης, θα ήθελα να ευχαριστήσω τους φίλους μου που στάθηκαν κοντά μου καθ' όλη τη διάρκεια των σπουδών μου.

Αυτή η διατριβή, όπως και ό,τι άλλο, οφείλεται και αφιερώνεται σε όλη την οικογένειά μου και ιδιαίτερα στους γονείς μου, Άννα και Κώστα και στο σύζυγό μου, Γαβριήλ. Ελπίζω να αποτελεί μία μικρή ανταπόδοση για την αγάπη, φροντίδα και υπομονή τους για όλα αυτά τα χρόνια.

Αθήνα, Ιούλιος 2009

Περίληψη

Οι σύγχρονες τάσεις στην ανάπτυξη συστημάτων μαθησιακής τεχνολογίας έχουν δασκαλο-κεντρικό και χρηστο-κεντρικό χαρακτήρα, οδηγώντας την έρευνα σε προσωποποιημένη πρόσβαση στις πληροφορίες. Η παρούσα διατριβή προσεγγίζει το στόχο αυτό με ανάπτυξη και χρησιμοποίηση αποτελεσματικών τεχνικών κατηγοριοποίησης και τεκμηρίωσης του εκπαιδευτικού περιεχομένου. Προτείνει επιπλέον τεχνικές αυτόματης αξιολόγησης της κατάστασης των χρηστών, αναλύοντας τη συμπεριφορά τους (κινήσεις κεφαλής, προσοχή), όπως και ανάλυση του προφίλ κάθε χρήστη για αποτελεσματικότερη εκμάθηση του περιεχομένου. Η ανάγκη για αυτόματη κατηγοριοποίηση του πολυμεσικού περιεχομένου οδήγησε την έρευνά μας στην ψηφιακή επεξεργασία των εικόνων και ειδικότερα στην εξόρυξη γνώσης από αυτές. Στο πλαίσιο αυτό, διαπιστώθηκε η συχνή εμφάνιση προσώπων μέσα σε εικόνες πολιτιστικού, όπως ιστορικού και θρησκευτικού, περιεχομένου. Η διατριβή προτείνει ένα σύστημα ανίχνευσης προσώπων και χαρακτηριστικών τους, που βασίζεται στην υπάρχουσα γνώση για αυτά. Τέλος, για να καλυφθεί η ανάγκη για προστασία του ψηφιοποιημένου εκπαιδευτικού περιεχομένου στο σύστημα μαθησιακής τεχνολογίας, προτείνεται μια μέθοδος προστασίας του με τεχνικές ψηφιακής υδατογράφησης.

Abstract

The teacher- and student-based nature of current learning technology systems necessitates the personalized access to annotated educational material. The development of techniques for categorization of learning content tackles this issue. This dissertation proposes additional techniques for automatic users' evaluation, through the analysis of their behavior (head movements, attention) and for user profile's analysis towards an effective learning procedure. The requirement for automatic categorization of multimedia content led us to use image processing and data mining methods. In this framework, we decided to use pictures of cultural, historical and religious, content. The dissertation proposes a system of facial features' detection, based on experts knowledge and a technique of digital watermarking for educational semantic content protection.

Κατάλογος Απόδοσης Όρων

affect interpretation	:	συναισθηματική ερμηνεία
affective computing	:	υπολογιστική με βάση το συναίσθημα
agents	:	πράκτορες
agglomerative algorithms	:	συσσωρευτικοί αλγόριθμοι
alphabet	:	αλφάβητο
Assertion Box - ABox	:	Σώμα ισχυρισμών
association rules	:	κανόνες συσχετίσεων
atomic concepts	:	ατομικές έννοιες
atomic roles	:	ατομικοί ρόλοι
average link	:	τεχνική μέσου συνδέσμου
blind	:	τυφλά
browsing	:	πλοήγηση
calibration	:	ευθυγράμμιση
capacity	:	χωρητικότητα
cascade of classifiers	:	διαδοχικά συνδεδεμένοι ταξινομητές
categorical	:	κατηγορηματικούς
centroid	:	εικονικό κέντρο ομάδας
city block	:	δομική απόσταση
classification rate	:	βαθμός (σωστής) κατηγοριοποίησης
classification	:	κατηγοριοποίηση
clique potentials	:	δυναμικά κλικών
clustering	:	συσταδοποίηση
Color Layout Descriptor	:	περιγραφέας διάταξης χρώματος
color predicates	:	αποδόσεις των χρωμάτων
competitive learning	:	ανταγωνιστική μάθηση
complete link	:	τεχνική πλήρους συνδέσμου
completeness	:	πληρότητα
context	:	πλαίσιο γνώσης
conversational models	:	μοντέλα συζήτησης
covariance	:	συμμεταβλητότητα
cropping	:	κοπή τμήματος
Description Logics	:	Περιγραφική λογική
descriptive model	:	περιγραφικό μοντέλο
dimensionality curse	:	κατάρρα των υψηλών διαστάσεων

domain of interpretation	:	χώρος ερμηνείας
Dominant Color Descriptor	:	περιγραφέας κύριων χρωμάτων
down-to-top	:	ανοδική ή από κάτω προς τα πάνω
dual triple	:	δυική τριάδα
Edge Histogram Descriptor	:	περιγραφέας ιστογράμματος ακμών
equivalence axioms	:	αξίωμα ισοδυναμίας
expert user	:	έμπειρος χρήστης
expressive power	:	εκφραστική δύναμη
eye gaze	:	βλέμμα
face tilt	:	γωνία περιστροφής του προσώπου
false positive	:	ψευδώς θετικά
feature	:	μετρήσιμο χαρακτηριστικό
First order logics	:	Λογική πρώτης τάξης
forest-like	:	δασική δομή
formal meaning	:	τυπικό νόημα
fuzzy concepts	:	ασαφείς έννοιες
Fuzzy logic	:	Ασαφής λογική
Gaussian probability density function	:	Γκαουσιανή συνάρτηση πυκνότητας πιθανότητας
gradient descent	:	μέθοδος καθόδου κλίσης βαθμίδας
Graphical User Interface	:	γραφικό περιβάλλον αλληλεπίδρασης
ground truth	:	υπόθεση αλήθειας
heterogeneous	:	ετερογενής
Hidden Markov Models-HMM	:	
hierarchical	:	ιεραρχικός
Higher order logics	:	Λογική υψηλότερης τάξης
Homogeneous Texture Descriptor	:	περιγραφέας ομοιογενούς υφής
implicit knowledge	:	υπονοούμενη γνώση
indexing	:	δεικτοδότηση
individuals	:	άτομα
information retrieval system (IRS)	:	σύστημα αναζήτησης πληροφορίας / συστήματα ανάκτησης πληροφορίας
interconnectivity	:	συνδεσιμότητα
interoperability	:	διαλειτουργικότητα
interpretation	:	ερμηνεία
knowledge-based system	:	σύστημα γνώσης ή σύστημα βασισμένο στη γνώση
LargeDB	:	μεγάλες βάσεις δεδομένων
maximum likelihood algorithm	:	αλγόριθμος μέγιστης πιθανοφάνειας
median filter	:	φίλτρο ενδιάμεσης τιμής
metadata	:	μεταδεδομένα
mixed mode variables	:	δεδομένα μεικτού τύπου
modifier / linguistic hedge	:	ασαφής τροποποιητής
NLP	:	επεξεργασία φυσικής γλώσσας

non-blind	:	εμφανής
non-intrusive	:	μη παρεμβατικός
partition function	:	παράγοντας κανονικοποίησης
partitioning clustering	:	διαμεριστική μέθοδος ομαδοποίησης
pattern	:	πρότυπο
perceptual transparency	:	διαφάνεια
precision	:	ακρίβεια
prediction	:	πρόβλεψη
predictive model	:	προβλέψιμα μοντέλα
quantization	:	κβαντισμός
query by example (QbE)	:	ερώτημα μέσω παραδείγματος
radial basis function networks	:	νευρωνικά δίκτυα συναρτήσεων ακτινικής βάσης
reasoner	:	μηχανή συμπερασματολογίας
recall	:	ανάκληση
regression	:	παλινδρόμηση
relevance feedback	:	συσχετιστική ανάδραση
resizing	:	αλλαγή μεγέθους
reusability	:	επαναχρησιμοποίηση
robustness	:	ευρωστία
rotation	:	περιστροφή
Scalable Color Descriptor	:	κλιμακωτός περιγραφέας χρώματος
scalar cardinality	:	βαθμωτός πληθικός αριθμός, βαθμωτή πληθικότητα
semantic interoperability	:	Σημασιολογική διαλειτουργικότητα
Semantic segmentation	:	Σημασιολογική κατάτμηση
semantics	:	σημασιολογία
sequential analysis	:	ακολουθιακή ανάλυση
sharing	:	διαμοιρασμός
single link	:	τεχνική απλού συνδέσμου
soundness	:	ορθότητα
spatial coherency	:	χωρική συνοχή
Statistical reasoning	:	Στατιστική συμπερασματολογία
subsumption/inclusion axioms	:	αξίωμα υπαγωγής
summarization	:	παρουσίαση συνόψεων
supervised	:	εποπτευμένος με επίβλεψη
syntax	:	συντακτικό
tableaux algorithms	:	αλγόριθμοι πινάκων
template matching	:	ταίριασμα προτύπων
terminological axioms	:	αξιώματα ορολογίας
Terminology Box - TBox	:	Σώμα ορολογίας
testbed	:	πλατφόρμα δοκιμών
thematic categorization	:	θεματική κατηγοριοποίηση
Theory of Mind - ToM	:	Γνωστικές Θεωρίες
thesaurus	:	λεξικό συνωνύμων θησαυρός

time series analysis	:	ανάλυση χρονολογικών σειρών
top-to-down	:	καθοδική ή από πάνω προς τα κάτω
trade off	:	αντιστάθμιση
training data	:	δεδομένα εκπαίδευσης
tree-like	:	δεντρική δομή
unsupervised	:	ανεπίβλεπτος χωρίς επίβλεψη
visual descriptors	:	οπτικοί περιγραφείς
watermarking	:	υδατογράφηση

Κεφάλαιο 1

Εισαγωγή

1.1 Συστήματα μαθησιακής τεχνολογίας και κατηγοριοποίηση περιεχομένου και χρηστών

Οι ραγδαίες αλλαγές σε όλους τους τομείς που σηματοδοτούν τη σύγχρονη εποχή, προβάλλουν επιτακτική την ανάγκη για διαρκή αναβάθμιση των γνώσεων και δεξιοτήτων των πολιτών, προκειμένου αυτοί να ανταποκριθούν στις ολοένα αυξανόμενες απαιτήσεις της προσωπικής και εργασιακής τους ζωής. Η ανάγκη για συνεχιζόμενη εκπαίδευση και επαγγελματική κατάρτιση οδήγησε στη δημιουργία ευέλικτων προγραμμάτων μάθησης σε ένα ευρύτατο φάσμα θεματικών πεδίων. Σε αυτό το πλαίσιο, ο ρόλος της εξ αποστάσεως εκπαίδευσης προσδίδει μια ευελιξία στη δια βίου παροχή εκπαιδευτικών ευκαιριών, περιορίζοντας αποκλεισμούς λόγω γεωγραφικών αποστάσεων και ελευθερίας χρόνου και εξασφαλίζοντας την ενεργό συμμετοχή στη μάθηση.

Η σπουδαιότητα των συστημάτων μαθησιακής τεχνολογίας στην κατεύθυνση της ενεργούς μάθησης άρχισε ήδη να διαφαίνεται με την έλευση του 21ου αιώνα. Τα συστήματα που προτάθηκαν αντιμετώπιζαν με όμοιο τρόπο τους χρήστες τους, χωρίς να λαμβάνουν υπόψη τις ιδιαιτερότητές τους.

Η ανάγκη για προσωποποιημένη πρόσβαση σε ψηφιακό υλικό μέσω συστημάτων μαθησιακής τεχνολογίας οδήγησε την έρευνά μας στη δημιουργία συστημάτων με δυνατότητα συλλογής και επεξεργασίας δεδομένων του χρήστη και με απώτερο σκοπό την εξόρυξη γνώσης από αυτά. Τα δεδομένα αυτά δύναται να αντληθούν από τις ανάγκες, τους στόχους, τις γνώσεις, τα ενδιαφέροντα, τις προτιμήσεις και τη συμπεριφορά του χρήστη του συστήματος.

Το ηλεκτρονικό ερωτηματολόγιο αποτέλεσε το πρώτο βήμα για τη συλλογή των δεδομένων. Για την εξόρυξη γνώσης από τα δεδομένα ερευνήθηκαν κλασικές μέθοδοι από το χώρο της μηχανικής μάθησης και της τεχνητής νοημοσύνης και προτάθηκε μια συνδυαστική μέθοδος για την κατηγοριοποίηση των δεδομένων με παράλληλη μείωση των διαστάσεών τους.

Επιπλέον, προκειμένου να αντληθούν πληροφορίες σχετικές με τη μη λεκτική αλληλεπίδραση μεταξύ χρήστη και συστήματος, καταγράφηκε η εικόνα του χρήστη, με τη βοήθεια μιας κάμερας, καθ' όλη τη διάρκεια αλληλεπίδρασής του με το σύστημα. Για την ανάλυση των εκφράσεων του προσώπου του χρήστη μελετήθηκαν μέθοδοι ανάλυσης χαρακτηριστικών του προσώπου. Τα αποτελέσματα αυτά συνδυάστηκαν με τις γνωστικές θεωρίες [18] για την εξαγωγή της κατάστασης του χρήστη. Η συλλογή και η αξιοποίηση αυτού του συνόλου των δεδομένων του κάθε χρήστη, τόσο μέσω

ηλεκτρονικών ερωτηματολογίων όσο και με εκτίμηση της μη λεκτικής αλληλεπίδρασής του με το σύστημα, οδήγησαν στη δημιουργία του προφίλ του.

Ένα σύστημα ηλεκτρονικής μάθησης για να είναι προσαρμοσμένο στις ανάγκες του χρήστη, θα πρέπει να είναι εφοδιασμένο με τεχνικές κατηγοριοποίησης όχι μόνο του προφίλ του χρήστη, αλλά και του εκπαιδευτικού περιεχομένου που διαθέτει. Η ανάγκη για αυτόματη κατηγοριοποίηση πολυμεσικού περιεχομένου οδήγησε την έρευνα στην περιοχή της ψηφιακής επεξεργασίας εικόνας, έχοντας ως στόχο την εξόρυξη γνώσης ικανής για την αυτόματη κατηγοριοποίηση του. Μελετώντας διάφορα πολυμεσικά εκπαιδευτικά υλικά, διαπιστώθηκε η συχνή εμφάνιση προσώπων μέσα σε εικόνες και ειδικότερα σε πολιτιστικό εκπαιδευτικό υλικό όπως φωτογραφίες με ιστορικά πρόσωπα, έργα τέχνης, κ.α.

Η έρευνα στράφηκε στην αναζήτηση καλλιτεχνικής ιστορικής περιόδου με ιδιαίτερη έμφαση στην απεικόνιση προσώπων και στην εύκολη πρόσβαση σε πλούσιο ψηφιοποιημένο υλικό. Έτσι οδηγηθήκαμε στη χρήση ψηφιοποιημένων έργων ζωγραφικής της Βυζαντινής τέχνης. Η συνεργασία που αναπτύξαμε με ειδικούς ερευνητές της βυζαντινής τέχνης του Ιδρύματος του Όρου Σινά, μας έδειξε πως η αυτόματη τεκμηρίωση του υλικού αυτού, εκτός από την κατηγοριοποίησή του, θα μπορούσε να χρησιμοποιηθεί περαιτέρω και για την εκπαίδευση φοιτητών, μελετητών και ερευνητών, καθώς και φίλων της τέχνης με ενδιαφέρον για αυτήν την ιστορική περίοδο.

Η Βυζαντινή αγιογραφία στηρίζεται πάνω σε αυστηρούς κανόνες σχεδίασης, οι οποίοι μελετήθηκαν και αποδόθηκαν στο σύστημα μαθησιακής τεχνολογίας με τη χρήση οντολογιών. Η προτεινόμενη τεχνική εντοπισμού προσώπου που απεικονίζεται σε Βυζαντινές εικόνες και των χαρακτηριστικών του, αξιοποιώντας τους αυστηρούς κανόνες σχεδίασης που ακολουθούν οι αγιογράφοι, οδήγησε σε εξαγωγή ασαφούς σημασιολογικής γνώσης από την εικόνα.

Η ιδέα της αξιοποίησης της γνώσης που εξάγεται από ανάλυση της Βυζαντινής εικόνας με σκοπό τη μάθηση, μας οδήγησε στην έρευνα για τεκμηρίωση του υλικού. Σε αυτό το σημείο, κρίθηκε απαραίτητη η ενασχόληση με τη περιοχή των τεχνολογιών γνώσης. Η γνώση που εξήχθη από την ανάλυση της εικόνας αποδόθηκε ως σώμα ισχυρισμών στο σύστημα μαθησιακής τεχνολογίας. Το σώμα ισχυρισμών δόθηκε σε μία μηχανή συμπερασματολογίας, εφοδιασμένη με τη γνώση του πεδίου μέσω της οντολογίας του, η οποία αποτελεί το σώμα ορολογίας του συστήματος, με αποτέλεσμα την αυτόματη τεκμηρίωση της Άγιας μορφής που απεικονίζεται στην εικόνα. Η τεκμηρίωση αυτή αποτελεί μεταδεδομένο της Βυζαντινής εικόνας, το οποίο μπορεί να χρησιμοποιηθεί περαιτέρω για εκπαιδευτικούς σκοπούς.

Επιπλέον, μελετήθηκε η χρήση τεχνολογιών γνώσης με σκοπό τη σύνδεση του προφίλ του χρήστη με το εκπαιδευτικό υλικό και τον τρόπο παρουσιάσής του μέσα από το προτεινόμενο σύστημα μαθησιακής τεχνολογίας. Τέλος, η ανάγκη για προστασία του διακινούμενου ψηφιοποιημένου εκπαιδευτικού υλικού στο σύστημα μαθησιακής τεχνολογίας μας οδήγησε στη μελέτη αντίστοιχων τεχνικών. Σε αυτό το πλαίσιο, προτάθηκε μέθοδος προστασίας σημασιολογικού πολυμεσικού υλικού.

Λαμβάνοντας, λοιπόν, υπόψη τα δεδομένα καθώς και τα ζητούμενα ενός συστήματος μαθησιακής τεχνολογίας, έγινε προσπάθεια, στο πλαίσιο της παρούσας διατριβής, να εργαστούμε προς τη συστηματική βελτίωσή του.

1.2 Συνεισφορές και δομή διδακτορικής διατριβής

Οι ερευνητικές συνεισφορές της παρούσας διατριβής σε συνδυασμό με τη δομή της παρουσιάζονται στη συνέχεια :

Κεφάλαιο 1ο: Εισαγωγή

Το παρόν κεφάλαιο αποτελεί την εισαγωγή της διατριβής, όπου παρουσιάζεται ο στόχος και η δομή της.

Κεφάλαιο 2ο: Συστήματα μαθησιακής τεχνολογίας

Σε αυτό το κεφάλαιο αναφέρεται η συνεισφορά των τρεχόντων συστημάτων εξ αποστάσεως εκπαίδευσης στη σύγχρονη κοινωνία μέσα από το πλαίσιο των διαφόρων αρχιτεκτονικών που διεθνείς οργανισμοί έχουν αναπτύξει. Ιδιαίτερη έμφαση δίνεται στην προτεινόμενη από την ΙΕΕΕ αρχιτεκτονική καθώς και στις ελλείψεις που αυτή παρουσιάζει. Έχοντας ως στόχο τη δημιουργία ενός συστήματος ηλεκτρονικής μάθησης που να είναι ικανό να παρέχει προσωποποιημένη πρόσβαση στους χρήστες του καθώς και κατάλληλο εκπαιδευτικό υλικό, παρουσιάζεται η προτεινόμενη επέκταση της αρχιτεκτονικής του ΙΕΕΕ.

Κεφάλαιο 3ο: Εξόρυξη γνώσης και προσωποποιημένη πρόσβαση

Στο κεφάλαιο αυτό παρουσιάζεται η ερευνητική συνεισφορά στο πεδίο της εξόρυξης γνώσης και της προσωποποιημένης πρόσβασης. Για την αντιμετώπιση των ερευνητικών προβλημάτων που αφορούν στην προσωποποιημένη πρόσβαση συνδυάστηκαν τεχνικές συσταδοποίησης και κατηγοριοποίησης δεδομένων παράλληλα με μείωση των διαστάσεών τους. Στη συνέχεια, προτάθηκε μια τεχνική αυτόματης αξιολόγησης της κατάστασης των χρηστών, η οποία βασίζεται στη μη λεκτική αλληλεπίδραση μεταξύ ανθρώπου-μηχανής και όπως αποδείχθη είναι ικανή να συμβάλει στην αυτόματη κατηγοριοποίηση των χρηστών ενός συστήματος μαθησιακής τεχνολογίας.

Κεφάλαιο 4ο: Ανάλυση και τεκμηρίωση πολυμεσικού υλικού

Στο πλαίσιο αυτού του κεφαλαίου, αρχικά εξετάστηκαν τεχνικές ανάλυσης εικόνων και ειδικότερα τεχνικές εντοπισμού προσώπου καθώς και τεχνικές αυτόματης κατάτμησης σημασιολογικών περιοχών εικόνων. Στη συνέχεια, οι τεχνικές που κρίθηκαν καταλληλότερες χρησιμοποιήθηκαν για την εξαγωγή σημασιολογικών περιοχών των εικόνων, κυρίως εικόνων πολιτιστικού περιεχομένου και ειδικότερα Βυζαντινών αγιογραφιών. Ο συνδυασμός της τυποποιημένης γνώσης που χρησιμοποιείται στη σχεδίαση των Βυζαντινών αγιογραφιών με τις τεχνικές που επελέγησαν, οδήγησε, όπως αποδείχθηκε, σε αξιόπιστη τεκμηρίωση των εικόνων.

Κεφάλαιο 5ο: Αναπαράσταση γνώσης και συστήματα εξ αποστάσεως εκπαίδευσης

Σε αυτό το κεφάλαιο, ερευνήθηκαν οι τεχνολογίες γνώσεις και πώς αυτές μπορούν να συνδυαστούν τόσο με τα εξαγόμενα χαρακτηριστικά που προέκυψαν από την ανάλυση του πολυμεσικού υλικού, όσο και με το προφίλ των χρηστών. Σε αυτό το πλαίσιο ερευνήθηκαν οι τυπικές μέθοδοι αναπαράστασης ασαφούς γνώσης και συλλογιστικής και πώς, σε συνδυασμό με μηχανισμούς εξαγωγής συμπερασμάτων, μπορεί να παραχθεί νέα γνώση με μεγάλη συνέπεια.

Κεφάλαιο 6ο: Ασφάλεια πολυμεσικού υλικού

Στο κεφάλαιο αυτό ερευνήθηκαν οι τεχνικές που θα μπορούσαν να χρησιμοποιηθούν για την προστασία του πολυμεσικού υλικού που διακινείται σε ένα περιβάλλον εξ αποστάσεως εκπαίδευσης. Μελετήθηκαν ιδιαίτερα τεχνικές υδατογράφησης πολυμεσικού υλικού και προτάθηκε μια τεχνική για προστασία σημασιολογικών περιοχών, η οποία βασίζεται σε μετασχηματισμό κυματιδίων. Η ευρωστία της προτεινόμενης τεχνικής ελέγχθηκε τόσο σε εικόνες όσο και σε έργα τέχνης.

Κεφάλαιο 7ο: Εφαρμογές προτεινόμενου συστήματος μαθησιακής τεχνολογίας

Σε αυτό το κεφάλαιο παρουσιάζονται τρία διαφορετικά σενάρια εφαρμογής του προτεινόμενου συστήματος μαθησιακής τεχνολογίας. Σε αυτά τα σενάρια παρουσιάζεται πώς η κατηγοριοποίηση των χρηστών και η ανάλυση και τεκμηρίωση του εκπαιδευτικού υλικού μπορεί να οδηγήσει σε ολοκληρωμένα συστήματα μαθησιακής τεχνολογίας, όπου επίκεντρό τους είναι ο εκπαιδευόμενος, δηλαδή ο χρήστης του συστήματος.

Κεφάλαιο 8ο: Συμπεράσματα και κατευθύνσεις μελλοντικής έρευνας

Το τελευταίο κεφάλαιο της διατριβής παρουσιάζει τα συμπεράσματα που προέκυψαν σε όλες τις φάσεις της έρευνας που πραγματοποιήθηκε καθώς και τις προτεινόμενες κατευθύνσεις για μελλοντική έρευνας.

Κεφάλαιο 2

Συστήματα Μαθησιακής Τεχνολογίας

2.1 Εισαγωγή

Στις μέρες μας, είναι πιο έντονη από ποτέ, η ανάγκη για δημιουργία ενός ανθρωποκεντρικού μοντέλου ανάπτυξης στο πλαίσιο της Κοινωνίας της Πληροφορίας. Το μοντέλο αυτό θα πρέπει να προσφέρει στον άνθρωπο ένα περιβάλλον αυξημένων ευκαιριών και δυνατοτήτων που θα αξιοποιεί τις νέες τεχνολογίες πληροφορικής και επικοινωνιών, βελτιώνοντας παράλληλα την ποιότητα ζωής και την παραγωγικότητά του. Σε αυτό το πλαίσιο έχουν προταθεί πολλά συστήματα μαθησιακής τεχνολογίας που όμως υστερούν ως προς τον ανθρωποκεντρικό χαρακτήρα. Λαμβάνοντας υπόψη τις παραπάνω διαπιστώσεις, εξετάστηκε η δυνατότητα ανάπτυξη συστήματος ηλεκτρονικής μάθησης με κέντρο τον άνθρωπο.

Σε αυτό το κεφάλαιο αρχικά παρουσιάζεται ο ρόλος της εξ αποστάσεως εκπαίδευσης. Έπειτα παρουσιάζονται διάφορες αρχιτεκτονικές που διεθνείς οργανισμοί έχουν προτείνει και αναπτύξει με σκοπό τη δημιουργία ολοκληρωμένων συστημάτων μαθησιακής τεχνολογίας, ενώ δίνεται έμφαση στην προτεινόμενη αρχιτεκτονική του IEEE. Τέλος, παρουσιάζεται η προσπάθειά μας για την επέκταση της αρχιτεκτονικής του IEEE συμπεριλαμβάνοντας νέα στοιχεία και διαδικασίες, ικανές να προσδώσουν τη δυναμική της περιβάλλουσας νοημοσύνης (ambient intelligence) στο υπάρχον αρχιτεκτονικό σχήμα του IEEE.

2.2 Εξ αποστάσεως εκπαίδευση

Η σημερινή εξέλιξη της τεχνολογίας δεν προσθέτει απλώς κάτι καινούργιο, τα αλλάζει όλα. Είναι, πλέον, ένα τεχνολογικό εργαλείο που διαμορφώνει ένα διαφορετικό κοινωνικο-πολιτισμικό περιβάλλον και αλλάζει τον τρόπο δράσης και σκέψης του ανθρώπου. Οι νέες τεχνολογίες απαιτούν από το σύγχρονο άνθρωπο να βρίσκεται σε διαρκή ενημέρωση. Η παγκοσμιοποίηση και η δημιουργία νέων συνθηκών και διαδικασιών σε όλες τις πτυχές της ανθρώπινης ζωής υπαγορεύουν την ανάγκη δια βίου μάθησης. Η δια βίου μάθηση παρέχει γνώσεις και δυνατότητες στο σύγχρονο άνθρωπο, ώστε να μπορεί να βρίσκει τρόπους έκφρασής του μέσα στη σύγχρονη κοινωνία καθώς και να αναδεικνύει τις δεξιότητές του.

Τα εκπαιδευτικά συστήματα, σύμφωνα με τις σύγχρονες απαιτήσεις, δεν περιορίζονται σε καθορισμένα χρονικά πλαίσια αλλά ούτε και σε συγκεκριμένους χώρους. Οπότε δημιουργείται η ανάγκη επαναπροσδιορισμού της διαδικασίας μάθησης σε νέα περιβάλλοντα. Η ταχύτατη εξέλιξη της τεχνολογίας οδηγεί στο σχεδιασμό

ηλεκτρονικών συστημάτων μάθησης τα οποία μπορούν και παρέχουν στους χρήστες τους τη δυνατότητα δια βίου εκπαίδευσης. Τα νέα αυτά περιβάλλοντα μπορούν να χρησιμοποιούν είτε παραδοσιακά μέσα εκμάθησης είτε σύγχρονα, όπως η μάθηση εξ αποστάσεως. Η εξ αποστάσεως εκπαίδευση παρέχει ένα σύνολο από εργαλεία τα οποία είναι ικανά να ενισχύσουν θετικά τις υπάρχουσες τεχνικές μάθησης χρησιμοποιώντας νέες τεχνολογίες.

Η χρήση των υπολογιστών στην εκπαίδευση ξεκίνησε ήδη με την εμφάνισή τους. Όμως, μέχρι σήμερα, ήταν κυρίως προσανατολισμένη σε απομονωμένες, αυτόνομες εφαρμογές, συνήθως ηλεκτρονικά ανάλογα βιβλίων ή εφαρμογές αυτοαξιολόγησης και εικονικών πειραμάτων πάντα αναφερόμενη σε ένα συγκεκριμένο μαθησιακό αντικείμενο. Με την έλευση των νέων Τεχνολογιών Πληροφορικής και Επικοινωνιών, ΤΠΕ, (Information and Communication Technologies, ICT), με κυρίαρχα στοιχεία τα δίκτυα υπολογιστών και τα υπερμέσα, εξαπλώθηκε ταχύτατα η χρήση των προσωπικών υπολογιστών και των δικτύων υπολογιστών, ειδικότερα του Διαδικτύου και του Παγκόσμιου Ιστού ανοίγοντας νέες προοπτικές για την εκπαίδευση [56]. Για το διαχωρισμό αυτών των τεχνολογιών χρησιμοποιείται διεθνώς ο όρος Προηγμένες Μαθησιακές Τεχνολογίες (Advanced Learning Technologies). Το άμεσο αποτέλεσμα κάθε τεχνολογίας είναι τα προϊόντα έτσι και στη περίπτωση των προηγμένων μαθησιακών τεχνολογιών προϊόντα είναι τα Συστήματα Μαθησιακής Τεχνολογίας (Learning Technology Systems) δηλαδή όλα εκείνα τα λογισμικά συστήματα που βασικό τους αντικείμενο είναι η υλοποίηση υπηρεσιών στο πλαίσιο ενός μαθησιακού περιβάλλοντος.

Οι ευεπίδωτες προοπτικές, που φαίνεται ότι έχουν οι Μαθησιακές Τεχνολογίες για την επίλυση του σημερινού εκπαιδευτικού προβλήματος, έχουν αυξήσει το ενδιαφέρον για τη χρήση τους από εκπαιδευτικούς οργανισμούς. Πανεπιστήμια, σχολεία, εταιρείες και άλλοι οργανισμοί, αντιμετωπίζουν σήμερα την πρόκληση να αξιοποιήσουν αυτές τις τεχνολογίες για να επεκτείνουν τα μαθησιακά τους περιβάλλοντα με την ανάπτυξη Συστημάτων Μαθησιακής Τεχνολογίας (ΣΜΤ) και να χρησιμοποιήσουν νέες μορφές μαθησιακών πόρων [56].

Δυστυχώς τα μέχρι τώρα αποτελέσματα από την ανάπτυξη μαθησιακών συστημάτων και πόρων με χρήση των νέων αυτών τεχνολογιών, είναι μάλλον απογοητευτικά. Η αξιοποίηση του δυναμικού αυτών των τεχνολογιών δεν είναι εύκολη υπόθεση. Όσο και αν η τεχνολογία κεντρίζει το ενδιαφέρον και εμπνέει τη φαντασία των ανά τον κόσμο τεχνολόγων, δεν πρόκειται από μόνη της να μας δώσει αυτόματα τα αναμενόμενα αποτελέσματα. Η ιστορία της τεχνολογίας έχει να επιδείξει αρκετά απογοητευτικά παραδείγματα δυνητικών και πολύ αναμενόμενων "επαναστάσεων", οι οποίες όμως δεν πραγματοποιήθηκαν ποτέ. Στην πλειοψηφία αυτών των περιπτώσεων, η έμφαση στα μέσα επισκίασε τον πραγματικό σκοπό. Η έλλειψη καθιερωμένων προσεγγίσεων τόσο για την αρχιτεκτονική περιγραφή ΣΜΤ όσο και για τη δημιουργία, περιγραφή και διανομή μαθησιακών πόρων θέτει εμπόδια στην αποδοτική και αποτελεσματική αξιοποίηση των δυνατοτήτων που προσφέρουν οι νέες τεχνολογίες στην εκπαίδευση και την κατάρτιση [56].

Από την άλλη, το διαδίκτυο χρησιμοποιείται, σε ολοένα αυξανόμενο βαθμό, ως εργαλείο παροχής ανοικτής και εξ αποστάσεως εκπαίδευσης. Οι λόγοι είναι πολλοί: είναι εύκολα προσβάσιμο, υποστηρίζει πολλαπλές επιλογές αποθήκευσης και απεικόνισης, παρέχει μία απλή αλλά συγχρόνως ισχυρή φόρμα δημοσίευσης και τρόπους ώστε να ενσωματώνονται πολλαπλά μέσα [2]. Το διαδίκτυο επιπλέον ενισχύει την ενεργή μάθηση αφού μπορεί να θεωρηθεί ως ένα περιβάλλον δημιουργικότητας που

υποστηρίζει και ενθαρρύνει συμπεριφορές για εξερεύνηση και αναζήτηση. Σε αυτό το πλαίσιο έχει ενισχυθεί η δημιουργία ολοκληρωμένων συστημάτων εξ αποστάσεως εκπαίδευσης. Θεωρείται πλέον επιτακτική η υλοποίηση νέων δομών εκμάθησης με σκοπό την ενεργή συμμετοχή όλων των πολιτών στις δυνατότητες που προσφέρει η κοινωνία της πληροφορίας με σκοπό όχι μόνο το μη αποκλεισμό τους από αυτές αλλά και την εύκολη προσαρμογή τους.

Η εξ αποστάσεως εκπαίδευση μπορεί να θεωρηθεί επαναστατικός τρόπος παροχής δεξιοτήτων και γνώσης για το σύγχρονο πολίτη. Μια τέτοιου είδους εκπαίδευση επιταχύνει τη διαδικασία εκμάθησης, μειώνει το κόστος μάθησης και εξασφαλίζει την πρόσβαση στην Κοινωνία της Πληροφορίας, γεγονός που οδηγεί στη θέσπιση διεθνών προτύπων προκειμένου να καταστεί δυνατή η υλοποίηση δομών εξ αποστάσεως εκπαίδευσης. Στο χώρο της προτυποποίησης της μαθησιακής τεχνολογίας υπάρχουν αρκετοί οργανισμοί και ομάδες που ασχολούνται με διάφορες πτυχές της μαθησιακής τεχνολογίας. Ο συνδυασμός των αποτελεσμάτων αυτών των ομάδων έχει οδηγήσει στη δημιουργία προδιαγραφών για Συστήματα Μαθησιακής Τεχνολογίας (Learning Technology System) [14] καθώς και στη δημιουργία προτύπων για Σύστημα Μαθησιακής Τεχνολογίας.

2.3 Προτυποποίηση μαθησιακών τεχνολογιών

Είναι γεγονός ότι η σύγχρονη κοινωνία κινείται πλέον στους ρυθμούς της κοινωνίας της πληροφορίας. Η βέλτιστη χρήση της τεχνολογία με κατεύθυνση τη μάθηση αποτελεί ίσως τη βασικότερη προσδοκία μας: τη δημιουργία μιας κοινωνίας χωρίς αποκλεισμό των μελών της. Η ένσταση που προβάλλουν πολλοί από τους επικριτές των νέων τεχνολογιών έγκειται στο ότι οι τεχνολογίες αυτές μέχρι στιγμής έχουν εφαρμοστεί αποσπασματικά είτε χωρίς να είναι εμφανή τα οφέλη τους ή έχοντας αποτελέσματα αμφίβολης ποιότητας. Η διεξοδική μελέτη και ανάπτυξη συγκεκριμένων μοντέλων και μεθοδολογιών χρήσης των νέων τεχνολογιών και η προτυποποίηση των μαθησιακών τεχνολογιών σίγουρα θα συμβάλλουν στη λύση αυτού του προβλήματος.

Στη συνέχεια παρουσιάζονται διάφορα διεθνή πρότυπα και προδιαγραφές που έχουν προταθεί καθώς και μια επισκόπηση συστημάτων μαθησιακής τεχνολογίας.

Η ανάπτυξη διεθνών προτύπων στο χώρο των Μαθησιακών Τεχνολογιών με στόχο την εξασφάλιση διαλειτουργικότητας μεταξύ των διαφορετικών ΣΜΤ και τη δυνατότητα διαμοιρασμού, επαναχρησιμοποίησης και ολοκλήρωσης μονάδων μαθησιακού υλικού είναι μείζονος σημασίας και μπορεί να αποτελέσει τη βάση για τη δημιουργία ευέλικτων και αποδοτικών λύσεων [56].

Θέλοντας η ευρωπαϊκή ένωση να προωθήσει την έρευνα γύρω από συστήματα μαθησιακής τεχνολογίας εξέδωσε μια ανακοίνωση το 1999 προς τους κυριότερους επίσημους φορείς προτυποποίησης στο χώρο των τεχνολογιών της πληροφορικής και των τηλεπικοινωνιών αναφέροντας ότι ο σημαντικότερος ανασταλτικός παράγοντας για μια ευρεία διάδοση των μαθησιακών τεχνολογιών είναι η έλλειψη προτύπων με ταυτόχρονη ικανοποίηση της απαίτησης του εκπαιδευτικού χώρου για περιορισμό στο κόστος. Γιαυτό το λόγο η Ευρωπαϊκή Ένωση απαίτησε τη δημιουργία ενός ευρύ φάσματος από πρότυπα που θα περιλαμβάνει όλα όσα σχετίζονται με τις πλατφόρμες των συστημάτων που θα χρησιμοποιηθούν. Αν και είναι νωρίς να μιλάμε για περιεκτικά πρότυπα που θα καλύπτουν όλο αυτό το φάσμα, πρότυπα που στοχεύουν συγκεκριμένες περιοχές μπορούν να αποδειχθούν εξαιρετικά πολύτιμα, καθώς είναι αναγκαία προϋπόθεση για την υλοποίηση νέων υπηρεσιών σε πανευρωπαϊκό επίπεδο

για τη δια βίου μάθηση.

Σε αυτό το πλαίσιο κινήθηκαν διάφοροι φορείς και συγκεκριμένα το 1999 ο φορέας προτυποποίησης ISO/IEC JTC1 ίδρυσε την υποεπιτροπή SC36 για την αντιμετώπιση των αναγκών προτυποποίησης στο χώρο των μαθησιακών τεχνολογιών. Σκοπός της επιτροπής αυτής ήταν η προτυποποίηση στο χώρο των τεχνολογιών της πληροφορικής για τη μάθηση, την εκπαίδευση και την κατάρτιση με σκοπό την υποστήριξη ατόμων, ομάδων και οργανισμών και την επίτευξη διαλειτουργικότητας και επαναχρησιμοποίησης των πόρων και των εργαλείων. Δικαίωμα συμμετοχής, ως μέλη στο φορέα αυτό, είχαν οι εθνικοί φορείς προτυποποίησης. Η SC36 συνεργάζεται στενά με αρκετούς οργανισμούς ανάπτυξης προδιαγραφών που αναπτύσσουν δράση τόσο στον αμερικάνικο όσο και στον ευρωπαϊκό χώρο. Οι ομάδες αυτές εργάστηκαν και εργάζονται με σκοπό την εισήγηση και ανάπτυξη τεχνικών συστάσεων και προδιαγραφών που άπτονται του συγκεκριμένου πεδίου. Στόχος είναι οι προδιαγραφές που φτάνουν σε ένα συγκεκριμένο επίπεδο ωρίμανσης να υποβάλλονται για έγκριση στην JTC1/SC36. Κλειδί για την έγκριση μιας προδιαγραφής είναι η ευρεία αποδοχή από την κοινότητα των χρηστών και τη βιομηχανία.

Οι κυριότεροι από τους οργανισμούς αυτούς είναι οι [56]:

- IEEE LTSC (Learning Technology Standards Committee): Η επιτροπή αυτή έχει περίπου 20 υπο-επιτροπές που ασχολούνται με την ανάπτυξη προδιαγραφών προτυποποίησης στο χώρο της Μαθησιακής Τεχνολογίας. Η LTSC καταθέτει στην SC36 προδιαγραφές που έχουν φτάσει σε ένα καλό επίπεδο ωριμότητας (<http://ltsc.ieee.org>), πολλές από τις οποίες έχουν προτυποποιηθεί.
- ARIADNE (Alliance of Remote Instructional Authoring and Distribution Networks for Europe): πρόκειται για ένα ευρωπαϊκό ίδρυμα που αναπτύσσει εννοιολογικά πλαίσια και εργαλεία για εξ αποστάσεως διδασκαλία και μάθηση με χρήση των νέων τεχνολογιών της πληροφορικής και των επικοινωνιών και με ιδιαίτερη έμφαση στο διαμοιρασμό και την επαναχρησιμοποίηση ηλεκτρονικού μαθησιακού υλικού. Το ίδρυμα συμμετέχει σε διεθνείς δραστηριότητες προτυποποίησης, με έμφαση στα μετα-δεδομένα (metadata), και έχει αναπτύξει αξιοσημείωτη λειτουργική υποδομή. Επιπλέον, διατηρεί στενή συνεργασία με την ομάδα εργασίας Learning Object Metadata της IEEE LTSC (<http://www.ariadne-eu.org>).
- IMS Global Learning Consortium: ο οργανισμός αυτός ιδρύθηκε στις Ηνωμένες Πολιτείες το 1997 με μέλη τόσο από τον ακαδημαϊκό όσο και από τον επιχειρηματικό χώρο. Σύντομα διέυρνε την εμβέλεια δραστηριότητάς του σε διεθνές επίπεδο και αυτή τη στιγμή πάνω από 200 μέλη του 'δικτύου ανάπτυξης' επιθεωρούν και χρησιμοποιούν τις προδιαγραφές που αναπτύσσει. Οι προδιαγραφές αυτές στοχεύουν στην επίτευξη διαλειτουργικότητας των συστημάτων που υποστηρίζουν τη συνδεδεμένη (online) μάθηση, εκπαίδευση και κατάρτιση. Πολλοί από τους συμμετέχοντες στον IMS συμμετέχουν επίσης σε εθνικούς φορείς, την IEEE LTSC, τον AICC, κ.α. (<http://imglobal.org>).
- DCMi (Dublin Core Metadata Initiative): η ομάδα εργασίας DC-Education ιδρύθηκε το 1999 με σκοπό την περιγραφή των εκπαιδευτικών πόρων για την επίτευξη διαλειτουργικότητας στο πλαίσιο του DCMi. Η επιτροπή SC36 συνεργάζεται με τον φορέα αυτό (<http://dublincore.org/groups/education>).

- ADL (Advanced Distributed Learning): ο οργανισμός αυτός ιδρύθηκε το 1997 με πρωτοβουλία του αμερικάνικου υπουργείου άμυνας για την ανάπτυξη μιας στρατηγικής σχετικά με τη χρήση των μαθησιακών τεχνολογιών. Σκοπός του είναι ο εκσυγχρονισμός της εκπαίδευσης και της κατάρτισης και η προώθηση της συνεργασίας μεταξύ της κυβέρνησης, της βιομηχανίας και του ακαδημαϊκού χώρου για την ανάπτυξη μαθησιακών προτύπων (<http://www.adlnet.org>).

Στη συνέχεια παρουσιάζονται οι σημαντικότεροι τομείς για τους οποίους έχουν προταθεί, από διάφορους οργανισμούς, ιδέες που έχουν προτυποποιηθεί [56]:

- Αρχιτεκτονική Συστημάτων Μαθησιακής Τεχνολογίας (Learning Technology Systems Architecture)

Ο σκοπός της αρχιτεκτονικής ενός τέτοιου συστήματος είναι η σχεδίαση και η υλοποίηση υποσυστημάτων τα οποία μπορούν να αλληλεπιδρούν και να ανταλλάσσουν πληροφορίες ανεξάρτητα της πλατφόρμας στην οποία τρέχουν. Προς το παρόν υπάρχουν οι εξής προσεγγίσεις: η IEEE LTSC έχει αναπτύξει το πρότυπο Learning Technology Systems Architecture [86] στο οποίο έχουν καθοριστεί τα υποσυστήματα των συστημάτων μαθησιακής τεχνολογίας καθώς και οι τρόποι αλληλεπίδρασης μεταξύ τους. Η επιτροπή ISO/IEC-JTC1-SC36 μετά από συνεργασία με την IEEE LTSC καθόρισαν ένα γενικό σύστημα μαθησιακής τεχνολογίας.

- Γλωσσάρι (Glossary)

Ασχολείται με την απαρίθμηση και ορισμό των όρων που χρησιμοποιούνται στα πρότυπα με σκοπό να επιτρέψει στους συγγραφείς των προτύπων να χρησιμοποιούν προκαθορισμένες ορολογίες για τη συγγραφή των προτύπων και να επιτρέψει στους χρήστες να κατανοούν το νόημα των όρων που υπάρχουν στα πρότυπα. Ένα γλωσσάρι για ΣΜΤ υπό μορφή προτύπου έχει καθορισθεί από την IEEE LTSC και την ISO/IEC JTC1 SC36.

- Μεταδεδομένα (Metadata)

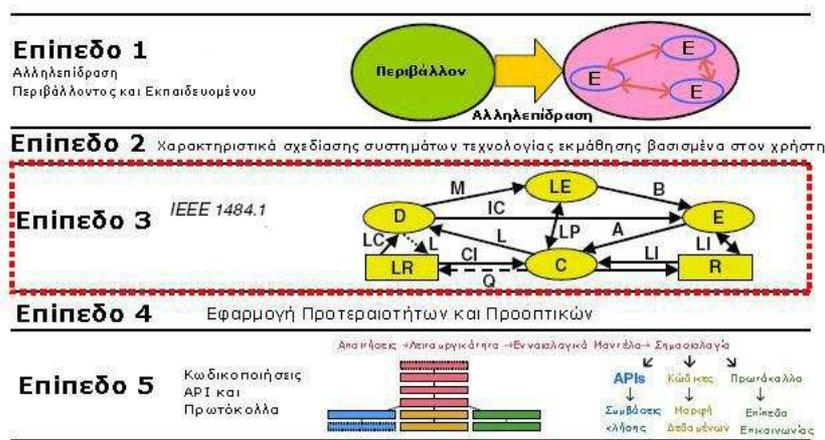
Είναι εμφανής η ανάγκη για δυνατότητα αναζήτησης, διαμοιρασμού και χρήσης μαθησιακού υλικού. Δεδομένου ότι οι μαθησιακοί χώροι στο πλαίσιο της προηγμένης μαθησιακής τεχνολογίας αυξάνονται εκθετικά, θα πρέπει να υπάρχει ένα σύνολο από χαρακτηριστικά γνωρίσματα το οποίο θα απαριθμεί και θα ορίζει τους όρους που χρησιμοποιούνται στα πρότυπα και που απαιτούνται για την πλήρη και ακριβή περιγραφή μαθησιακών αντικειμένων. Οι διάφορες προσεγγίσεις των επιτροπών πάνω σε αυτό το τομέα είναι τα IEEE LTSC Standard for Learning Object Metadata, IMS Global Learning Consortium Metadata Specification, έκδοση του Dublin Core Metadata Initiative που αφορά στα μαθησιακά αντικείμενα.

- Προφίλ Εκπαιδευομένων (Learner Profile)

Ένας εκπαιδευόμενος μπορεί να καταρτίζεται από διαφορετικούς οργανισμούς. Άρα θα ήταν επιθυμητό να μπορούν τα στοιχεία των εκπαιδευομένων όπως προσωπικά στοιχεία, ιστορικό εκπαίδευσης τους, μαθησιακές προτιμήσεις, προσόντα, τυχόν αξιολογήσεις κ.τ.λ. να είναι προσπελάσιμα από διαφορετικά συστήματα. Οι προσεγγίσεις που έχουν γίνει στον τομέα αυτό από τις αντίστοιχες επιτροπές είναι το πρότυπο Personal and Private Information (PAPI) της IEEE LTSC και το πρότυπο Learner Information Package (LIP) του οργανισμού IMS.

2.4 Επισκόπηση αρχιτεκτονικής του συστήματος του IEEE για μαθησιακή τεχνολογία

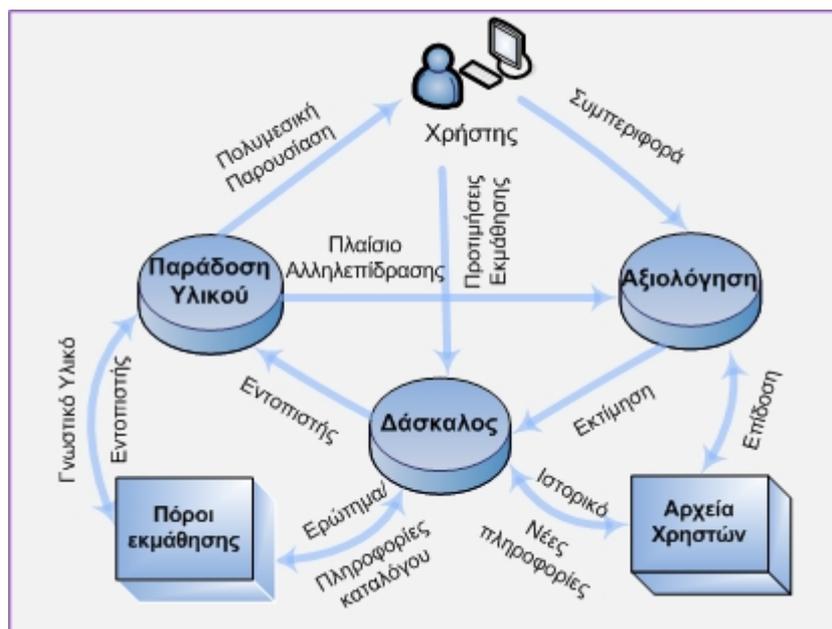
Η αρχιτεκτονική του συστήματος του IEEE για μαθησιακή τεχνολογία (Learning Technology Systems Architecture) είναι ένα υψηλού επιπέδου πλαίσιο για ένα αρχιτεκτονικό σχέδιο μιας σειράς συστημάτων τεχνολογίας εκμάθησης, συμπεριλαμβανομένης της εκπαίδευσης της κατάρτισης, της βασισμένης σε υπολογιστή μάθησης και της ευφυούς μάθησης. Η προτεινόμενη αρχιτεκτονική αποτελείται από πέντε επίπεδα στα οποία καθορίζονται τόσο οι ανθρώπινες δραστηριότητες όσο και οι υπολογιστικές διαδικασίες. Το πρότυπο στοχεύει σε μια γενικευμένη οπτική για τα Συστήματα Μαθησιακής Τεχνολογίας αφήνοντας έτσι χώρο για την ανάπτυξη παιδαγωγικών και πολιτισμικών εννοιών.



Σχήμα 2.1: Επίπεδα αρχιτεκτονικής συστήματος μαθησιακής τεχνολογίας

Τα επίπεδα που υπάρχουν στην αρχιτεκτονική απεικονίζονται στο Σχήμα 2.1 και είναι τα εξής πέντε:

- **Επίπεδο 1 - Αλληλεπίδραση περιβάλλοντος και εκπαιδευόμενου (Learner and Environment Interactions):**
Αφορά στην απόκτηση, μεταφορά, ανταλλαγή της γνώσης του συστήματος με τον εκπαιδευόμενο μέσω της αλληλεπίδρασης με το περιβάλλον μάθησης.
- **Επίπεδο 2 - Χαρακτηριστικά σχεδίασης συστημάτων τεχνολογίας εκμάθησης σχετιζόμενα με τον εκπαιδευόμενο (Learner-Related Design Features):**
Αφορά στην υλοποίηση σχεδίων των συστημάτων τεχνολογίας εκμάθησης που σχετίζονται με τη συμπεριφορά και το γνωστικό επίπεδο των εκπαιδευομένων.
- **Επίπεδο 3 - Στοιχεία συστήματος (System Components):**
Περιγράφει τα συστατικά μέρη του συστήματος μαθησιακής τεχνολογίας με κέντρο τον εκπαιδευόμενο.
- **Επίπεδο 4 - Εφαρμογή προτεραιοτήτων και προοπτικών (Implementation Perspectives and Priorities):**
Περιγράφει τα συστήματα τεχνολογίας εκμάθησης και τις ποικίλες προοπτικές και προτεραιότητες που μπορούν οι εκπαιδευόμενοι να θέσουν στο σύστημα.



Σχήμα 2.2: Σύστημα Μαθησιακής Τεχνολογίας του IEEE

- Επίπεδο 5 - Λειτουργικά συστατικά και διαλειτουργικότητα (Operational Components and Interoperability - codings, APIs, protocols): Περιγράφει τις διεπαφές (interfaces) της προτεινόμενης αρχιτεκτονικής με την αναπτυξιακή διαδικασία και προτείνει τρόπους και τεχνικές υλοποίησης του προτύπου.

Αυτά τα πέντε επίπεδα παρουσιάζουν πέντε ανεξάρτητες περιοχές τεχνικής ανάλυσης, ενώ το μόνο επίπεδο το οποίο μπορεί σαφώς να προτυποποιηθεί είναι το Επίπεδο 3, ο καθορισμός του οποίου απαιτείται για τη δημιουργία ενός Συστήματος Μαθησιακής Τεχνολογίας. Τα στοιχεία που απαρτίζουν το Σύστημα Μαθησιακής Τεχνολογίας του IEEE (επίπεδο 5), όπως απεικονίζεται στο Σχήμα 2.2, μπορούν να κατηγοριοποιηθούν ως εξής:

- **Διαδικασίες** (Χρήστης, Δάσκαλος, Παράδοση Υλικού, Αξιολόγηση)
Αποτελούν τις εισόδους και τις εξόδους του Συστήματος Μαθησιακής Τεχνολογίας.
- **Αποθήκες** (Αρχεία Χρηστών Πόροι εκμάθησης)
Περιγράφουν τον τύπο των δεδομένων που αποθηκεύονται καθώς και τις μεθόδους αναζήτησης, ανάκτησης και αναβάθμισης του υλικού.
- **Ροές** (Πολυμεσικό Υλικό, Προτιμήσεις Εκμάθησης, Πλαίσιο Αλληλεπίδρασης, Συμπεριφορά, Εκτίμηση, Ιστορικό, Τρέχουσες Πληροφορίες Εκπαιδευομένων, Νέες Πληροφορίες Εκπαιδευομένων, Εντοπιστής, Πληροφορίες Καταλόγου, Ερώτημα)
Περιγράφουν το είδος της επικοινωνίας (μονόδρομη, αμφίδρομη, στατική, δυναμική επικοινωνία) μεταξύ των Διαδικασιών και των Αποθηκών καθώς και τις πληροφορίες που διαβιβάζονται.

Η κύρια διαδικασία του Συστήματος Μαθησιακής Τεχνολογίας αποτελείται από την οντότητα Χρήστης, η οποία εκπροσωπεί είτε έναν απλό χρήστη είτε ένα σύνολο

χρηστών. Η διαδικασία *Χρήστης* λαμβάνει ως είσοδο μία πολυμεσική παρουσίαση από τη διαδικασία *Παράδοση Υλικού* ενώ εξάγει παρατηρήσεις σχετικές με τη δραστηριότητα του χρήστη καθόλη τη διάρκεια παρουσίασης του πολυμεσικού υλικού. Αυτές οι παρατηρήσεις μπορεί να περιέχουν πληροφορίες σχετικές με το μέσο χρόνο απόκρισης του χρήστη στο υλικό της παρουσίασης και με τη γενική του εικόνα κατά τη διάρκεια της παρουσίασης. Οι παρατηρήσεις αυτές διοχετεύονται στη διαδικασία *Αξιολόγηση* διαμέσου της ροής *Συμπεριφοράς*.

Η ροή *Πλαίσιο Αλληλεπίδρασης* η οποία έχει μονόδρομη κατεύθυνση (από τη διαδικασία παράδοσης υλικού προς τη διαδικασία αξιολόγησης), παρέχει το απαραίτητο πλαίσιο στη διαδικασία *Αξιολόγηση* για την ερμηνεία της συμπεριφοράς του χρήστη.

Έτσι η διαδικασία *Αξιολόγηση* δέχεται ως κύρια είσοδο τη συμπεριφορά του χρήστη και ως βοηθητική είσοδο το πλαίσιο αλληλεπίδρασης. Τα αποτελέσματα που εξάγονται από τη διαδικασία *Αξιολόγηση* προωθούνται μέσω της ροής *Επίδοση Χρήστη* προς την αποθήκη *Αρχεία Χρηστών*. Η διαδικασία *Αξιολόγηση* και η αποθήκη *Αρχεία Χρηστών* συνδέονται αμφίδρομα. Η διαδικασία *Αξιολόγηση* μπορεί, οποιαδήποτε στιγμή, να ανακτήσει πληροφορίες, μέσω της ροής *Επίδοση* για το χρήστη από την αποθήκη *Αρχεία Χρηστών*, η οποία υλοποιείται με μία βάση δεδομένων.

Η διαδικασία *Δάσκαλος* μπορεί οποιαδήποτε στιγμή να λάβει πληροφορίες σχετικά με την απόδοση του εκπαιδευόμενου από την αποθήκη *Αρχεία Χρηστών* μέσω της ροής *Ιστορικό*. Ο *Δάσκαλος* επίσης μπορεί να παρέχει στοιχεία για την απόδοση, πιστοποιήσεις και τις προτιμήσεις του χρήστη, πληροφορίες που μπορούν να αποθηκευτούν στο *Αρχεία Χρηστών*.

Το αποτέλεσμα της διαδικασίας *Αξιολόγηση* του χρήστη, το οποίο και εξάγεται από την αντίστοιχη διαδικασία, αποστέλλεται στη διαδικασία *Δάσκαλος* μέσω της ροής *Εκτίμηση*. Με αυτόν το τρόπο παρέχεται στη διαδικασία *Δάσκαλος* η άμεση δυνατότητα επιλογής του κατάλληλου εκπαιδευτικού υλικού για τον εκπαιδευόμενο μέσα από την αποθήκη των *Πόρων Εκμάθησης*.

Η διαδικασία *Χρήστης* μπορεί να λάβει ή να στείλει τις προτιμήσεις εκμάθησης με τη διαδικασία *Δάσκαλος* μέσω της ροής *Προτιμήσεις Εκμάθησης*. Με τον τρόπο αυτό η μάθησης μπορεί να καθοριστεί από τον εκπαιδευόμενο, από το δάσκαλο ή κατόπιν συνεργασίας τους. Οι προτιμήσεις εκμάθησης μπορεί να διαμορφώνονται από τη γλώσσα και την πολιτισμική προσαρμογή, καθώς επίσης και από τις ειδικές ανάγκες των χρηστών.

Η διαδικασία *Δάσκαλος* δέχεται ως εισόδους την εκτίμηση που εξάγεται από τη διαδικασία *Αξιολόγηση*, τις *Προτιμήσεις Εκμάθησης* από τη διαδικασία *Χρήστης* καθώς και το *Ιστορικό* του χρήστη το οποίο εξάγεται από την αποθήκη *Αρχεία Χρηστών*.

Η αποθήκη *Πόροι Εκμάθησης* είναι μια βάση δεδομένων που αντιπροσωπεύει τη γνώση, τις πληροφορίες καθώς και άλλους πόρους που χρησιμοποιούνται με σκοπό την απόκτηση γνώσης. Μπορεί να θεωρηθεί ως συλλογή των ερωτηματολογίων, των τεστ αξιολόγησης, των πειραμάτων, των παρουσιάσεων και των μαθημάτων. Οι *Πόροι Εκμάθησης* προωθούν μετά από ερώτημα του δασκάλου (που μεταβιβάζεται μέσω της ροής *Ερώτημα*) τις πληροφορίες καταλόγων, δηλαδή έναν κατάλογο με δείκτες εντοπισμού του υλικού που αναζητήθηκε βάσει του ερωτήματος. Αυτός ο κατάλογος των δεικτών στέλνεται στη διαδικασία *Παράδοση* μέσω της ροής *Πληροφορίες Καταλόγου*. Τέλος, η διαδικασία *Παράδοση* αναλαμβάνει να ανακτήσει το υλικό από τη βάση των πόρων εκμάθησης και να το μεταφέρει προς την οντότητα

Χρήστης μέσω της ροής Πολυμεσική Παρουσίαση.

2.5 Προτεινόμενη επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE

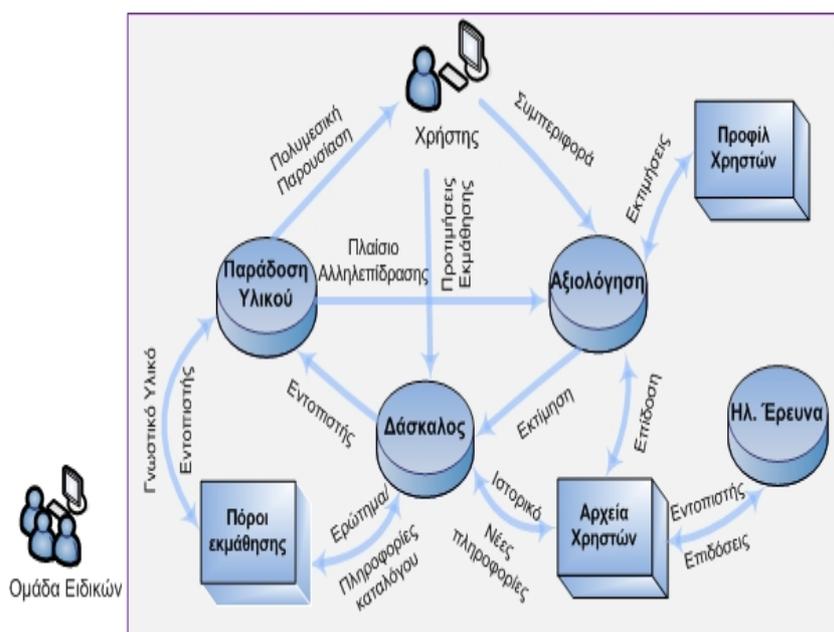
Το προτεινόμενο Σύστημα Μαθησιακής Τεχνολογίας (ΣΜΤ) (Σχήμα 2.3) ενισχύει τη γενική περιγραφή του IEEE LTSA Draft Standard [6], συνεισφέροντας στην κατηγοριοποίηση τόσο των χρηστών του όσο και του υλικού με χρήση τεχνολογιών εξόρυξης δεδομένων, ανάλυσης εικόνας και μηχανικής μάθησης. Η συνεισφορά μας επεκτείνεται και στην ασφαλή μετάδοση των δεδομένων και ειδικότερα των ψηφιακών εικόνων και βίντεο, προς τους χρήστες του προτεινόμενου συστήματος ΣΜΤ. Όλες οι διαδικασίες του προτεινόμενου, από την IEEE, συστήματος έχουν εμπλουτιστεί με νέες λειτουργίες οι οποίες περιγράφονται λεπτομερώς στη συνέχεια.

Το προτεινόμενο σύστημα θεωρείται ένα ολοκληρωμένο σύστημα ηλεκτρονικής μάθησης που η βάση του στηρίζεται στη θεωρία περί εποικοδόμησης της σκέψης (constructivism) [9]. Ένας από τους κυριότερους παράγοντες που συνέβαλε στην προώθηση της θεωρίας περί εποικοδόμησης της σκέψης ήταν η υπόθεση ότι από τη στιγμή που η μάθηση θεωρείται μια μετάδοση πληροφοριών χάνει το νόημά της. Η μάθηση αρχίζει σταδιακά να θεωρείται ότι βασίζεται στην ενεργό διερεύνηση και στην προσωπική δημιουργία μάλλον, παρά σε ένα μοντέλο απλής μετάδοσης. Σύμφωνα με αυτή τη θεωρία, οι μαθητές μαθαίνουν όταν εποικοδομούν ενεργητικά τις νέες γνώσεις στις υπάρχουσες γνώσεις τους [1]. Βασική διαδικασία σε αυτή την εποικοδόμηση είναι η αυτορύθμιση, η ικανότητα δηλαδή του μαθητή να λειτουργεί σαν δάσκαλος του εαυτού του και η μεταγνώση που αφορά κυρίως στην αναθεώρηση της γνώσης. Εποικοδόμηση νέων γνώσεων σημαίνει είτε επέκταση των προϋπαρχουσών γνώσεων μέσα από συσσώρευση των νέων γνώσεων που συνδέονται με τις προϋπαρχουσες είτε με μικρή ή μεγάλη τροποποίηση της γνωστικής δομής στην περίπτωση που προϋπαρχουσες και νέες γνώσεις είναι ασύμβατες ή αντίθετες.

Κατά συνέπεια ο δάσκαλος δεν αρκεί να γνωρίζει καλά τη νέα γνώση ούτε να τη διδάξει "σωστά", αλλά πρέπει να μπορεί να επεκτείνει τις γνώσεις των μαθητών ή να τροποποιεί τις παρανοήσεις τους. Για να μπορέσει το σύστημα μαθησιακής τεχνολογίας του IEEE, το οποίο έχει αντικαταστήσει τον κλασικό δάσκαλο με έναν ηλεκτρονικό δάσκαλο, να ευθυγραμμισθεί με τη θεωρία της εποικοδομητικής υπόθεσης θα πρέπει να συμπεριλάβει την έννοια της εξατομικευμένης μάθησης.

Γιαυτό το λόγο, η προσέγγισή μας εισάγει τις έννοιες των χρηστών και των ειδικών καθώς και την έννοια της αυτοματοποιημένης εξαγωγής προφίλ χρηστών, η οποία και εξετάζεται αναλυτικά στο τρίτο κεφάλαιο. Η νέα αυτή προσέγγιση δίνει στο χρήστη τη δυνατότητα να διερευνήσει το παρουσιαζόμενο υλικό με βάση τις ανάγκες και τις προτιμήσεις του, ακολουθώντας μία προσωπική διαδρομή. Κατά αυτό τον τρόπο ο μαθητής- χρήστης αρχίζει αυτόματα να έχει ενεργό ρόλο στη μάθηση. Στόχος της προτεινόμενης επέκτασης είναι το σύστημα να είναι σε θέση να στραφεί προς τις εσωτερικές πτυχές του μαθητή-χρήστη, τη συμπεριφορά του και τις γνωστικές διαδικασίες που υπεισέρχονται στην αλληλεπίδραση του με τον υπολογιστή, μετατρέποντας τη διαδικασία της μάθησης από παθητική σε ενεργητική.

Όσον αφορά στις γνωστικές διαδικασίες, ιδιαίτερη προσοχή δόθηκε στον τρόπο χειρισμού του εκπαιδευτικού περιεχομένου. Σύμφωνα με έρευνες [27] που έχουν γίνει στον χώρο της εκπαίδευσης έχει αποδειχθεί ότι μια δραστηριότητα μπορεί να αποδώσει



Σχήμα 2.3: Αρχιτεκτονική προτεινόμενου ΣΜΤ

καλύτερα αν δεν βασίζεται μόνο σε κείμενο. Αξιοποιώντας τα αποτελέσματα αυτής της έρευνας, το προτεινόμενο σύστημα μπορεί να περιλάβει, όπου απαιτείται, τη χρήση εικόνων, αριθμών, σχεδίων, ταινιών video και ήχου και επιπλέον επιτρέπει στους χρήστες να έχουν διαλογική σχέση με πληθώρα αντικειμένων στην οθόνη ώστε να μπορούν να αποκτήσουν γνώσεις με ένα διαφορετικό και πιο εποικοδομητικό τρόπο.

Το προτεινόμενο σύστημα περιλαμβάνει δύο ομάδες - αυτή των χρηστών και αυτή των ειδικών - καθώς και έναν εξυπηρετητή.

Στην ομάδα χρηστών ανήκουν όλοι οι χρήστες του συστήματος οι οποίοι μπορούν να έχουν πρόσβαση στους πόρους εκμάθησης του εξυπηρετητή. Κάθε φορά που ο χρήστης συνδέεται με τον εξυπηρετητή, ενεργοποιείται το προφίλ του, το οποίο και καθορίζει τον τρόπο παρουσίασης του εκπαιδευτικού υλικού.

Η ομάδα των ειδικών περιλαμβάνει ένα πλήθος ατόμων όπως έμπειρους εκπαιδευτικούς, εκπαιδευτικούς με εμπειρία σε θέματα της ειδικής εκπαίδευσης, ειδικούς σε θέματα ΤΠΕ, ειδικούς σε διάφορα ερευνητικά θέματα και ειδικούς σε συστήματα εξ αποστάσεως εκπαίδευσης, μηχανικούς λογισμικού, στατιστικούς αναλυτές, ψυχολόγους και προγραμματιστές. Ο ρόλος αυτής της ομάδας είναι πολλαπλός καθώς είναι υπεύθυνη για τη σχεδίαση των πόρων εκπαίδευσης, τον εμπλουτισμό του συστήματος με κατάλληλα εργαλεία για την ανάλυση των αποκρίσεων των χρηστών και την κατηγοριοποίηση τους όσο και για την ανάλυση και ασφαλή μετάδοση των πολυμεσικών εκπαιδευτικών υλικών που το σύστημα παρέχει.

Η μονάδα του εξυπηρετητή αποτελείται από ένα ή περισσότερα ισχυρά υπολογιστικά μηχανήματα (κεντρική ή κατακεντρωμένη) όπου είναι εγκατεστημένος ένας εξυπηρετητής διαδικτύου (Web Server), μια βάση δεδομένων (RDBMS) που συσσωρεύει τις πληροφορίες που συλλέγονται μέσω του συστήματος, συστήματα επεξεργασίας εικόνας και βίντεο, συστήματα εξόρυξης δεδομένων, συστήματα μηχανικής μάθησης, συστήματα υδατογράφησης καθώς και το εκπαιδευτικό υλικό.

Στην προτεινόμενη επέκταση, η διαδικασία Χρήστης αποτελεί την κεντρική οντότητα γύρω από την οποία έχει αναπτυχθεί το προτεινόμενο σύστημα. Σε αυτό το πλαίσιο, η διαδικασία Παράδοση Υλικού έχει εμπλουτιστεί με ένα εργαλείο λογισμικού

το οποίο επιτρέπει την παραγωγή πολύγλωσσων ηλεκτρονικών ερωτηματολογίων. Το προτεινόμενο εργαλείο λογισμικού παρέχει τη δυνατότητα στους ειδικούς του συστήματος μάθησης να εισάγουν ερωτήσεις αλλά και τις πιθανές απαντήσεις. Το σύστημα αναλαμβάνει να αποθηκεύσει αυτά τα κείμενα στην αποθήκη Πόροι Εκμάθησης. Όταν ο χρήστης προσπελάσει για πρώτη φορά το προτεινόμενο σύστημα εκπαίδευσης, η διαδικασία Παράδοση Υλικού μορφοποιεί κατάλληλα τα κείμενα των ερωτηματολογίων που είναι αποθηκευμένα στην αποθήκη Πόροι Εκμάθησης και μέσω της ροής Πολυμεσική Παρουσίαση τα αποστέλλει στη διαδικασία Χρήστης. Οι πληροφορίες που συλλέγονται από τα ερωτηματολόγια (ροή συμπεριφορά) εκτιμούνται από τη διαδικασία Αξιολόγηση και φυλάσσονται στην αποθήκη Αρχεία Χρηστών.

Οι παρατηρήσεις αυτές διοχετεύονται στη διαδικασία Αξιολόγηση διαμέσου της ροής Συμπεριφορά. Επιπλέον η διαδικασία Αξιολόγηση συλλέγει από τη ροή Συμπεριφορά δεδομένα σχετικά με τη δραστηριότητα του χρήστη καθ'όλη τη διάρκεια παρουσίασης του πολυμεσικού υλικού. Αυτή η συλλογή περιέχει πληροφορίες σχετικές με το μέσο χρόνο απόκρισης του χρήστη στο υλικό της παρουσίασης και με τη γενική του εικόνα κατά τη διάρκεια της παρουσίασης. Όλες αυτές οι πληροφορίες οικοδομούν το τρέχον προφίλ του χρήστη το οποίο και φυλάσσεται στην αποθήκη Αρχεία Χρηστών. Συγκεκριμένα, η διαδικασία Αξιολόγηση έχει εμπλουτιστεί με ένα αυτόματο σύστημα το οποίο υλοποιεί αυτόματη εξαγωγή των προφίλ των χρηστών (η λειτουργία του οποίου θα περιγραφεί αναλυτικά στο Κεφάλαιο 3).

Γενικά, τα προφίλ των χρηστών, τα οποία αρχικά καθορίζονται από την ομάδα ειδικών περιέχονται στην αποθήκη Προφίλ Χρηστών. Κάθε ένα από αυτά περιγράφει τα χαρακτηριστικά των κατηγοριών των χρηστών καθώς και τις πιθανές εκπαιδευτικές ανάγκες κάθε κατηγορίας. Οι απαντήσεις των χρηστών χρησιμοποιούνται από τη διαδικασία Αξιολόγηση με σκοπό την επιλογή του κατάλληλου προφίλ στο οποίο θα καταταγεί ένας νέος χρήστης. Μόλις γίνει η κατάταξη του χρήστη, φυλάσσεται στην αποθήκη Αρχεία Χρηστών μαζί με πιθανές προσωπικές προτιμήσεις του.

Η διαδικασία Δάσκαλος μπορεί να λάβει τις πληροφορίες από τα Αρχεία Χρηστών οποιαδήποτε στιγμή. Πληροφορίες αξιολόγησης και πιστοποίησης αποθηκεύονται επίσης στη διαδικασία Αρχεία Χρηστών μέσω της διαδικασίας Δάσκαλος. Βασιζόμενος σε αυτές τις πληροφορίες, η διαδικασία Δάσκαλος δημιουργεί ερωτήματα και τα προωθεί στην αποθήκη Υλικό Εκμάθησης η οποία θα παρέχει το κατάλληλο υλικό στον χρήστη. Ανάλογα με την απόκριση (ροή Συμπεριφορά) του χρήστη στη ροή Πολυμέσα, μπορεί να εξαχθεί αυτόματα το προφίλ του. Ανάλογα με το προφίλ του κάθε χρήστη, παρουσιάζονται σε αυτόν κατάλληλοι πόροι από την αποθήκη Πόροι Εκμάθησης. Οι πολυμεσικοί πόροι εκμάθησης (κείμενα, εικόνες, βίντεο) έχουν κατηγοριοποιηθεί, χρησιμοποιώντας τεχνικές επεξεργασίας εικόνων και βίντεο και διατίθενται μέσω του συστήματος στους χρήστες έχοντας προηγουμένως προσαρμοστεί στο προφίλ τους.

Τέλος, η διαδικασία Παράδοση μεταφέρει το εκπαιδευτικό υλικό στο χρήστη. Η διαδικασία Δάσκαλος χρησιμοποιεί αυτή την πληροφορία για να ανακτήσει το εκπαιδευτικό υλικό που ταιριάζει σε κάθε περίπτωση, από την αντίστοιχη αποθήκη. Τυχόν αξιολογήσεις του χρήστη σχετικές με τις επιδόσεις του φυλάσσονται στην αποθήκη Αρχεία Χρηστών.

Ένα νέο προφίλ χρήστη μπορεί να δημιουργηθεί ή ένα υπάρχον μπορεί να τροποποιηθεί βασίζόμενο στη στατιστική έρευνα. Αυτό το νέο προφίλ φυλάσσεται στην αποθήκη Προφίλ Χρηστών. Όταν ένας χρήστης ο οποίος έχει περάσει τη διαδικασία Αξιολόγησης (μετά τη συμπλήρωση του σχετικού ερωτηματολογίου) επισκεφθεί ξανά

το προτεινόμενο σύστημα τότε το προφίλ του ανακτάται από την αποθήκη Προφίλ Χρηστών. Αλλαγές στο προφίλ του χρήστη μπορούν να πραγματοποιηθούν με βάση τη διαδικασία εκπαίδευσης.

Η διαδικασία Ηλεκτρονική Έρευνα έχει προστεθεί στο προτεινόμενο σύστημα. Ο σκοπός αυτής της διαδικασίας είναι η στατιστική επεξεργασία των στοιχείων των χρηστών, τα οποία έχουν συλλεχθεί από την διαδικασία Αξιολόγηση και έχουν αποθηκευτεί στην αποθήκη Αρχεία Χρηστών. Η διαδικασία αυτή εξάγει χρήσιμα ερευνητικά στοιχεία για οποιοδήποτε ενδιαφερόμενο.

Όσο αφορά στο εκπαιδευτικό υλικό, η διαδικασία Παράδοση Υλικού χρησιμοποιεί τεχνολογίες εξόρυξης δεδομένων, ανάλυσης εικόνας και μηχανικής μάθησης με σκοπό την αντιστοίχιση του υλικού με τα προφίλ. Κατά αυτόν τον τρόπο ο χρήστης θα παραλάβει εκπαιδευτικό υλικό που θα είναι κατάλληλο για το προφίλ του. Οι προτεινόμενες μέθοδοι ανάλυσης του εκπαιδευτικού υλικού περιγράφονται στο κεφάλαιο 4. Το προτεινόμενο σύστημα λειτουργεί δικτυακά με συνέπεια το εκπαιδευτικό υλικό να χρειάζεται προστασία για να μπορεί να μεταδίδεται με ασφάλεια. Στο 6ο κεφάλαιο περιγράφονται αναλυτικά οι προτεινόμενοι αλγόριθμοι προστασίας του υλικού.

Βασιζόμενοι στην προτεινόμενη επέκταση του συστήματος μαθησιακής τεχνολογίας, σχεδιάσαμε και υλοποιήσαμε τρία σενάρια χρήσης. Στο πρώτο σενάριο ως χρήστες συμμετείχαν εκπαιδευτικοί με διαφορετικό επίπεδο γνώσεων πάνω σε θέματα νέων τεχνολογιών και στόχος του σεναρίου ήταν η επιμόρφωσή τους πάνω σε θέματα σχετικά με τεχνολογίες πληροφορικής και υπολογιστών (ΤΠΕ). Το θέμα του δεύτερου σεναρίου ήταν η δημιουργία ενός εργαλείου υποβοήθησης ανάγνωσης για παιδιά με προβλήματα δυσλεξίας. Το τρίτο σενάριο περιλαμβάνει ως χρήστες φοιτητές σε μάθημα Βυζαντινής τέχνης και στόχος του σεναρίου είναι οι φοιτητές αυτοί να κατανοήσουν τις τεχνικές που χρησιμοποιούνται στην αγιογραφία τόσο μέσα από την ανάλυση όσο και από τη σύνθεση των μορφών αγίων. Τα προαναφερόμενα σενάρια αυτά καθώς και η υλοποίησή τους περιγράφονται αναλυτικά στο κεφάλαιο εφαρμογών συστημάτων μαθησιακής τεχνολογίας (Κεφάλαιο 7).

Κεφάλαιο 3

Εξόρυξη Γνώσης και Προσωποποιημένη Πρόσβαση

3.1 Εισαγωγή

Τα πρώτα συστήματα μαθησιακής τεχνολογίας [14], [129], [36], [37] που εμφανίστηκαν αντιμετωπίζουν με όμοιο τρόπο τους χρήστες τους, χωρίς να λαμβάνουν υπόψη την ιδιαιτερότητα τους. Η ανάγκη για προσωποποιημένη πρόσβαση σε ψηφιακό υλικό μέσω συστημάτων μαθησιακής τεχνολογίας οδήγησε στην δημιουργία συστημάτων με δυνατότητα συλλογής και επεξεργασίας δεδομένων του χρήστη με απώτερο σκοπό την εξόρυξη γνώσης από αυτά. Τα δεδομένα αυτά συνήθως αντλούνται από τις ανάγκες, τους στόχους, τις γνώσεις, τα ενδιαφέροντα, τις προτιμήσεις και τη συμπεριφορά ενός χρήστη. Η συλλογή και η αξιοποίηση αυτού του συνόλου των δεδομένων του κάθε χρήστη οδηγούν στη δημιουργία του προφίλ του.

Στο κεφάλαιο αυτό παρουσιάζουμε την ερευνητική συνεισφορά μας στο πεδίο της εξόρυξης γνώσης και της προσωποποιημένης πρόσβασης. Αρχικά, θα παρουσιάσουμε ορισμένες θεμελιώδεις μεθοδολογίες που είναι σχετικές με το ερευνητικό πεδίο της εξόρυξης γνώσης από δεδομένα. Η προτεινόμενη διαδικασία κατηγοριοποίησης χρηστών παρουσιάζεται κατόπιν μέσα από το πρίσμα τόσο των κλασικών προσεγγίσεων από το χώρο της μηχανικής μάθησης και της τεχνητής νοημοσύνης, καθώς και από την αξιοποίηση των τεχνολογιών επεξεργασία εικόνας και βίντεο. Για την αντιμετώπιση των ερευνητικών προβλημάτων που μας απασχόλησαν συνδυάσαμε τεχνικές συσταδοποίησης και κατηγοριοποίησης δεδομένων παράλληλα με μείωση των διαστάσεων των δεδομένων κατά τη συσταδοποίηση, προκειμένου να βελτιώσουμε την απόδοση και την επίδοσή τους. Τέλος προτάθηκε μια τεχνική αυτόματης αξιολόγησης της κατάστασης της συμπεριφοράς των χρηστών, η οποία ενισχύει τη μη λεκτική αλληλεπίδραση μεταξύ ανθρώπου και μηχανής και συμβάλλει στην αυτόματη κατηγοριοποίηση των χρηστών.

3.2 Εξόρυξη γνώσης από δεδομένα

Λόγω της εξάπλωσης και της ευρείας χρήσης υπολογιστών, ο καθημερινά αυξανόμενος αριθμός δεδομένων και βάσεων δεδομένων έχει οδηγήσει σε μια καινούρια θεώρηση των προβλημάτων ανάλυσης και ομαδοποίησης των δεδομένων. Σε αυτό το πλαίσιο παρεμβαίνει η εξόρυξη γνώσης από δεδομένα αποβλέποντας στην αποκάλυψη ή

παραγωγή λειτουργικής γνώσης, μέσω της ανάλυσης αυτών των δεδομένων. Η εξόρυξη γνώσης από δεδομένα μπορεί να εφαρμοστεί χάρη στην εξέλιξη των ερευνητικών πεδίων της μηχανικής μάθησης, της στατιστικής και των βάσεων δεδομένων. Οι τεχνικές που έχουν αναπτυχθεί αποτελούνται συνήθως από τρία στάδια: στο πρώτο στάδιο πραγματοποιείται η εξέταση των δεδομένων με σκοπό τον καθορισμό μοντέλου που να είναι πλησιέστερο στα χαρακτηριστικά των δεδομένων που εξετάζονται, στο δεύτερο στάδιο χρησιμοποιούνται κάποια κριτήρια για το ταίριασμα στα δεδομένα ενός μοντέλου έναντι κάποιου άλλου, ενώ στο τρίτο στάδιο εφαρμόζονται τεχνικές αναζήτησης δεδομένων.

3.2.1 Εξέταση δεδομένων

Η επιτυχία της διαδικασίας εξαγωγής γνώσης σχετίζεται άμεσα τόσο με το είδος των δεδομένων όσο και με το είδος της εφαρμογής. Έχοντας λοιπόν τα δεδομένα, συνήθως επιλέγεται δείγμα από αυτά με σκοπό τον καθορισμό του συνόλου των μεταβλητών. Τα δεδομένα αυτά συνήθως είναι πολυδιάστατα διανύσματα, όπου κάθε διάσταση εκφράζει και μια μεταβλητή. Οι μεταβλητές αυτές μπορεί να είναι είτε ποσοτικές είτε ποιοτικές. Οι ποσοτικές μεταβλητές είναι συνεχείς, δυαδικές ή διακριτές, ενώ οι ποιοτικές είναι κατηγορικές (ονομαστικής κλίμακας, σε κλίμακα κατάταξης) ή διάταξης (βαθμολογία: λίαν καλώς/ άριστα)).

Η επιλογή των δεδομένων που θα χρησιμοποιηθούν επηρεάζει σημαντικά τόσο την αναπαράσταση των δεδομένων όσο και την ποιότητα της εξαγόμενης γνώσης. Σε πολλές περιπτώσεις απαιτείται μια προεπεξεργασία των δεδομένων που θα κάνει εφικτή την περαιτέρω επεξεργασία τους. Η προεπεξεργασία των δεδομένων μπορεί να περιλαμβάνει τη μετατροπή τους σε ίδια κλίμακα, την κανονικοποίηση τους ως προς μια μονάδα μέτρησης ή και τη χρήση βαρών με σκοπό τη δημιουργία κατάλληλου διανύσματος αναπαράστασης των δεδομένων. Στην περίπτωση απουσίας τιμών συνήθως γίνεται αντικατάσταση των ελλιπών δεδομένων με τη μέση τιμή του χαρακτηριστικού ή με τη μεγαλύτερης συχνότητας εμφανιζόμενη τιμή.

Κατά τη διαδικασία εκπαίδευσης, πιθανή εισαγωγή θορύβου στο σύστημα μπορεί να οδηγήσει τον αλγόριθμο σε λανθασμένα συμπεράσματα. Η εισαγωγή θορύβου μπορεί να συμβεί με την εισαγωγή λανθασμένων δεδομένων στο σύστημα από το χρήστη. Η εκτεταμένη παρουσία θορύβου μπορεί να οδηγήσει σε λανθασμένα αποτελέσματα τον αλγόριθμο μάθησης και σε επιλογή μιας υποβέλτιστης υπόθεσης, ελάχιστα συνεπώς με τα δεδομένα ελέγχου [31]. Στη βιβλιογραφία έχουν προταθεί διάφοροι τρόποι ανίχνευσης σφαλμάτων [156]. Συνήθως πραγματοποιείται ο καθορισμός των σφαλμάτων χωρίς τη χρήση προγενέστερης γνώσης των δεδομένων. Η προσέγγιση αυτή θεωρεί ότι τα δεδομένα ακολουθούν μια κανονική κατανομή και αφαιρεί όσες τιμές δεν ανήκουν στην κατανομή αυτή. Επιπλέον, τα δεδομένα πρέπει να ελεγχθούν για να διαπιστωθεί αν υπάρχουν μεταβλητές με σταθερές τιμές, οι οποίες και παραλείπονται.

Το πλήθος των δεδομένων που συγκεντρώνεται είναι τεράστιο και συχνά περιλαμβάνει μεταβλητές άχρηστες ή πλεονάζουσες οπότε κρίνεται απαραίτητη η συμπύκνωση της πληροφορίας και ο περιορισμός του πλήθους των μεταβλητών με τη δημιουργία νέων μεταβλητών. Συχνά πραγματοποιείται προεπεξεργασία με σκοπό τη μείωση του χώρου διαστάσεων των δεδομένων, διαδικασία που οδηγεί στη μείωση της υπολογιστικής πολυπλοκότητας για την εκπαίδευση και τη λειτουργία του ταξινομητή. Η προεπεξεργασία γίνεται συνήθως με την απευθείας επιλογή χαρακτηριστικών

γνωρισμάτων ή με διαδικασία εξαγωγή των χαρακτηριστικών γνωρισμάτων από το πλήθος των δεδομένων. Οι δύο πιο δημοφιλείς μέθοδοι που εφαρμόζονται για την επιλογή των πιο σημαντικών χαρακτηριστικών είναι του περιτυλίγματος [79] και του φίλτρου [39]. Από την άλλη, η εξαγωγή χαρακτηριστικών γνωρισμάτων εφαρμόζει μια αντιστοίχιση του πολυδιάστατου χώρου σε ένα χώρο λιγότερων διαστάσεων. Αυτό σημαίνει ότι ο αρχικός χώρος των χαρακτηριστικών γνωρισμάτων μετασχηματίζεται με την εφαρμογή, για παράδειγμα, ενός γραμμικού μετασχηματισμού μέσω μιας ανάλυσης κύριων συνιστωσών (principal components analysis).

3.2.2 Επιλογή μεθόδου μέτρησης ομοιότητας

Έπειτα από τον καθορισμό των μεταβλητών που θα χρησιμοποιηθούν για την κατηγοριοποίηση των δεδομένων, επιλέγεται μια μέθοδος μέτρησης ομοιότητας των μεταβλητών αυτών [72]. Το πλήθος των ομάδων στις οποίες θα χωρισθούν τα δεδομένα και το αν αυτό θα είναι προκαθορισμένο ή όχι επιλέγεται ανάλογα με τη εφαρμογή. Αυτή η επιλογή μπορεί να οδηγήσει στον καθορισμό του καταλληλότερου, για την εφαρμογή, αλγορίθμου μάθησης, ο οποίος θα είναι ικανός να οδηγήσει σε ακριβή αποτελέσματα.

Οι έννοιες της απόστασης και της ομοιότητας είναι αντίθετες αφού παρατηρήσεις που είναι όμοιες θα παρουσιάζουν ομοιότητα και θα απέχουν μεταξύ τους κατά μικρή απόσταση. Οι έννοιες αυτές είναι πολύ χρήσιμες καθώς μας επιτρέπουν να μετρήσουμε πόσο μοιάζουν οι παρατηρήσεις μεταξύ τους και να τις τοποθετήσουμε στην ίδια ή σε διαφορετική ομάδα.

Ας υποθέσουμε αρχικά ότι ενδιαφερόμαστε για δύο συνεχείς μεταβλητές που λήφθηκαν ως απαντήσεις ενός χρήστη σε δύο ερωτήσεις π.χ. Α και Β, δηλαδή το δεδομένο μας αντιπροσωπεύεται από ένα διδιάστατο διάνυσμα, κάθε διάσταση του οποίου εκφράζει και μία μεταβλητή. Αν συμβολίσουμε τα δύο δεδομένα ως $x = (x_1, x_2)$ και $y = (y_1, y_2)$ τότε μια πρώτη προσέγγιση για την επιλογή μιας απόστασης ανάμεσα σε αυτά τα διανύσματα θα ήταν η ευκλείδια απόσταση:

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (3.1)$$

την οποία μπορούμε να γενικεύσουμε για την περίπτωση που έχουμε n -διάστατες μεταβλητές. δηλαδή $x = (x_1, x_2, \dots, x_n)$ και $y = (y_1, y_2, \dots, y_n)$, οπότε η ευκλείδια απόσταση δίνεται από τον τύπο :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3.2)$$

Αν οι μεταβλητές των διανυσμάτων είναι σε διαφορετική κλίμακα, η ευκλείδια απόσταση αποδεικνύεται ότι δεν παρέχει αξιόπιστα αποτελέσματα αφού οι μεταβλητές που έχουν μεγάλες απόλυτες τιμές θα έχουν πολύ μεγαλύτερο βάρος και σχεδόν θα καθορίζουν την απόσταση ανάμεσα στα δεδομένα. Επιπλέον, η ευκλείδια απόσταση εξαρτάται πολύ από την κλίμακα μέτρησης κι επομένως, αλλάζοντας την κλίμακα, λαμβάνουμε διαφορετικές αποστάσεις.

Ανάλογα με το είδος των δεδομένων πρέπει να επιλεγούν και τα κατάλληλα μέτρα αποστάσεων. Για παράδειγμα, όταν τα δεδομένα είναι συνεχή χρησιμοποιούνται αποστάσεις όπως η ευκλείδια απόσταση, η city-block απόσταση, η Chebychev απόσταση [72]. Όλες αυτές οι αποστάσεις αντιμετωπίζουν το πρόβλημα της κλίμακας

των μετρήσεων. Ένας γνωστός τρόπος με τον οποίο είναι εφικτή η αντιμετώπιση του προβλήματος αυτού, είναι η μετατροπή της κάθε μεταβλητής σε συγκρίσιμη κλίμακα.

Αν οι μεταβλητές αφορούν σε διαφορετικούς τύπους μεταβλητών, όπως για παράδειγμα ηλικία (συνεχής), φύλο (δυναδική), οικογενειακή κατάσταση (ονομαστική κατηγορική) τότε τα δεδομένα θεωρούνται δεδομένα μεικτού τύπου (mixed mode variables) [69]. Ανάμεσα στους τρόπους αντιμετώπισης τέτοιας μορφής δεδομένων είναι ο ορισμός ψευδομεταβλητών για όλους τους τύπους των δεδομένων που τελικά οδηγεί σε ένα σύνολο από δυναδικές μεταβλητές. Στην περίπτωση συνεχών μεταβλητών αυτό προϋποθέτει μια κβαντοποίηση-διακριτοποίησή τους, δηλαδή αρχικά τις μετατρέπουμε σε μικρότερα δυναδικά διαστήματα και στη συνέχεια ορίζουμε ψευδομεταβλητές για τις διακριτοποιημένες μεταβλητές. Προφανώς αυτή η προσέγγιση μπορεί να οδηγήσει σε σοβαρή απώλεια πληροφορίας κατά τη διάρκεια των μετασχηματισμών αυτών οπότε και απαιτείται μεγάλη προσοχή.

Η πιο διαδεδομένη μεθοδολογία βασίζεται στον υπολογισμό μιας απόστασης, κατάλληλης για μικτής μορφής δεδομένα. Η απόσταση αυτή ορίστηκε από τον Gower το 1971 [54] και έχει τη μορφή:

$$d(x, y) = \frac{\sum_{i=1}^n w_i(x, y) \cdot d_i(x, y)}{\sum_{i=1}^n w_i(x, y)} \quad (3.3)$$

όπου $d(x, y)$ είναι η απόσταση ανάμεσα στις παρατηρήσεις x και y για την i μεταβλητή. Η ιδέα είναι να χρησιμοποιήσει κανείς διαφορετικές αποστάσεις για κάθε τύπο μεταβλητής, όμως αυτές οι αποστάσεις να είναι στην ίδια κλίμακα ώστε να είναι συγκρίσιμες και να συνεισφέρουν ισόποσα στη συνολική απόσταση.

Συγκεκριμένα:

- Αν η μεταβλητή είναι κατηγορική (είτε ονομαστικής κλίμακας είτε σε κλίμακα κατάταξης), τότε $d_i(x, y) = \begin{cases} 1, & \text{ομοια} \\ 0, & \text{ανομοια} \end{cases}$
- Αν η μεταβλητή είναι συνεχής, $d_i(x, y) = 1 - \frac{|x_i - y_i|}{\mu_i}$, όπου μ_i είναι το εύρος της μεταβλητής. Ουσιαστικά, η διαίρεση με το εύρος πραγματοποιείται με σκοπό η τελική τιμή να κυμαίνεται στο διάστημα $[0, 1]$.
- Τα βάρη παίρνουν την τιμή 0 ή 1 ανάλογα με το αν η σύγκριση ανάμεσα στις δύο παρατηρήσεις για τη συγκεκριμένη μεταβλητή έχει έννοια. Όταν η σύγκριση έχει νόημα η τιμή είναι 1 αντί για 0. Περιπτώσεις με βάρος 0 έχουμε όταν μια εκ των δύο τιμών δεν υπάρχει (missing values) ή όταν οι μεταβλητές είναι κατηγορικές και ασύμμετρες με κοινή απουσία του χαρακτηριστικού, το οποίο δεν έχει ενδιαφέρον.

3.2.3 Ταίριασμα μοντέλου στα δεδομένα

Η επιλογή του καταλληλότερου αλγορίθμου μάθησης για το δοθέν πρόβλημα περιλαμβάνει αποφάσεις για το ποιο μοντέλο και ποιες παράμετροι είναι οι πλέον κατάλληλες. Διάφορες μέθοδοι μπορεί να βοηθήσουν σε αυτή την απόφαση όπως τεχνητά νευρωνικά δίκτυα διαφόρων τύπων, δέντρα απόφασης, Bayesian ταξινομητές, κ.λ.π. Η επιλογή του μοντέλου που θα χρησιμοποιήσουμε για την τελική λύση είναι ένα περίπλοκο πρόβλημα που καλούμαστε να μελετήσουμε.

Όσον αφορά στα μοντέλα, υπάρχουν δύο μεγάλες κατηγορίες, τα προβλέψιμα (predictive model) και τα περιγραφικά (descriptive model). Τα προβλέψιμα μοντέλα ουσιαστικά δημιουργούν ή επιλέγουν στατιστικά πρότυπα με σκοπό την καλύτερη προβλέψη της πιθανότητας μιας μελλοντικής συμπεριφοράς των δεδομένων βασιζόμενα στη χρήση της πρότερης γνώσης που έχει εξαχθεί από αυτά. Σε αυτή την κατηγορία υπάγονται οι τεχνικές κατηγοριοποίησης (classification), παλινδρόμησης (regression), πρόβλεψης (prediction) και ανάλυσης χρονολογικών σειρών (time series analysis). Τα περιγραφικά μοντέλα, αντίθετα με τα προβλέψιμα, διερευνούν τις ιδιότητες των δεδομένων που εξετάζονται, με σκοπό να αναγνωρίσουν πρότυπα και συσχετίσεις σε αυτά. Περιγραφικά μοντέλα θεωρούνται η συσταδοποίηση (clustering), η παρουσίαση συνόψεων (summarization), οι κανόνες συσχετίσεων (association rules) και η ακολουθιακή ανάλυση (sequential analysis).

Η κατηγοριοποίηση είναι από τις πιο γνωστές τεχνικές της εξόρυξης γνώσης και ανήκει στα προβλέψιμα μοντέλα. Η κατηγοριοποίηση χρησιμοποιείται σε πολλούς επιστημονικούς τομείς και αναφέρεται συχνά σαν εποπτευόμενη μάθηση επειδή οι κατηγορίες καθορίζονται πριν την εξέταση των δεδομένων. Οι αλγόριθμοι κατηγοριοποίησης αντιστοιχούν το σύνολο των δεδομένων σε ένα πεπερασμένο αριθμό κατηγοριών. Η παλινδρόμηση προϋποθέτει ότι τα σχετικά δεδομένα ταιριάζουν με μερικά γνωστά είδη συνάρτησεων (γραμμική, λογαριθμική κ.α.) και καθορίζει την καλύτερη συνάρτηση αυτού του είδους που μοντελοποιεί τα δεδομένα που έχουν δοθεί. Για την απόδοση της καλύτερης συνάρτησης χρησιμοποιούνται τεχνικές ανάλυσης σφάλματος. Με την ανάλυση χρονοσειρών μελετάται η μεταβολή ενός γνωρίσματος στο πέρασ του χρόνου, λαμβάνοντας τιμές σε ίσα διαστήματα. Η πρόβλεψη μπορεί να θεωρηθεί σαν κατηγοριοποίηση της τιμής ενός γνωρίσματος σε μία προκαθορισμένη κλάση, λαμβάνοντας υπόψη τις τρέχουσες καθώς και τις παρελθόντες τιμές του συγκεκριμένου γνωρίσματος.

Η συσταδοποίηση, τεχνική των περιγραφικών μοντέλων, οργανώνει τα δεδομένα σε συστάδες (clusters), οι οποίες δεν είναι προκαθορισμένες (σε αντίθεση με την κατηγοριοποίηση). Σκοπός των αλγορίθμων συσταδοποίησης είναι η εύρεση ομοιοτήτων ως προς προκαθορισμένα γνωρίσματα μεταξύ των δεδομένων και η ομαδοποίηση των ομοίων σε ίδιες συστάδες. Στην τεχνική παρουσίασης συνόψεων, τα δεδομένα συνοδεύονται από απλές περιγραφές οι οποίες και χρησιμοποιούνται για να εξάγουν σχετικές πληροφορίες. Οι τεχνικές κανόνων συσχέτισης ασχολούνται με την εξαγωγή γνώσης που αποκαλύπτει συσχετίσεις μεταξύ των δεδομένων ενώ η ακολουθιακή ανάλυση εξετάζει την συμπεριφορά των προτύπων σε δεδομένα που συλλέγονται από δύο διαφορετικές χρονολογικές σειρές για να βρει την ομοιότητά τους.

Το πρόβλημα της συσταδοποίησης σχετίζεται με την τμηματοποίηση (partitioning) ενός συνόλου δεδομένων σε συστάδες, έτσι ώστε τα στοιχεία του συνόλου των δεδομένων που ανήκουν σε μία συστάδα να είναι περισσότερο όμοια μεταξύ τους από ό,τι είναι με τα στοιχεία των άλλων συστάδων. Δύο βασικές έννοιες για την ανάλυση κατά συστάδες είναι η απόσταση και η ομοιότητα. Θεωρώντας ως παρατήρηση το κάθε διάλυμα δεδομένων, μπορούμε να πούμε ότι οι παρατηρήσεις που έχουν σχετικά όμοιες τιμές θα έχουν πολύ μεγάλη τιμή ομοιότητας και θα απέχουν ελάχιστα. Το αντίστροφο θα ισχύει για παρατηρήσεις που έχουν ανόμοιες τιμές. Γίνεται φανερό ότι σκοπός της ανάλυσης σε συστάδες είναι η δημιουργία ομάδων των οποίων οι παρατηρήσεις θα απέχουν λίγο, ενώ παρατηρήσεις διαφορετικών ομάδων θα απέχουν αρκετά μεταξύ τους. Η συσταδοποίηση μπορεί να προσφέρει και λύση σε προβλήματα

όπως η εύρεση ομοιότητας μεταξύ δεδομένων, η εύρεση διακριτικής ικανότητας των μεταβλητών και η μείωση των διαστάσεων του προβλήματος, εξαλείφοντας επικαλύψεις και μη σημαντικές μεταβλητές και επικεντρώνοντας σε μεταβλητές που παρουσιάζουν ενδιαφέρον. Επίσης, μέσω τεχνικών συσταδοποίησης είναι ικανή η δημιουργία και ο έλεγχος υποθέσεων σχετικά με τα δεδομένα καθώς και η πρόβλεψη καινούριων παρατηρήσεων.

Ο αριθμός των συστάδων δεν είναι εξ αρχής γνωστός, οπότε η δημιουργία τους είναι μια δυναμική διαδικασία. Η ποικιλία τεχνικών για την αναπαράσταση των δεδομένων, έκφραση της ομοιότητας μεταξύ στοιχείων και ομαδοποίηση των δεδομένων έχει ως αποτέλεσμα την ύπαρξη μιας πλούσιας συλλογής μεθόδων συσταδοποίησης.

Οι αλγόριθμοι συσταδοποίησης μπορεί να διαχωριστούν σε κατηγορίες όπως ιεραρχικούς (hierarchical), διαμεριστικούς (partitional), κατηγορηματικούς (categorical) και μεγάλων βάσεων δεδομένων (LargeDB) και καθένα από αυτά μπορεί να χωριστεί σε υποκατηγορίες. Στην ιεραρχική συσταδοποίηση δημιουργείται ένα εμφωλιασμένο σύνολο από συστάδες. Κάθε επίπεδο της ιεραρχίας έχει ένα ξεχωριστό σύνολο συστάδων όπου στο κατώτερο επίπεδο, κάθε αντικείμενο βρίσκεται στη δική του συστάδα ενώ στο ανώτερο επίπεδο, όλα τα αντικείμενα ανήκουν στην ίδια συστάδα. Οι ιεραρχικοί αλγόριθμοι χωρίζονται σε συσσωρευτικούς (agglomerative) όπου οι συστάδες δημιουργούνται ανοδική προσέγγιση και διαιρετικούς (divisive) όπου οι συστάδες δημιουργούνται καθοδική προσέγγιση. Στη διαμεριστική συσταδοποίηση, οι προτεινόμενες αλγοριθμικές προσεγγίσεις χρησιμοποιούν τον επιθυμητό αριθμών συστάδων προκειμένου να καταλήξουν σε ένα τελικό σύνολο συστάδων.

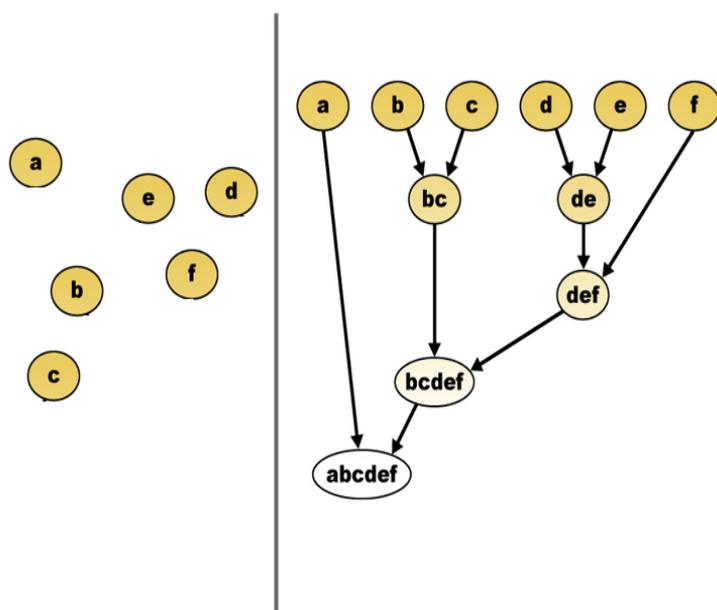
Ο γενικός ορισμός του προβλήματος της συσταδοποίησης δίνεται στη συνέχεια:

Ορισμός: Δοθείσης μια βάσης δεδομένων $D = \{t_1, t_2, \dots, t_n\}$ που αποτελείται από πλειάδες, ενός μέτρου ομοιότητας $sim(t_1, t_2)$, μεταξύ οποιωνδήποτε δύο πλειάδων, $t_i, t_l \in D$, και μιας ακέραιας τιμής k , το πρόβλημα της συσταδοποίησης συνίσταται στο να ορισθεί μια αντιστοιχισή $f : D \rightarrow \{1, \dots, k\}$ όπου κάθε t_i ανατίθεται σε μία συστάδα $K_j, 1 \leq j \leq k$. Για μία συστάδα $K_j, \forall t_{jl}, t_{jm} \in K_j$ και $t_i \notin K_j$, $dist(t_{jl}, t_{jm}) > dist(t_{jl}, t_i)$.

Η απόσταση είναι μια θεμελιώδης έννοια που έχει ως σκοπό να μετρήσει την απόσταση μεταξύ δύο παρατηρήσεων, ώστε να μπορέσει να προσδιορίσει αν μοιάζουν ή όχι. Στην Ενότητα 3.2.3.1 περιγράφονται οι αλγόριθμοι συσταδοποίησης που χρησιμοποιήθηκαν και πως αυτές τροποποιήθηκαν στη διάρκεια αυτής της διατριβής.

3.2.3.1 Ιεραρχικοί αλγόριθμοι

Όπως προαναφέρθηκε, οι ιεραρχικοί αλγόριθμοι συσταδοποίησης δημιουργούν στην πραγματικότητα σύνολα συστάδων. Ένας ιεραρχικός αλγόριθμος μπορεί να αποδοθεί με ένα δενδροδιάγραμμα (Σχήμα 3.1) το οποίο παρουσιάζει τις συγχωνεύσεις στοιχείων για τη δημιουργία συστάδων και τα επίπεδα ομοιότητας βάσει των οποίων αλλάζουν οι ομάδες και διαμορφώνονται οι συστάδες. Τα φύλλα του δενδρογράμματος περιέχουν ατομικές συστάδες, δηλαδή κάθε φύλλο περιέχει μία συστάδα με ένα στοιχείο, ενώ η ρίζα του δενδρογράμματος αποτελείται από μία συστάδα η οποία περιέχει όλα τα δοθέντα στοιχεία. Χαρακτηριστικό του διαγράμματος είναι τα διάφορα επίπεδα στα οποία αποδίδονται διαφορετικές ομαδοποιήσεις. Έτσι, σε κάθε επίπεδο,



Σχήμα 3.1: Ιεραρχικός αλγόριθμος: στοιχεία και δενδρόγραμμα

οι κόμβοι του δενδρογράμματος αναπαριστούν νέες συστάδες που προκύπτουν από τη συνένωση των συστάδων-παιδιών τους. Κάθε επίπεδο του δενδρογράμματος συνδέεται με ένα μέτρο απόστασης που χρησιμοποιείται για τη συγχώνευση των συστάδων, ώστε οι συστάδες-παιδιά να απέχουν μεταξύ τους απόσταση σύνδεσης του συγκεκριμένου επιπέδου στο δενδρόγραμμα.

Η μέθοδος αυτή παρουσιάζεται σχηματικά στο Σχήμα 3.1. Στο αριστερό τμήμα του σχήματος φαίνονται έξι στοιχεία $\{a, b, c, d, e, f\}$ τα οποία θέλουμε να ομαδοποιήσουμε σε συστάδες. Στο δεξί τμήμα του σχήματος εμφανίζεται το δενδρόγραμμα. Στο επίπεδο των φύλλων του δενδρογράμματος παρουσιάζονται έξι συστάδες $\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f\}$ καθεμία από τις οποίες αποτελείται από ένα μόνο στοιχείο. Στο αμέσως επόμενο επίπεδο υπάρχουν δύο νέα σύνολα συστάδων $\{b, c\}, \{d, e, f\}$, κάθε ένα από τα οποία περιέχει δύο στοιχεία που είναι πιο κοντά μεταξύ τους από ό,τι με τα υπόλοιπα στοιχεία. Στο επόμενο επίπεδο φαίνεται μία νέα συστάδα $\{d, e, f\}$ ενώ, στο προτελευταίο επίπεδο, οι συστάδες των δύο $\{b, c\}$ και τριών $\{d, e, f\}$ στοιχείων συγχωνεύονται σε μία νέα συστάδα την $\{b, c, d, e, f\}$. Τέλος, στο τελευταίο επίπεδο, τη ρίζα, τα στοιχεία συγχωνεύονται σε μια συστάδα $\{a, b, c, d, e, f\}$.

Συσσωρευτικοί αλγόριθμοι

Οι συσσωρευτικοί αλγόριθμοι (agglomerative) θεωρούν αρχικά ότι κάθε στοιχείο ανήκει στη δική τους συστάδα και επαναληπτικά συγχωνεύουν τις συστάδες, ώσπου όλα τα στοιχεία να ανήκουν στην ίδια συστάδα. Διαφοροποιήσεις των συσσωρευτικών αλγορίθμων προκύπτουν από τον τρόπο που γίνονται οι συγχωνεύσεις των συστάδων σε κάθε επίπεδο. Ένα παράδειγμα τέτοιου αλγορίθμου είναι ο τυπικός συσσωρευτικός αλγόριθμος, ο οποίος δέχεται ως είσοδο ένα σύνολο στοιχείων και τις αποστάσεις μεταξύ τους με μορφή μιας $n \times n$ μήτρας γειτνίασης κορυφών A , με τιμές $A[i, j] = dist(t_i, t_j)$. Η έξοδος του αλγορίθμου είναι ένα δενδρόγραμμα, DE , το οποίο αναπαρίσταται ως ένα σύνολο διατεταγμένων τριάδων d, k, K , όπου d είναι η τιμή του κατωφλίου, k το πλήθος των συστάδων και K το σύνολο

των συστάδων. Η έξοδος του τυπικού συσσωρευτικού αλγορίθμου θα είναι ένα δενδρόγραμμα και συγκεκριμένα για το παράδειγμα του Σχήμα 3.1 θα έχει τη μορφή: $\langle (0, 5, \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}\}), (1, 3, \{\{a\}, \{b, c\}, \{d, e\}\}), (2, 2, \{\{a\}, \{b, c, d, e\}\}), (3, 1, \{\{a, b, c, d, e\}\}) \rangle$

Ο χρήστης μπορεί να καθορίσει (με βάση το κατώφλι της απόστασης) ποιες από τις συστάδες επιθυμεί να χρησιμοποιήσει. Οι συσσωρευτικοί αλγόριθμοι διαφέρουν μεταξύ τους στον αριθμό των συστάδων που συγχωνεύονται σε κάθε επίπεδο, στο ποιες συστάδες θα συγχωνεύονται όταν υπάρχουν πολλαπλές συστάδες με τις ίδιες ακριβώς αποστάσεις και στην τεχνική που χρησιμοποιείται για τον καθορισμό της απόστασης μεταξύ των συστάδων.

Οι πιο γνωστές τεχνικές προέρχονται από τη θεωρία των γράφων και είναι οι τεχνικές απλού συνδέσμου (single link), πλήρους συνδέσμου (complete link) και μέσου συνδέσμου (average link).

Αναλυτικότερα, στην τεχνική απλού συνδέσμου (πλησιέστερου γείτονα) δύο συστάδες συγχωνεύονται αν η ελάχιστη απόσταση μεταξύ οποιωνδήποτε δύο σημείων των συστάδων είναι μικρότερη ή ίση με την απόσταση κατωφλίου. Η είσοδος στην τεχνική απλού συνδέσμου είναι ένας γράφος που αναπαρίσταται από μία μήτρα γειτνίασης κορυφών και ένα σύνολο κορυφών και ως έξοδο δίνει το πλήθος του συνόλου συνεκτικών συνιστωσών καθώς και έναν πίνακα που περιέχει πληροφορία μέλους για κάθε συνιστώσα. Η τεχνική αυτή παρουσιάζει το πρόβλημα της *επίπτωσης της αλυσίδας* σύμφωνα με το οποίο η συγχώνευση δύο συστάδων μπορεί να συμβεί αρκεί δύο μόνο στοιχεία τους να είναι κοντά το ένα με το άλλο ενώ το σύνολο των υπολοίπων στοιχείων της κάθε συστάδας να απέχουν πολύ μεταξύ τους.

Στην τεχνική του πλήρους συνδέσμου, ο αλγόριθμος ψάχνει για κλίκες (cliques - μέγιστος γράφος στον οποίο υπάρχει ακμή μεταξύ δύο κορυφών) και έτσι οι συστάδες που προκύπτουν είναι πιο συμπαγείς από αυτές του απλού συνδέσμου. Οι διαφορές μεταξύ των αλγορίθμων αυτών έγκειται στον τρόπο με τον οποίο ορίζουν την ομοιότητα μεταξύ στοιχείων και κατά συνέπεια των συστάδων πριν τη συγχώνευσή τους. Στην περίπτωση του απλού συνδέσμου η απόσταση μεταξύ δύο συστάδων είναι η ελάχιστη από τις αποστάσεις μεταξύ όλων των ζευγών στοιχείων από τις δύο συστάδες (κάθε ζεύγος περιέχει ένα στοιχείο από τη μία συστάδα και ένα από το άλλη). Παρόλα αυτά ο αλγόριθμος απλού συνδέσμου είναι αρκετά ευέλικτος σε δύσκολες περιπτώσεις.

Στην τεχνική μέσου συνδέσμου, δύο συστάδες συγχωνεύονται αν η μέση απόσταση μεταξύ οποιωνδήποτε δύο σημείων των συστάδων είναι μικρότερη ή ίση με την απόσταση κατωφλίου.

Στα βασικά μειονεκτήματα των συσσωρευτικών αλγορίθμων συγκαταλέγονται η υψηλή υπολογιστική πολυπλοκότητα τους που παρουσιάζεται λόγω της επαναληπτικής τους διαδικασίας, το ότι δεν είναι αυξητικοί αλγόριθμοι δηλαδή αν προστεθούν νέα στοιχεία ή αφαιρεθούν ή τροποποιηθούν τα ήδη υπάρχοντα θα πρέπει να τρέξει πάλι ο αλγόριθμος και τέλος τυχόν λάθη στα αρχικά βήματα του αλγορίθμου επηρεάζουν το τελικό αποτέλεσμα (αρχική διάδοση λάθους).

3.2.3.2 Διαμεριστικοί αλγόριθμοι

Στη διαμεριστική συσταδοποίηση, οι προτεινόμενες αλγοριθμικές προσεγγίσεις χρησιμοποιούν τον επιθυμητό αριθμό συστάδων για να καταλήξουν σε ένα τελικό σύνολο συστάδων. Ως είσοδο αυτών των αλγορίθμων δίδονται τα δεδομένα μαζί με τον αριθμό των συστάδων, K , που επιθυμούμε να λάβουμε στην έξοδο. Ο αριθμός αυτός καθορίζεται κυρίως από την προσπάθεια βελτιστοποίησης του αλγορίθμου. Ο

αλγόριθμος εκτελείται για διάφορους αριθμούς από συστάδες και επιλέγεται εκείνη η τελική κατάσταση η οποία βελτιστοποιεί μια συνάρτηση. Το κριτήριο που κυρίως χρησιμοποιείται σε διαμεριστικούς αλγορίθμους για την τελική απόφαση του αριθμού των συστάδων βασίζεται σε μέτρα ποιότητας, όπως μετρικές (π.χ. μέση απόσταση μεταξύ συστάδων, τετραγωνικό σφάλμα απόστασης σημείων της συστάδας από το κέντρο της) ή σε συναρτήσεις κριτηρίων, για τον προσδιορισμό της καταλληλότητας των προτεινόμενων λύσεων.

Η πολυπλοκότητα των διαμεριστικών αλγορίθμων είναι πολύ υψηλή λόγω του πλήθους των πιθανών λύσεων. Για να αποφευχθεί αυτό το πρόβλημα, οι περισσότεροι διαμεριστικοί αλγόριθμοι κάνουν αναζήτηση μόνο σε ένα μικρό υποσύνολο όλων των συστάδων χρησιμοποιώντας τεχνικές για τον εντοπισμό των σημαντικών συστάδων.

3.2.3.3 Αλγόριθμος K-μέσων

Ο αλγόριθμος K-means διαμερίζει το πολυεπίπεδο που δημιουργούν τα δεδομένα σε περιοχές και αντιστοιχεί μια περιοχή σε κάθε ομάδα. Η ελαχιστοποίηση του αθροίσματος των τετραγώνων των αποστάσεων όλων των στοιχείων μίας κλάσης από το κέντρο της κλάσης αυτής είναι το πρόβλημα που ο αλγόριθμος K-Means καλείται να επιλύσει δηλαδή αναζητούμε να προσδιορίσουμε το :

$$\min \sum_{x \in S_j(k)} |x - z_j|^2 \quad (3.4)$$

όπου $S_j(k)$ είναι η κλάση στην k επανάληψη, z_j είναι το κέντρο της κλάσης και το $|\cdot|$ εκφράζει την ευκλείδεια μετρική.

Η συνάρτηση που πρέπει να ελαχιστοποιηθεί για όλες τις κλάσεις θα είναι η :

$$J(Z) = \sum_j \sum_{x \in S_j(k)} |x - z_j|^2 \quad (3.5)$$

Η βέλτιστη λύση θα δοθεί από το μηδενισμό της πρώτης παραγώγου, οπότε έχουμε:

$$\frac{\partial J(z)}{\partial z_i} = \sum_{x \in S_j(k)} (x - z_j) = 0, \forall j \Rightarrow z_j = \frac{1}{n_j} \sum_{x \in S_j(k)} x \quad (3.6)$$

όπου είναι ο αριθμός των στοιχείων της κλάσης $S_j(k)$.

Έτσι στο K-means αλγόριθμο αρχικά καθορίζεται το πλήθος των κλάσεων, έστω K , στις οποίες τα δεδομένα πρόκειται να κατηγοριοποιηθούν και έπειτα επιλέγονται τυχαία ή προσεγγιστικά K στοιχεία τα οποία θα αποτελέσουν τα κέντρα των κλάσεων. Για τα υπόλοιπα $M - K$ στοιχεία προσδιορίζεται η απόστασή τους από τα κέντρα των K κλάσεων και τοποθετούνται στην κατάλληλη κλάση έχοντας ως κριτήριο την ελάχιστη απόσταση. Έπειτα, υπολογίζονται ξανά τα κέντρα των κλάσεων (δηλαδή η μέση τιμή των στοιχείων κάθε κλάσης) καθώς και οι αποστάσεις των στοιχείων από τα κέντρα των κλάσεων και γίνεται επανατοποθέτηση των στοιχείων ώσπου τα κέντρα των κλάσεων να μην έχουν καμιά μεταβολή, οπότε ο αλγόριθμος συγκλίνει.

Η ουσιαστική λειτουργία του αλγορίθμου είναι η ελαχιστοποίηση του αθροίσματος των αποστάσεων των παρατηρήσεων από τα κέντρα των ομάδων που ανήκουν και η δημιουργία ομάδων με παρόμοιο αριθμό παρατηρήσεων Το μειονέκτημα του αλγορίθμου είναι ότι εξαρτάται από τις αρχικές τιμές οι οποίες, αν δεν είναι κατάλληλες, μπορεί να οδηγήσουν σε διαφορετική ομαδοποίηση. Η βελτίωση του

αλγόριθμοι επιτυγχάνεται με την επιλογή των αρχικών κέντρων των συστάδων με μέγιστη απόσταση μεταξύ τους. Ένας τρόπος αποφυγής της υλοποίησης μεγάλου αριθμού διαφορετικών ομαδοποιήσεων είναι η μελέτη της λύση που ήδη διαθέτουμε, προσπαθώντας να ενώσουμε ή να χωρίσουμε ομάδες που θα μπορούσαν να βελτιώσουν τη μέση απόσταση των παρατηρήσεων από το κέντρο της ομάδας που ανήκουν. Η βέλτιστη λύση μπορεί να επιτευχθεί με δοκιμές στην επιλογή διαφορετικού αριθμού ομάδων. Πρέπει να σημειωθεί ότι δεν είναι απαραίτητος ένας μεγάλος αριθμός επαναλήψεων καθώς η βασική δομή της ομαδοποίησης διαφαίνεται από τις πρώτες επαναλήψεις. Στις υπόλοιπες επαναλήψεις πραγματοποιούνται μετακινήσεις των παρατηρήσεων που βρίσκονται ανάμεσα σε δύο ομάδες.

3.2.3.4 Αλγόριθμος k-πλησιέστερου γείτονα

Η βασική ιδέα της μεθόδου του k-πλησιέστερου γείτονα έγκειται στο ότι κάθε απόφαση για κατηγοριοποίηση ενός δεδομένου x λαμβάνεται αποκλειστικά με βάση τις ετικέτες των k πλησιέστερων γειτόνων του. Η αναζήτηση σύμφωνα με τον κανόνα του k πλησιέστερου γείτονα ξεκινάει από το σημείο ελέγχου x και απλώνεται σφαιρικά ώπου να περιλάβει k δείγματα εκπαίδευσης. Το σημείο ελέγχου παίρνει την ετικέτα που έχει η πλειοψηφία αυτών των δειγμάτων. Ας θεωρήσουμε c_1 και c_2 κλάσεις με N_1 και N_2 διανύσματα εκπαίδευσης. Αρχικά επιλέγεται η τιμή του k και με κέντρο το \vec{x} αυξάνουμε τον όγκο, ώστε να περιλάβει k δείγματα από το c_1 και k δείγματα από το c_2 . Έστω οι αντίστοιχοι όγκοι είναι V_1 και V_2 .

Οπότε έχουμε: αν ισχύει $P(c_1)(k/N_1 \cdot V_1) > P(c_2)(k/N_2 \cdot V_2)$, το $x \in c_1$ αλλιώς αν $P(c_1)(k/N_1 \cdot V_1) < P(c_2)(k/N_2 \cdot V_2)$, το $x \in c_2$

Για να καθοριστεί πλήρως ο k-NN αλγόριθμος, του οποίου τα δεδομένα θεωρούνται πως ανήκουν στον n -διάστατο χώρο R_n , πρέπει να ορισθεί η απόσταση μεταξύ των κλάσεων, όπου συνήθως επιλέγεται η Ευκλείδεια απόσταση. Επιπλέον πρέπει να διαλέξουμε τη βέλτιστη τιμή του k , δηλαδή του αριθμού των πιο κοντινών στο x παρατηρήσεων.

Η μέθοδος του πλησιέστερου γείτονα είναι μια παραλλαγή της μεθόδου του k-πλησιέστερου γείτονα. Ας θεωρήσουμε c_1 και c_2 κλάσεις με N_1 και N_2 διανύσματα εκπαίδευσης. Αρχικά επιλέγεται η τιμή του k να είναι περιττή για δύο κλάσεις. Με κέντρο το x , αυξάνεται ο όγκος της περιοχής ώπου να περιλάβει συνολικά k διανύσματα εκπαίδευσης. Έπειτα, μετρίεται το σύνολο των διανυσμάτων εκπαίδευσης που ανήκουν σε κάθε μία από τις δύο κλάσεις: έστω k_1 ανήκουν στην c_1 και k_2 ανήκουν στην c_2 . Αν $k_1 > k_2$, το x ανήκει στην κλάση c_1 αλλιώς αν $k_1 < k_2$, το x ανήκει στην κλάση c_2 .

3.3 Προτεινόμενη τεχνική εξόρυξης γνώσης

Η εξαγωγή γνώσης από πολυδιάστατα δεδομένα θεωρείται ανοικτό ερευνητικό θέμα, γεγονός που οδήγησε στην ενασχόληση με την ανίχνευση προτύπων σε πολυδιάστατα δεδομένα, τα οποία δεν έχουν υποστεί καμιά επεξεργασία, ενώ δεν έχει διευκρινισθεί το πλήθος των διακριτών προτύπων σε αυτά ούτε και η σχέση μεταξύ των χαρακτηριστικών εισόδου.

Ο προτεινόμενος αλγόριθμος είναι μια επέκταση της ιεραρχικής ομαδοποίησης και είναι βασισμένος σε μια ασαφή επιλογή χαρακτηριστικών που λαμβάνονται υπόψη κατά

τη σύγκριση των δεδομένων καθώς και στην ασαφοποίηση αυτών. Τα αποτελέσματα της αρχικής ομαδοποίησης βελτιώνονται μέσω ενός βήματος επαναταξινόμησης χωρίς επίβλεψη, το οποίο είναι βασισμένο στις αρχές του ιεραρχικού αλγορίθμου συσταδοποίησης και συμβάλλει στην πειραματική αξιολόγηση της αποδοτικότητας της μεθόδου.

3.3.1 Επιλογή χαρακτηριστικών και καθορισμός απόστασης

Η εξέταση των δεδομένων οδηγεί στην επιλογή χαρακτηριστικών με σκοπό τη δημιουργία ομοιογενών ομάδων. Η επιλογή, όμως, απολύτων (σαφών) χαρακτηριστικών δεν είναι πάντα δυνατή αφού το κάθε χαρακτηριστικό μπορεί να επηρεάσει το κάθε στοιχείο του συνόλου κατά ορισμένο βαθμό. Κάθε σύνολο αποτελείται από στοιχεία των οποίων κάποια χαρακτηριστικά συσχετίζονται, άρα η απόσταση μεταξύ τους είναι μικρή, ενώ κάποια άλλα όχι, άρα η απόσταση μεταξύ τους είναι μεγάλη [141]. Για τον καθορισμό μιας ομάδας εκμεταλλευόμαστε αυτήν τη διαφορά που παρουσιάζεται στην κατανομή των τιμών απόστασης.

Αναλυτικότερα, έστω c_1 και c_2 δύο ομάδες στοιχείων, f_i , $i \in N_F$ η μετρική που συγκρίνει το i -οστο χαρακτηριστικό και F το πλήθος των χαρακτηριστικών (η διάσταση του χώρου εισόδου). Ένα μέτρο απόστασης μεταξύ των δύο ομάδων, θεωρώντας μόνο το i -στο χαρακτηριστικό, δίδεται από τη σχέση:

$$f_i(c_1, c_2) = \sqrt[k]{\sum_{a \in c_1, b \in c_2} \frac{r_i(a_i, b_i)^k}{|c_1||c_2|}} \quad (3.7)$$

όπου a_i , b_i είναι το i -οστο χαρακτηριστικό των στοιχείων a , b , αντίστοιχα, $|c|$ είναι η πληθικότητα της ομάδας c και $k \in \mathcal{R}$ είναι μια σταθερά που σχετίζεται με το πλήθος των επαναλήψεων ενός ιεραρχικού αλγορίθμου που χρειάστηκαν για τη δημιουργία των συγκεκριμένων ομάδων στοιχείων. Μια ασαφής επιλογή από χαρακτηριστικά μπορεί να μας δώσει τη δυνατότητα δημιουργίας διαφορετικού πλαισίου δεδομένων. Η απόσταση ανάμεσα στα χαρακτηριστικά των στοιχείων των ομάδων που συσχετίζονται είναι μικρή, οπότε, αν βρούμε το πλαίσιο το οποίο παράγει τη μικρότερη συνολική απόσταση ανάμεσα στις ομάδες, μπορούμε να θεωρήσουμε ότι έχουμε βρει τη βέλτιστη λύση, δηλαδή έχουμε ομαδοποιήσει τα δεδομένα μας κατά το βέλτιστο τρόπο. Αναλυτικότερα στοιχεία για τη μέθοδο δίνονται στην [141].

Έστω x , ένα ασαφές σύνολο ορισμένο στο N_F , με βαθμωτή πληθικότητα ίση με 1. Τότε η συνολική απόσταση μεταξύ c_1 και c_2 υπολογίζεται από τη σχέση:

$$d(c_1, c_2) = \sum_{i \in N_F} x_i (c_1, c_2)^\lambda f_i(c_1, c_2) \quad (3.8)$$

όπου x_i είναι ο βαθμός στον οποίο το i , άρα και το f_i , περιέχεται στο πλαίσιο, $i \in N_F$ και $\lambda \in \mathcal{R}$ είναι μία σταθερά. Τυπική τιμή για το λ είναι 2. Βασιζόμενη στο ότι τα στοιχεία με συσχετιζόμενα χαρακτηριστικά απέχουν μικρή απόσταση, επιλέγουμε τις τιμές του διανύσματος x μέσω της ελαχιστοποίησης της απόστασης d . Τα χαρακτηριστικά που συσχετίζουν τα c_1 και c_2 είναι "περισσότερο πιθανά" από αυτά που παράγουν τις μικρότερες αποστάσεις f_i .

3.3.2 Βελτίωση συσταδοποίησης μέσω κατηγοριοποίησης

Η εξαγωγή γνώσης από πολυδιάστατα δεδομένα σχετίζεται με πολλά ερευνητικά πεδία και θεωρείται ένα ανοικτό ερευνητικό θέμα. Σε αυτό το πλαίσιο ασχοληθήκαμε με την ανίχνευση προτύπων σε πολυδιάστατα δεδομένα τα οποία δεν έχουν υποστεί καμιά επεξεργασία καθώς και δεν έχει διευκρινισθεί ούτε το πλήθος των διακριτών προτύπων στα δεδομένα αλλά ούτε και η σχέση μεταξύ των χαρακτηριστικών εισόδου.

Σε αυτήν την ενότητα θα παρουσιαστεί η προτεινόμενη τεχνική, στόχος της οποίας είναι η βέλτιστη αντιμετώπιση κλασικών προβλημάτων του ιεραρχικού αλγορίθμου, όπως είναι η κατάρα της διαστατικότητας (curse of dimensionality) και η αρχική διάδοση λάθους (initial error propagation), ώστε να αυξηθεί η ευρωστία της τεχνικής. Η βελτίωση αυτή επιτυγχάνεται με τη χρήση ενός διαμεριστικού αλγορίθμου σε ένα τμήμα των δεδομένων.

Ο αλγόριθμος που υιοθετήθηκε, έχει προταθεί στην [141] και είναι βασισμένος στην ασαφή επιλογή των χαρακτηριστικών που λαμβάνονται υπόψη κατά τη σύγκριση των δεδομένων. Τα αποτελέσματα αυτής της αρχικής ομαδοποίησης βελτιώνονται μέσω ενός βήματος επαναταξινόμησης χωρίς επίβλεψη, το οποίο είναι βασισμένο στις αρχές του ιεραρχικού αλγορίθμου και το οποίο συμβάλλει στην πειραματική αξιολόγηση της αποδοτικότητας της μεθόδου.

Αρχικά, χρησιμοποιώντας την τεχνική που παρουσιάστηκε στην Ενότητα 3.3.1, προχωρούμε σε μια αποδοτικότερη αντιμετώπιση του προβλήματος ελαττώνοντας τις διαστάσεις των δεδομένων εισαγωγής και δημιουργώντας χαρακτηριστικά διανύσματα ικανά να οδηγήσουν σε σαφή συσταδοποίηση.

Έπειτα, χρησιμοποιώντας την αρχή που αναφέρει ότι όταν τα χαρακτηριστικά ενός συνόλου συσχετίζονται, η απόσταση μεταξύ τους είναι μικρή [141], πραγματοποιούμε μια ανάλυση συσχέτισης των χαρακτηριστικών. Για να το πετύχουμε αυτό, εφαρμόζουμε κατηγοριοποίηση σε ένα τμήμα των δεδομένων και επεκτείνουμε τα αποτελέσματα σε ολόκληρο το σύνολο δεδομένων.

Επιλέξαμε να εργαστούμε με τον ταξινομητή του k -κοντινότερου γείτονα (kNN) (Ενότητα 3.2.3.2), αφού κρίθηκε, μετά από εφαρμογή του, αποδοτικότερος από άλλους ταξινομητές που χρησιμοποιήθηκαν. Ο kNN έχειδειχθεί ότι είναι αποτελεσματικός ακόμη και για στιγμιότυπα που αντιστοιχούν σε σημεία του \mathcal{R}^n . Ο ταξινομητής του k -κοντινότερου γείτονα εξετάζει όλες τις ιδιότητες από όλα τα στοιχεία, το οποίο κρίνεται ως μειονέκτημα για μεγάλο σύνολο δεδομένων. Αυτό μας οδήγησε στην εφαρμογή μιας αρχικής διαδικασίας συσταδοποίησης επιλέγοντας μόνο ένα μικρό υποσύνολο από τα δείγματα δεδομένων.

Συγκεκριμένα, ο αλγόριθμος kNN υποθέτει ότι όλα τα στοιχεία αντιστοιχούν σε σημεία στο n -διάστατο χώρο \mathcal{R}^n . Με τη χρήση του σταθμισμένου μέτρου απόστασης (Τύπος 3.8) καθορίζονται οι γείτονες ενός στοιχείου. Κάθε απόφαση για ταξινόμηση ενός στοιχείου λαμβάνεται αποκλειστικά με βάση τις ετικέτες των k πλησιέστερων γειτόνων του. Η αναζήτηση σύμφωνα με τον κανόνα του k πλησιέστερου γείτονα ξεκινάει από το σημείο ελέγχου x και μεγαλώνει σφαιρικά, ώσπου να περιλαμβάνει k δείγματα εκπαίδευσης. Το στοιχείο παίρνει την ετικέτα που έχει η πλειοψηφία αυτών των δειγμάτων. Με αυτήν την τεχνική, το υποσύνολο των δεδομένων που χρησιμοποιήθηκε κατηγοριοποιήθηκε σε ένα επιθυμητό σύνολο κλάσεων.

Το επόμενο βήμα της προτεινόμενης μεθόδου είναι η συσταδοποίηση όλων των δεδομένων. Αυτό πραγματοποιείται με τη χρήση του ιεραρχικού αλγορίθμου. Εάν η αρχική κατηγοριοποίηση (με τη χρήση του αλγορίθμου kNN) είναι επιτυχής στην αποκάλυψη των προτύπων που αντιπροσωπεύουν τα δεδομένα, τότε αυτή η διαδικασία

θα βελτιώσει την έξοδο και το ποσοστό συσταδοποίησης με την απομάκρυνση μερικών μελών συστάδων που προήλθαν ως αποτέλεσμα λαθών κατά τα αρχικά βήματα.

Κατά συνέπεια, αυτή η διαδικασία προσφέρει μία ένδειξη της αληθινής απόδοσης του ιεραρχικού αλγορίθμου. Επιπλέον, καθιστά το συνολικό αλγόριθμο πιο εύρωστο, σε αντίθεση με τον απλό ιεραρχικό αλγόριθμο, δεδομένου ότι είναι πιο ανθεκτικός στα λάθη που μπορεί να συμβούν στα αρχικά βήματα.

3.4 Προσωποποιημένη πρόσβαση και συστήματα μαθησιακής τεχνολογίας

3.4.1 Κατηγοριοποίηση χρηστών βασισμένη σε ερωτηματολόγια

Η εφαρμογή της προτεινόμενης τεχνικής αφορά σε ένα σύστημα ηλεκτρονικής μάθησης (Ενότητα 7.2) συμβάλλοντας στην κατηγοριοποίηση των χρηστών του. Αυτή η κατηγοριοποίηση μπορεί να χρησιμοποιηθεί για πολλούς εκπαιδευτικούς σκοπούς, όπως για παροχή εκπαιδευτικού υλικού κατάλληλου για το γνωστικό επίπεδο του κάθε χρήστη ή για προσαρμογή του περιβάλλοντος μάθησης στις ανάγκες του χρήστη.

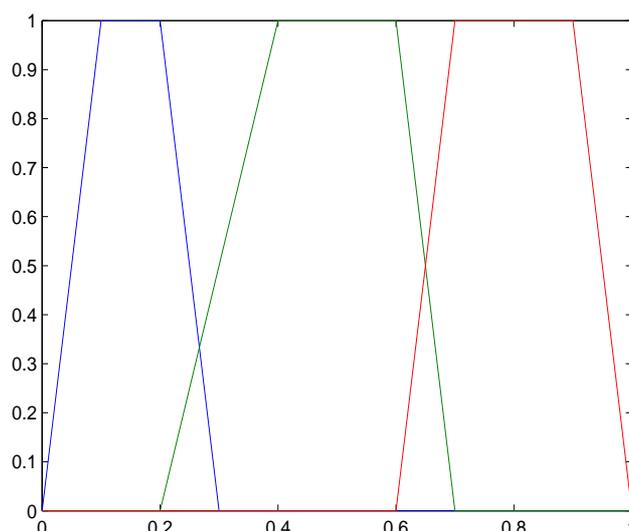
3.4.1.1 Στατική εξαγωγή προφίλ χρηστών

Όπως αναφέρθηκε στην Ενότητα 2.5, το σύστημα μαθησιακής τεχνολογίας εκτός από τους χρήστες του, περιλαμβάνει και μια ομάδα ειδικών σε θέματα σχετικά με το μαθησιακό αντικείμενο που πραγματεύεται. Αυτή η ομάδα είναι υπεύθυνη για τη δημιουργία και λειτουργία των ηλεκτρονικών ερωτηματολογίων. Επιπλέον, η ομάδα αυτή αποδίδει την κατάλληλη βαθμονόμηση σε κάθε πιθανή απάντηση και εξετάζει περαιτέρω τις απαντήσεις βασιζόμενη στη γνώση και την εμπειρία της με σκοπό να καθορίζει το σύνολο των συστάδων - αρχικών, στατικών προφίλ στα οποία θα κατηγοριοποιηθούν οι χρήστες.

Αναλυτικότερα, τα ερωτηματολόγια αυτά περιλαμβάνουν ερωτήσεις κειμένου (text), πολλαπλών επιλογών (multiple choices), σωστό/λάθος (true/false) και μοναδικής επιλογής (radio button). Η κάθε ερώτηση καταγράφεται σε βάση δεδομένων λαμβάνοντας ένα συγκεκριμένο διαγνωστικό (id), ενώ οι πιθανές απαντήσεις αποθηκεύονται στη βάση μαζί με το μοναδικό τους διαγνωστικό καθώς και το διαγνωστικό του χρήστη.

Ανάλογα με τον τύπο των ερωτήσεων οι ειδικοί αποδίδουν την κατάλληλη βαρύτητα σε κάθε απάντηση. Ερωτήσεις του τύπου σωστό/λάθος ή μοναδικής επιλογής περιέχουν σαφήνεια και έτσι οι ειδικοί αποδίδουν σαφή βαθμονόμηση σε κάθε απάντηση αυτής της κατηγορίας. Για παράδειγμα, η ερώτηση "Έχετε ηλεκτρονικό υπολογιστή στο σπίτι;" με πιθανές απαντήσεις (α) όχι, (β) ναι, απέκτησα πρόσφατα και (γ) ναι έχει βαθμονομηθεί ως εξής: η απάντηση (α) 1 που αντιστοιχεί στην δεύτερη κατηγορία προφίλ, τους αρχάριοι χρήστες, η (β) 2 που αντιστοιχεί στη τρίτη κατηγορία προφίλ, τους προηγμένους χρήστες και η (γ) 3 που αντιστοιχεί στην πρώτη κατηγορία προφίλ, τους ειδήμονες χρήστες.

Αντίθετα, οι ερωτήσεις τύπου κειμένου και πολλαπλών επιλογών περιέχουν ασάφεια και για αυτό είναι απαραίτητο η κάθε απάντηση να μπορεί να εκφράζεται με ασαφή τρόπο, με τη βοήθεια της ασαφούς θεωρίας συνόλων [151]. Για παράδειγμα, στην ερώτηση "B.2.4. Για ποιο από τα παρακάτω χρησιμοποιήσατε τον Η/Υ τουλάχιστον μια φορά στο μήνα που πέρασε;" η οποία είναι μια ερώτηση πολλαπλών



Σχήμα 3.2: Συνάρτηση συμμετοχής: απόδοση ασάφεια σε πιθανό σύνολο απαντήσεων

επιλογών, όπου ο χρήστης μπορεί να δώσει τις ακόλουθες πέντε απαντήσεις: α) Πρόσβαση στο Διαδίκτυο, β) Λήψη/αποστολή ηλεκτρονικών μηνυμάτων, γ) Παιχνίδια, δ) Επεξεργασία κειμένου και ε) Εκπαιδευτικό λογισμικό.

Οι ειδικοί ορίζουν μία συνάρτηση συμμετοχής για κάθε ερώτηση που οι απαντήσεις της περιέχουν ασάφεια. Η συνάρτηση συμμετοχής (Σχήμα 3.2) καθορίζει το βαθμό στον οποίο κάθε σημείο του πεδίου ορισμού διαθέτει μια συγκεκριμένη ιδιότητα. Έτσι, ένα ασαφές σύνολο A του X αντιστοιχεί κάθε τιμή x του X σε ένα πραγματικό αριθμό στο διάστημα $[0, 1]$, με μία τιμή της συνάρτησης συμμετοχής $f_A(x)$ στο x αντιπροσωπεύοντας το βαθμό συμμετοχής του x στο A .

Βασιζόμενοι στην ασαφή λογική, οι ειδικοί ορίζουν την κατάλληλη ασαφή τμηματοποίηση (partition) - μια σειρά από ασαφή σύνολα - του X και συνδέουν κάθε ασαφές σύνολο της τμηματοποίησης με ένα από τα προφίλ, όπως αυτό καθορίζεται από το κατάλληλο ασαφές σύνολο.

Πιο αναλυτικά, στο Σχήμα 3.2 το πρώτο ασαφές σύνολο της τμηματοποίησης αντιστοιχεί στο προφίλ A (προφίλ ειδημόνων), το δεύτερο στο προφίλ B (προφίλ αρχαρίων) και το τρίτο στο προφίλ C (προφίλ προχωρημένων). Αν, για παράδειγμα, ένας χρήστης έχει δώσει ως απάντησεις την α) και τη γ) στην προαναφερθείσα ερώτηση την τιμή τότε η συνάρτηση συμμετοχής (στην προκειμένη περίπτωση έχει χρησιμοποιηθεί η τραπεζοειδής συνάρτησης) δίνει 0.7 βαθμό συμμετοχής του x στο προφίλ A και 0.3 στο προφίλ B (Σχήμα 3.2).

Στην εφαρμογή που περιγράφεται στο [133], τα ερωτηματολόγια (Παράρτημα Β) απευθύνονται σε εκπαιδευτικούς με σκοπό την αξιολόγησή τους σε θέματα νέων τεχνολογιών και την κατηγοριοποίησή τους με σκοπό να τους δοθεί κατάλληλο εκπαιδευτικό υλικό (Ενότητα 7.2).

Η συλλογή των δεδομένων από τους χρήστες του συστήματος γίνεται με τη χρήση ηλεκτρονικών ερωτηματολογίων. Οι απαντήσεις στα ερωτηματολόγια είναι τα δεδομένα του συστήματος τα οποία χαρακτηρίζονται από πολυάριθμα, μετρήσιμα χαρακτηριστικά γνωρίσματα, οπότε είναι δυνατό να καθοριστούν περισσότερες από μία μετρικές ομοιότητας. Μια αποδοτική αντιμετώπιση του προβλήματος είναι η ελάττωση των διαστάσεων εισαγωγής [145], που μπορεί να πραγματοποιηθεί με το να αγνοηθούν

Χαρακτηριστικό γνώρισμα	Αναγνωριστικό ερώτησης	Περιγραφή
1	125	Are you a teacher dedicated to or working in Special Education Needs?
2, 3, 4	126, 127, 128	Qualification/training in Special Education Needs?
5	130	Teaching Experience
6	133	Do you have a computer at home?
7	134	Do you have access to the Internet from your home?
8	139	How often do you personally use your Internet connection at home?
9, 10, 11, 12, 13	141, 142, 143, 144, 145	For which of the following did you use the computer at least once in the past month?
14, 15, 16, 17, 18, 19, 20	147, 148, 149, 150, 151, 152, 153	Which of the following tasks have you performed at least once, without any help?
21	155	Are there any computers in your work environment?
22, 23	300, 301	How often did you use the computer last week at the school?
24, 25	168, 169	Do you have access to the Internet or educational software in your work environment?
26, 27	174, 175	In your teaching, how many hours a week, on average, do you use the Internet or educational software with your students?
28, 29	176, 177	Do you use the Internet for search and retrieval of information relating to the needs and problems faced by SEN students?
30, 31	182, 183	Do you use the Internet from the school in order to find additional sources of educational material?
32, 33	400, 401	Do you use the Internet to connect with other schools?
34, 35, 36, 37	261, 262, 264, 265	Is your post permanent - temporary?
38, 39, 40, 41, 42, 43	267, 268, 270, 271, 273, 274	Age of your students
44	275	Area served by your school

Πίνακας 3.1: Επιλογή χαρακτηριστικών γνωρισμάτων

μερικά από τα διαθέσιμα χαρακτηριστικά γνωρίσματα/απαντήσεις.

Για να αγνοήσουμε δεδομένα πρέπει να είναι γνωστή η σχέση εξάρτησής τους. Αν αυτά είναι ανεξάρτητα μεταξύ τους, τότε μπορεί να οδηγηθούμε σε απώλεια πληροφορίας. Με τη χρήση των ερωτηματολογίων, η σχέση των χαρακτηριστικών γνωρισμάτων (απαντήσεων) είναι γνωστή εκ των προτέρων, άρα η μείωση των διαστάσεων των δεδομένων είναι εφικτή. Η περαιτέρω εξέταση των δεδομένων οδηγεί στην επιλογή χαρακτηριστικών με σκοπό τη δημιουργία ομοιογενών ομάδων. Ένα χαρακτηριστικό όμως μπορεί να είναι περισσότερο σημαντικό από άλλα, ενώ όλα τα χαρακτηριστικά είναι περισσότερο ή λιγότερο χρήσιμα, οπότε η επιλογή απόλυτων (σαφών) χαρακτηριστικών δεν είναι πάντα δυνατή.

Τα χαρακτηριστικά γνωρίσματα τα οποία αρχικά μπορούσαμε να εξάγουμε από τα ερωτηματολόγιο χωρίς περαιτέρω επεξεργασία ήταν 176. Έπειτα όμως από τη μείωση των διαστάσεων που πραγματοποιήσαν οι ειδικοί τα χαρακτηριστικά μειώθηκαν στα 44 (Πίνακας 3.1), δημιουργώντας ένα διάγραμμα ίσου μήκους το οποίο αντιστοιχούσε σε κάθε χρήση.

Αν και τα παραπάνω σημαντικά χαρακτηριστικά γνωρίσματα αποτέλεσαν μόνο ένα 25%, αυτό αποδείχθηκε επαρκές και αποδοτικό κατά τη διαδικασία. Στον

Πίνακα 3.1 εμφανίζονται τα χαρακτηριστικά γνωρίσματα (πρώτη στήλη, Πίνακας 3.1) και το αντίστοιχο αναγνωριστικό της ερώτησης (δεύτερη στήλη, Πίνακας 3.1) του ηλεκτρονικού ερωτηματολογίου από όπου εξάγονται, ενώ η λεκτική περιγραφή των ερωτήσεων παρουσιάζεται στην τρίτη στήλη του Πίνακα 3.1.

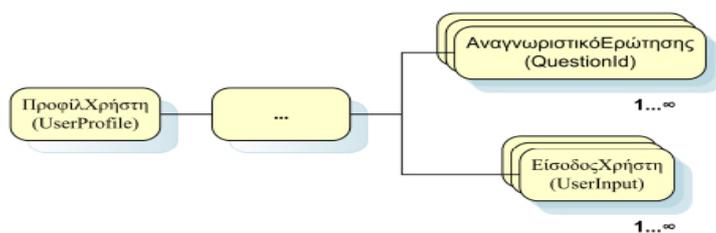
Οι ειδικοί, έχοντας επιλέξει τα χαρακτηριστικά γνωρίσματα των δεδομένων, είναι σε θέση, βασιζόμενοι στην εμπειρία τους να καθορίζουν ένα σύνολο χαρακτηρισμών χρηστών (στην προκειμένη περίπτωση τριών - Ειδημόνων, Αρχαρίων, Προηγμένων), που διαμορφώνουν μια στατική αντιπροσώπευση των προφίλ τους. Αυτοί οι αρχικοί χαρακτηρισμοί χρηστών χρησιμοποιούνται επιπρόσθετα και σε μεταγενέστερο στάδιο, κατά τη διάρκεια της ευφυούς διαδικασίας εξαγωγής των προφίλ και παρέχουν ένα στέρεο σημείο αναφοράς, αφού, αν και είναι στατικοί, παράγονται στην πραγματικότητα αυτόματα από το σύστημα, σύμφωνα με τους αρχικούς ορισμούς τους. Η βάση της εξαγωγής τους στηρίζεται στις πληροφορίες που παρέχονται από τα δεδομένα εισόδου των χρηστών, τα οποία λαμβάνονται από τα ηλεκτρονικά ερωτηματολόγια.

Τα ερωτηματολόγια αυτά, μαζί με τη βαθμονόμηση των απαντήσεών τους, αποθηκεύονται στην αποθήκη Πόροι Εκμάθησης της προτεινόμενης επέκτασης της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE (Σχήμα 2.5). Η διαδικασία Παράδοση Υλικού έχει εμπλουτιστεί με μηχανισμό αυτόματης δημιουργίας ηλεκτρονικών ερωτηματολογίων, ικανό να συλλέξει τα δεδομένα των χρηστών. Η επεξεργασία των δεδομένων πραγματοποιείται μέσω της προτεινόμενης μεθόδου η οποία έχει ενσωματωθεί στη διαδικασία Αξιολόγηση (Σχήμα 2.5) και τα αποτελέσματα της φυλάσσονται στην αποθήκη Προφίλ Χρηστών.

Καθώς αρχικοποιείται η διαδικασία εξαγωγής των προφίλ, όλα τα σχετικά προφίλ αποθηκεύονται μέσα σε μια ενιαία, κεντρική δομή αντιστοίχισης, της οποίας το αφηρημένο μοντέλο παρουσιάζεται στο Σχήμα 3.3. Το UserProfile αντιστοιχείται με τις πληροφορίες που ανακτώνται είτε από το ίδιο το εκάστοτε ηλεκτρονικό ερωτηματολόγιο, είτε άμεσα από την εισαγωγή δεδομένων, λεκτικών ή μη, των χρηστών. Εδώ, το πρώτο στοιχείο (QuestionID) φυλάσσει όλες τις πληροφορίες που απαιτούνται για τον προσδιορισμό της ερώτησης του ηλεκτρονικού ερωτηματολογίου, καθώς επίσης και του τύπου της, στοχεύοντας στην καλύτερη κατανόηση και ταίριασμα του παραγόμενου προφίλ. Το δεύτερο στοιχείο (UserInput) περιέχει τις σχετικές με τα στοιχεία πληροφορίες χρηστών, όπως είναι οι απαντήσεις των χρηστών. Έτσι, όταν ένας χρήστης απαντά σε μια ερώτηση του ηλεκτρονικού ερωτηματολογίου εισαγωγής, ένας βαθμός σχετικότητας συνδέεται με αυτή και προσαρμόζεται στο συγκεκριμένο στοιχείο QuestionID του. Με τον τρόπο αυτό ο βαθμός αυτός διαδίδεται και στο στοιχείο UserProfile.

Καθώς όλο και περισσότερες απαντήσεις από τον τελικό χρήστη εισάγονται στη δομή UserProfile, καταχωρούνται και οι πρόσθετοι βαθμοί σχετικότητας στα αντίστοιχα στοιχεία QuestionID. Ανάλογα με την εκάστοτε ερώτηση, καθώς επίσης και με το τμήμα του ηλεκτρονικού ερωτηματολογίου στο οποίο ανήκει αυτή, διαδίδονται και διαφορετικοί βαθμοί. Το τελευταίο βήμα βασίζεται στη σύγκριση των παρεχόμενων αριθμητικών τιμών με τη σειρά των τιμών που συνδέονται εκ των προτέρων με τα προφίλ.

Ο προτεινόμενος μηχανισμός εξαγωγής προφίλ, ο οποίος ενσωματώνεται στη διαδικασία Αξιολόγησης του προτεινόμενου συστήματος, δημιουργεί, αναπροσαρμόζει και χρησιμοποιεί τα προφίλ των χρηστών του συστήματος, ταίριαζοντας συγκεκριμένες απαντήσεις από τα ηλεκτρονικά ερωτηματολόγια με στατικά προφίλ.



Σχήμα 3.3: Δομή αντιστοίχισης Προφίλ Χρηστών

Η σχεδίαση του μοντέλου των προφίλ διευκολύνει τόσο τη διαδικασία χρήσης των προτιμήσεων των χρηστών κατά τη δημιουργία των προφίλ, όσο και τη διαδικασία της ανίχνευσης/παρακολούθησης των προτιμήσεων αυτών. Επιπλέον, έχει σχεδιαστεί με έναν τρόπο που επιτρέπει την αυτοματοποιημένη εξαγωγή των προφίλ του χρήστη με βάση αυτές τις προτιμήσεις και το ιστορικό εισαγωγής δεδομένων των χρηστών.

Ο προτεινόμενος μηχανισμός εφαρμόζεται στο σύνολο των δεδομένων (απαντήσεων του χρήστη). Οι ειδικοί αποδίδουν σημασιολογικές ετικέτες στο τμήμα του συνόλου των υπό εξέταση δεδομένων. Με αυτό τον τρόπο, κάθε τελικός χρήστης ταξινομείται σε ένα αρχικό, στατικό προφίλ, που χαρακτηρίζει τη συμπεριφορά του, τα ενδιαφέροντά του και την περαιτέρω επεξεργασία/διαχείρισή του από το σύστημα. Αυτός ο ιδιαίτερος χαρακτηρισμός προφίλ αποτελεί τη βάση της επόμενης διαδικασίας συσταδοποίησης, η οποία περιλαμβάνει τη δυναμική εξαγωγή και την ολοκλήρωση του προφίλ του χρήστη.

3.4.2 Δυναμική εξαγωγή προφίλ χρηστών

Κατά τη διαδικασία δυναμικής εξαγωγής των προφίλ χρηστών, θεωρείται ότι δεν υπάρχει διαθέσιμο κανενός είδους χειρωνακτικό σημασιολογικό μαρκάρισμα. Έτσι, η πιθανότητα να προκύψουν διάφορα προφίλ, διαφορετικά από αυτά που προέκυψαν με χειρωνακτικό χαρακτηρισμό των ειδικών του συστήματος είναι αυξημένη. Άρα θα πρέπει να εφαρμοστεί τεχνική κατηγοριοποίησης, προκειμένου να αναγνωρισθούν τα κυρίαρχα προφίλ. Αρχικά, η προτεινόμενη τεχνική (Ενότητα 3.3) εφαρμόζεται σε ένα μέρος του συνόλου των δεδομένων και έπειτα επεκτείνεται σε ολόκληρο το σύνολο των δεδομένων. Η απόδοση της προτεινόμενης μεθοδολογίας συγκρίνεται τα αποτελέσματα της στατικής εξαγωγής προφίλ, χρησιμοποιώντας τους προκαθορισμένους χαρακτηρισμούς προφίλ ως μία εκ των προτέρων γνωστή σημασιολογική πληροφορία ετικετών.

Εφαρμόζοντας την προτεινόμενη τεχνική στα δεδομένα - απαντήσεις στα ερωτηματολόγια που λαμβάνει το σύστημα από τον εκάστοτε χρήστη, πραγματοποιείται η αντιστοίχιση των απαντήσεων του κάθε χρήστη σε συγκεκριμένη συστάδα. Οι συστάδες αυτές έχουν παραχθεί με την χρήση της προτεινόμενης μεθόδου 3.3. Αρχικά, κάθε μία από αυτές τις συστάδες αποτελεί και ένα προφίλ των χρηστών του συστήματος. Κάθε ένα από τα προφίλ περιγράφει τα χαρακτηριστικά των κατηγοριών των χρηστών καθώς και τις εκπαιδευτικές ανάγκες που κάθε κατηγορία μπορεί να έχει.

Η τεχνική αυτή εντάσσεται στην διαδικασία Αξιολόγησης όπως αυτή περιγράφηκε στην Ενότητα 2.5. Οι απαντήσεις του χρήστη στα ηλεκτρονικά ερωτηματολόγια χρησιμοποιούνται για την κατηγοριοποίησή του σε κατάλληλο προφίλ του συστήματος

ή τη δημιουργία νέου προφίλ κατάλληλου να περιγράψει την τρέχουσα κατάσταση του χρήστη.

Επιπλέον αυτές οι απαντήσεις μπορούν να χρησιμοποιηθούν για να εξαχθούν στατιστικές μετρήσεις που αφορούν στις απαντήσεις όλων των χρηστών στα συγκεκριμένα ερωτηματολόγια. Αναλυτικότερα, για την μορφή των ερωτηματολογίων και για την στατιστική του ανάλυση καθώς και την χρήση της κατηγοριοποίησης σε περιβάλλον ηλεκτρονικής μάθησης ο αναγνώστης μπορεί να βρει στην Ενότητα 7.2.

Κατά την αρχική φάση της προτεινόμενη τεχνική, πραγματοποιεί συσταδοποίηση σε ένα μόνο υποσύνολο 10% του συνόλου των δεδομένων και όχι σε όλο το σύνολο των δεδομένων. Κατά αυτό τον τρόπο η προτεινόμενη τεχνική είναι όχι μόνο υπολογιστικά αποδοτικότερη, αλλά και αποδοτικότερη συγκρινόμενη με την προσέγγιση της εφαρμογής της ιεραρχικής διαδικασίας συσταδοποίησης σε ολόκληρο το σύνολο των δεδομένων. Αν και η συσταδοποίηση του 10% του συνόλου δεδομένων οδήγησε σε διαφορετικούς αριθμούς πιθανών συστάδων, τα βέλτιστα αποτελέσματα επιτεύχθηκαν για έναν αριθμό από 9 συστάδες, όπως προκύπτει από την παρατήρηση των πινάκων 3.2, 3.3 και 3.4.

Σε αυτούς τους πίνακες παρουσιάζονται τα αποτελέσματα συσταδοποίησης 100 χρηστών για διαφορετικό πλήθος συστάδων (3, 5 και 9 συστάδες). Για παράδειγμα, στον Πίνακα 3.2 παρουσιάζονται τα αποτελέσματα συσταδοποίησης της περίπτωσης όπου ο ιεραρχικός αλγόριθμος συσταδοποίησης ολοκληρώθηκε, όταν έφθασε σε ένα κατώτατο όριο 3 συστάδων με 3 κατηγορίες ανά συστάδα (1η κατηγορία: Ειδήμονες, 2η κατηγορία: Αρχάριοι και 3η κατηγορία: Προηγμένοι χρήστες). Ομοίως, στους Πίνακες 3.3 και 3.4, ο αλγόριθμος ολοκληρώθηκε όταν έφτασε σε ένα κατώτερο όριο 5 συστάδων με 5 κατηγορίες ανά συστάδα και 9 συστάδων με 9 κατηγορίες ανά συστάδα, αντίστοιχα.

Για παράδειγμα, η πρώτη συστάδα του Πίνακα 3.2 αποτελείται από 17 χρήστες, από τους οποίους 2 ανήκουν στην πρώτη κατηγορία, 6 στη δεύτερη κατηγορία και 9 στην τρίτη κατηγορία. Η συστάδα αυτή εξουσιάζεται από την τρίτη κατηγορία κατά 52,94% αφού σε αυτήν έχει ταξινομηθεί το μεγαλύτερο ποσοστό των χρηστών που ανήκουν στην πρώτη συστάδα. Όμοια, η δεύτερη και η τρίτη συστάδα εξουσιάζεται πάλι από την τρίτη κατηγορία με ποσοστό 65,79% και 66,67% αντίστοιχα.

Συστάδα	Στοιχεία	Επί τοις εκατό
1η	(2, 6, 9)	(11.77%, 35.29%, 52.94%)
2η	(11, 2, 25)	(28.95%, 5.26%, 65.79%)
3η	(14, 1, 30)	(31.11%, 2.22%, 66.67%)

Πίνακας 3.2: Αποτελέσματα συσταδοποίησης για 100 χρήστες - 3 συστάδες

Στον Πίνακα 3.3 παρουσιάζονται τα αποτελέσματα συσταδοποίησης των ίδιων 100 χρηστών, χρησιμοποιώντας ένα νέο κατώτατο όριο 5 συστάδων ολοκληρώνοντας τον αλγόριθμο συσταδοποίησης 2 βήματα νωρίτερα. Όπως είναι εμφανές, οι νέες συστάδες περιέχουν λιγότερους χρήστες σε σχέση με την προηγούμενη περίπτωση. Πάλι βλέπουμε ότι σε όλες οι συστάδες επικρατεί η τρίτη κατηγορία προφίλ.

Τέλος, στον Πίνακα 3.4 βλέπουμε ότι πλέον οι συστάδες δεν εξουσιάζονται όλες από την 3η κατηγορία. Συγκεκριμένα οι συστάδες 1η έως 3η εξουσιάζονται από το στατικό, ενδιάμεσο προηγμένο προφίλ. Αναλυτικότερα, παρατηρούμε ότι οι δύο πρώτες συστάδες που προσδιορίζονται από τον αλγόριθμό μας, κυριαρχούνται σαφέστατα από την τρίτη κατηγορία προφίλ, δηλ. τους Προηγμένους χρήστες.

Συστάδα	Στοιχεία	Επί τους εκατό
1η	(3, 1, 7)	(27.27%, 9.09%, 63.64%)
2η	(5, 1, 8)	(35.72%, 7.14%, 57.14%)
3η	(5, 1, 13)	(26.32%, 5.26%, 68.42%)
4η	(5, 9, 11)	(20.00%, 36.00%, 44.00%)
5η	(11, 1, 19)	(35.48%, 3.23%, 61.29%)

Πίνακας 3.3: Αποτελέσματα συσταδοποίησης για 100 χρήστες - 5 συστάδες

Επιπλέον, οι συστάδες 3, 4 και 5 δείχνουν και αυτές μια ξεκάθαρη πλειοψηφία της τρίτης κατηγορίας στα στοιχεία τους. Άρα, οι περισσότερες συστάδες (5 από τις 9 συστάδες σε ποσοστό 55.55%) δείχνουν ένα σαφές πλεονέκτημα των Προηγμένων χρηστών. Η 6η συστάδα μπορεί να θεωρηθεί ως ένας ξεκάθαρος αντιπρόσωπος των Ειδημόνων. Στην 7η συστάδα παρατηρούμε ότι εμφανίζεται μια ισορροπία στοιχείων ανάμεσα στην πρώτη και την τρίτη κατηγορία προφίλ. Αυτό αναδεικνύει μία νέα δυναμική κατηγορία προφίλ η οποία ενεργεί ως μεσάζων μεταξύ των Προηγμένων και των Ειδημόνων. Η 8η και η 9η κυριαρχούνται από την κατηγορία προφίλ Αρχαρίων με ποσοστά 52.63% και με 50.00% αντίστοιχα.

Συστάδα	Στοιχεία	Επί τους εκατό
1η	(1, 1, 4)	(16.66%, 16.66%, 66.66%)
2η	(0, 1, 6)	(0.00%, 14.28%, 85.71%)
3η	(4, 2, 5)	(36.36%, 18.18%, 45.45%)
4η	(3, 3, 6)	(25.00%, 25.00%, 50.00%)
5η	(4, 2, 5)	(36.36%, 18.18%, 45.45%)
6η	(8, 4, 5)	(47.05%, 23.52%, 29.41%)
7η	(4, 1, 4)	(44.44%, 11.11%, 44.44%)
8η	(3, 10, 6)	(15.78%, 52.63%, 31.57%)
9η	(1, 4, 3)	(12.50%, 50.00%, 37.50%)

Πίνακας 3.4: Αποτελέσματα συσταδοποίησης για 100 χρήστες - 9 συστάδες

Ανακεφαλαιώνοντας, η κάθε συστάδα που ανιχνεύεται συσχετίζεται άμεσα με τις απαντήσεις των χρηστών εξάγοντας τις σχετικές πληροφορίες από τα εκάστοτε προφίλ χρήστη. Βασιζόμενοι στις εξαγόμενες συστάδες, οι ειδικοί μπορούν να ξεχωρίσουν τους χρήστες σύμφωνα με τα ανιχνευμένα δυναμικά προφίλ. Η προτεινόμενη μεθοδολογία έχει χρησιμοποιηθεί στο σύστημα ηλεκτρονικής μάθησης το οποίο περιγράφεται αναλυτικά στην ενότητα 7.2.

3.4.3 Κατηγοριοποίηση χρηστών βασισμένη σε καταγραφή συμπεριφοράς

Τα περισσότερα συστήματα μαθησιακής τεχνολογίας συλλέγουν πληροφορίες για τους χρήστες τους, οι οποίες βασίζονται σε ερωτηματολόγια, όπως περιγράψαμε στην προηγούμενη ενότητα. Πολλές φορές αυτό δείχνει να μην επαρκεί μιας και οι χρήστες μπορεί για διάφορους λόγους να μην είναι ειλικρινείς στις απαντήσεις τους. Επιπλέον, μπορεί οι ειδικοί να μην έχουν καταφέρει, μέσω των ερωτήσεων τους, να συλλέξουν όλες τις δυνατές πληροφορίες για το χρήστη, με συνέπεια το σύστημα μαθησιακής τεχνολογίας να μη μπορεί να καλύψει τις πραγματικές ανάγκες των χρηστών. Αντιλαμβανόμενοι αυτήν την αδυναμία ενός συστήματος που στηρίζεται σε ερωτήσεις, προχωρήσαμε στην ανάπτυξη μιας τεχνικής αυτόματης

αξιολόγησης της κατάστασης της συμπεριφοράς των χρηστών. Στηριζόμενοι στην προτεινόμενη επέκταση της αρχιτεκτονικής συστήματος μαθησιακής τεχνολογίας του IEEE (Ενότητα 2.5), προσθέσαμε την εν λόγω τεχνική στη διαδικασία της αξιολόγησης. Σε αυτό το πλαίσιο, η ανάλυση της μη λεκτικής αλληλεπίδρασης μεταξύ ανθρώπου και μηχανής ενισχύει το σύστημα με αξιολογικά δεδομένα, ενώ η περαιτέρω ανάλυσή τους παρέχει χρήσιμες πληροφορίες που μπορούν αξιόπιστα να ισχυροποιήσουν το προφίλ του χρήστη.

3.4.3.1 Ανάλυση μη λεκτικών στοιχείων χρήστη και εξαγωγή κατάστασης

Η μη λεκτική επικοινωνία περιλαμβάνει όλα τα στοιχεία που σχετίζονται με την ανθρώπινη επικοινωνία και δεν μπορούν να εκφραστούν με τη χρήση του γραπτού ή του προφορικού λόγου. Αυτά τα στοιχεία περιλαμβάνουν συνειδητές ή ασυνειδητές εκφράσεις προσώπου και κινήσεις σώματος. Δεδομένου ότι το πρόσωπο είναι το πιο εκφραστικό μέρος του σώματος, έχει δείχθει ότι η περιοχή που σχηματίζεται μεταξύ μετώπου, φρυδιών, ματιών, μύτης και στόματος είναι η περιοχή μέσα στην οποία μεταδίδονται τα περισσότερα μη λεκτικά μηνύματα, δίδοντας μια ιδιαίτερη βαρύτητα στα μάτια ως το πιο εκφραστικό σημείο του προσώπου. Μελέτες έχουν αποδείξει ότι το περιεχόμενο ενός μηνύματος σε μία συνομιλία μεταδίδεται κατά 55% μέσω των εκφράσεων του προσώπου, κατά 35% μέσω της στάσης του σώματος και μόνο κατά 10% μέσω της φωνής. Η στάση του σώματος σε συνδυασμό με τις εκφράσεις του προσώπου και τις χειρονομίες μπορούν να δώσουν μια ολοκληρωμένη εικόνα για την κατάσταση ενός ατόμου. Η κατανόηση της μη λεκτικής επικοινωνίας είναι πολύ σημαντική μιας και μας παρέχει τη δυνατότητα να αντιλαμβανόμαστε αυτό που οι άλλοι νοιώθουν πραγματικά, ακόμη και όταν διστάζουν να το εκφράσουν μέσω της λεκτικής επικοινωνίας.

Οι αρχές και οι τεχνικές που οι άνθρωποι αναπτύσσουν προκειμένου να κατανοήσουν, να προβλέψουν και να χειριστούν τη συμπεριφορά άλλων ανθρώπων αναφέρονται στη βιβλιογραφία συλλογικά ως *Γνωστικές Θεωρίες - Theory of Mind - ToM* [18]. Το να μπορεί κάποιος να κατανοήσει τις προθέσεις των άλλων αποδεικνύει τη σημαντική ικανότητα του να αντιλαμβάνεται τη διανοητική κατάσταση τους.

Το να μπορέσουμε να κατασκευάσουμε μια μηχανή ικανή να κατανοεί τη διανοητική κατάσταση αυτών που τη χειρίζονται είναι σαν να προσπαθούμε να φτιάξουμε μια μηχανή που να προσαρμόζεται στις διανοητικές καταστάσεις του χρήστη της. Σε αυτό το πλαίσιο, έχουν προταθεί στη βιβλιογραφία διάφορα συστήματα που χρησιμοποιήσουν τις Γνωστικές Θεωρίες για δύο λόγους α) για να προβλέψουν τη συμπεριφορά των χρηστών [24] (π.χ. ενέργειες που πρόκειται να κάνουν), και β) για να χειριστούν τη συμπεριφορά (π.χ. προσπαθώντας να φέρουν το χρήστη σε μία συγκεκριμένη κατάσταση, ώστε να εκτελεσθούν επιθυμητές, ή να μην εκτελεσθούν ανεπιθύμητες ενέργειες).

Στην εργασία [88], παρουσιάζεται μια μελέτη σχετική με την αποτελεσματικότητα του συστήματος εντοπισμού και παρακολούθησης του προσώπου ενός χρήστη μπροστά από έναν υπολογιστή, προκειμένου να αξιολογηθούν οι αντιδράσεις του χρήστη, με βάση τα γεγονότα που παράγονται από τον υπολογιστή. Επιπλέον αναλύθηκαν οι διαφορές στις κινήσεις του προσώπου σε διαφορετικά γεγονότα. Κατασκευάστηκαν δύο πειράματα όπου συμμετείχαν 15 άτομα. Κατά τη διάρκεια των πειραμάτων, οι χρήστες χρησιμοποίησαν ένα τεστ γνώσεων, διασχεδαστικό όπου για το πρώτο πείραμα ενώ για το δεύτερο πείραμα προκάλυψε συναισθήματα έκπληξης στο χρήστη. Όπως αποδείχθηκε, τα ισχυρά ερεθίσματα προκάλεσαν σημαντική κίνηση

των χαρακτηριστικών του προσώπου, ενώ τα πιο ασθενή ερεθίσματα προκάλεσαν περιορισμένης έκτασης των κινήσεις των χαρακτηριστικών σημείων του προσώπου. Επιπλέον, οι φυσιολογικές αντιδράσεις και οι αντιδράσεις των χαρακτηριστικών του προσώπου δε συνέπιπταν πάντοτε. Μέσω αυτών των πειραμάτων αποδείχθηκε ότι ένα μη παρεμβατικό (non-intrusive) σύστημα παρακολούθησης χαρακτηριστικών γνωρισμάτων προσώπου μπορεί να είναι μια σημαντική πηγή πληροφόρησης για τεχνικές υπολογιστικής με βάση το συναίσθημα (affective computing).

Στην εργασία [89] έχει προταθεί ένας μη-παρεμβατικός τρόπος με σκοπό να αναλυθούν τα χαρακτηριστικά του προσώπου, να ταξινομηθούν οι εκφράσεις του προσώπου αλλά και να χρησιμοποιηθεί η συναισθηματική ερμηνεία (affect interpretation). Για τα πειράματα έχει χρησιμοποιηθεί μια υπέρυθρη φωτογραφική μηχανή και έχουν συμμετάσχει 21 άτομα. Σε αυτά τα άτομα του ζητήθηκε να εκφράσουν μέσω του προσώπου πέντε διαφορετικές καταστάσεις (ουδέτερη, ευτυχία, αποστροφή, θλίψη και φόβος). Έπειτα, οι θερμικές τιμές έντασης των συγκεκριμένων χαρακτηριστικών των υπέρυθρων εικόνων του προσώπου υποβλήθηκαν σε ανάλυση κύριων συνιστωσών με σκοπό την ταξινόμηση των εκφράσεων του προσώπου.

Στη βιβλιογραφία έχουν προταθεί διάφορες τεχνικές για την εκτίμηση της κίνησης του κεφαλιού, οι οποίες χρησιμοποιούν περισσότερες από μια κάμερες, ή και επιπλέον εξοπλισμό [40], [41], [46], τεχνικές που βασίζονται στην ανίχνευση των χαρακτηριστικών του προσώπου [52], [53] αλλά που υστερούν σε ευρωστία, ή τεχνικές που εκτιμούν τη στάση της κεφαλής [62], [68] εντοπίζοντας πρόσωπα και απαιτώντας οι ανιχνευμένες περιοχές να εναρμονίζονται με το σύνολο δοκιμών (training set). Επιπλέον, έχει μελετηθεί εκτεταμένα στην εργασία [76] και έχει δειχθεί ότι το βλέμμα (Eye gaze), δηλαδή η κατεύθυνση στην οποία έχουν επικεντρωθεί τα μάτια, είναι μια ισχυρή ένδειξη της εστίασης της προσοχής.

Στη βιβλιογραφία δεν έχει δημοσιευθεί στον τομέα της αλληλεπίδρασης ανθρώπου-υπολογιστή έρευνα σχετική με το θέμα της αναγνώρισης του επιπέδου προσοχής ενός χρήστη σε μη παρεμβατικά περιβάλλοντα. Μεγάλο τμήμα σχετικών ερευνών ασχολείται με την εκτίμηση της προσοχής του ατόμου σε περιβάλλοντα συνεδριάσεων (meeting). Για παράδειγμα, στην εργασία [75], παρουσιάζεται μια μέθοδος για την εκτίμηση του κέντρου προσοχής στο πλαίσιο μιας συνεδρίασης. Σε αυτή την προσέγγιση, η προσοχή του ατόμου σε μια συνεδρίαση βασίζεται στην ακουστική (ποιος μιλάει) και σε οπτικές πληροφορίες. Μια κατευθυντική κάμερα χρησιμοποιείται για την ανίχνευση προσώπων και νευρωνικά δίκτυα χρησιμοποιούνται για την εκτίμηση της στάσης κεφαλής κάθε συμμετέχοντος. Πληροφορίες σχετικές με το βλέμμα δε λαμβάνονται υπόψη σε αυτήν την εργασία, αλλά χρησιμοποιείται μία Μπεϋσιανή προσέγγιση με σκοπό να ανιχνευθεί η εστίαση της προσοχής κάθε συμμετέχοντος βασιζόμενη στην εκτίμηση της θέσης της κεφαλής. Σε αυτή την προσέγγιση, χρησιμοποιείται η στάση της κεφαλής καθώς και επισημαίνονται ως πιθανά σημεία προσοχής κάποιοι περιβαλλοντικοί στόχοι (συμπεριλαμβανομένων και των άλλων συμμετεχόντων). Η μέγιστη εκ των υστέρων πιθανότητα ή Μαρκοβιανά κρυφά μοντέλα (Hidden Markov Models -HMM) χρησιμοποιούνται για την αξιολόγηση του σημείου προσοχής του συμμετέχοντα.

Στην εργασία [83], η προσοχή του χρήστη εκτιμάται σε ένα περιβάλλον γραφείου και συγκεκριμένα μπροστά σε έναν υπολογιστή. Χρησιμοποιώντας ένα ζευγάρι από κάμερες καταγράφεται το πρόσωπο του χρήστη και χρησιμοποιούνται 3Δ μοντέλα αναπαράστασης προσώπου και ματιών με σκοπό τον καθορισμό της στάσης της κεφαλής και του βλέμματος. Οι συγγραφείς της εργασίας [87] διερευνούν τη

δομή της συζήτησης σε συνάντηση που συμμετέχουν τέσσερεις άνθρωποι, σε ένα κλειστό χώρο. Σε αυτήν την εργασία, οι κινήσεις του κεφαλιού και η παρουσία / απουσία αρθρώσεων χρησιμοποιούνται για εξαγωγή συμπερασμάτων σχετικά με το ρόλο των συμμετεχόντων στα μοντέλα συζήτησης (conversational models) (ομιλητής και ακροατές). Η κίνηση της κεφαλής υπολογίζεται χρησιμοποιώντας μαγνητικούς αισθητήρες που εφαρμόζονται στο κεφάλι του κάθε χρήστη και οικοδομούνται πιθανολογικά μοντέλα.

Η συσχέτιση της κίνησης της κεφαλής με την προσοχή έχει εξεταστεί και σε συστήματα αναγνώρισης της προσοχής των οδηγών αυτοκινήτων ([26], [93]). Στην [26], μια καθοδική (top-to-down) προσέγγιση έχει προταθεί για την εκτίμηση της προσοχής του οδηγού χρησιμοποιώντας ένα σύστημα με μία κάμερα χωρίς την ανάγκη της πρόσθετης οργάνωσης από την άποψη του υλικού ή βαθμονόμησης- ευθυγράμμιση (calibration) της κάμερας. Σε αυτό το σύστημα, ανιχνεύονται και ακολουθούνται τα μάτια και τα χείλια του οδηγού, χρησιμοποιώντας τις αντίστοιχες αποδόσεις των χρωμάτων (color predicates).

Το γεωμετρικό σχεδιάγραμμα αυτών των χαρακτηριστικών δίνει μια ικανοποιητική, από άποψη σχεδίασης, είσοδο για τον υπολογισμό της περιστροφής της κεφαλής, καθώς επίσης και την κατεύθυνση του βλέμματος των ματιών. Επιπλέον χρησιμοποιείται και το κλείσιμο των ματιών. Οι ανωτέρω μετρικές χρησιμοποιούνται για να αξιολογήσουν την προσοχή ενός οδηγού με τη χρήση τριών πεπερασμένων αυτόματων: ένα που χρησιμοποιεί το κλείσιμο των ματιών και άλλα δύο που χρησιμοποιούν τις κινήσεις άνω-κάτω και δεξιά-αριστερά της κεφαλής.

Στην εργασία [93] εξετάζεται η επαγρύπνηση του οδηγού με την ανίχνευση των ματιών του και τη διάκριση μεταξύ ανοικτών ή κλειστών οφθαλμών. Η αξιολόγηση της μεθόδου πραγματοποιήθηκε με πειράματα υπό κανονικές συνθήκες, έχοντας τοποθετήσει μία κάμερα σε κεντρικό σημείο στο ταμπλό ενός αυτοκινήτου και με τη χρήση ενός γκαουσιανού μοντέλου ικανού να περιγράψει τα μάτια του οδηγού υπό κανονικές συνθήκες. Άρα κλειστά μάτια για το συγκεκριμένο χρονικό διάστημα δείχνουν απροσεξία ή κούραση του οδηγού.

Παρόλο που δεν έχει πραγματοποιηθεί κάποια έρευνα σχετική με το συνδυασμό των πληροφοριών που εξάγουν οι τεχνικές ανίχνευσης βλέμματος και ανίχνευσης κίνησης κεφαλής, σε μη παρεμβατικό περιβάλλον, η προτεινόμενη μέθοδος χρησιμοποιεί ένα συνδυασμό των δύο τεχνικών μαζί με άλλα βιομετρικά στοιχεία εξάγει πληροφορίες σχετικές με τα επίπεδα προσοχής ενός χρήστη μπροστά από την οθόνη ενός υπολογιστή με τη χρήση μιας δικτυακής κάμερας και χωρίς την ανάγκη ειδικού εξοπλισμού.

Το προτεινόμενο σύστημα, το οποίο αποτελεί ένα σύστημα αλληλεπίδρασης ανθρώπου μηχανής (Human Computer Interaction), χρησιμοποιεί τις γνωστικές θεωρίες με σκοπό να ερμηνεύσει τη δυνατότητα του συστήματος να παρακολουθεί τη συμπεριφορά των χρηστών. Οι δυνατότητες αυτές βασίζονται στην καταγραφή της εικόνας του χρήστη καθώς αυτός αλληλεπιδρά με το σύστημα, με σκοπό τον προσδιορισμό διαφόρων χαρακτηριστικών του και την προσαρμογή του συστήματος στις τρέχουσες ανάγκες του χρήστη. Τα επίσημα πρότυπα για τις Γνωστικές Θεωρίες που έχουν συζητηθεί [24] συνήθως προσπαθούν να ανιχνεύσουν τα αίτια (επιθυμίες, πεποιθήσεις, προθέσεις, συναισθήματα, προσοχή) που οδήγησαν το χρήστη σε μια κατάσταση.

Στη δική μας προσέγγιση, δεν προσπαθούμε να προσδιορίσουμε τι κρύβεται πίσω από τα χαρακτηριστικά και τις καταστάσεις που το σύστημα μας εντοπίζει, αλλά

χρησιμοποιούμε τις πληροφορίες κατάστασης για να προσαρμόσουμε το περιβάλλον εκμάθησης του χρήστη. Πιο συγκεκριμένα, χρησιμοποιούμε τα συμπεράσματα από τις Γνωστικές Θεωρίες για να συσχετίσουμε τα διάφορα χαρακτηριστικά που εντοπίζει το σύστημα σχετικά με την τρέχουσα κατάσταση του χρήστη. Για παράδειγμα, η κατεύθυνση του βλέμματος του χρήστη προς την οθόνη, καθώς αυτός διαβάζει κάποιο ηλεκτρονικό κείμενο που προβάλλεται σε αυτήν, μπορεί να συνδεθεί με το επίπεδο του ενδιαφέροντος του χρήστη για αυτό που διαβάζει εκείνη τη στιγμή. Επιπλέον, ξαφνικές και απότομες κινήσεις της κεφαλής σύμφωνα με τις γνωστικές θεωρίες μπορούν να συνδεθούν με τη νευρική δραστηριότητα και την απογοήτευση του χρήστη.

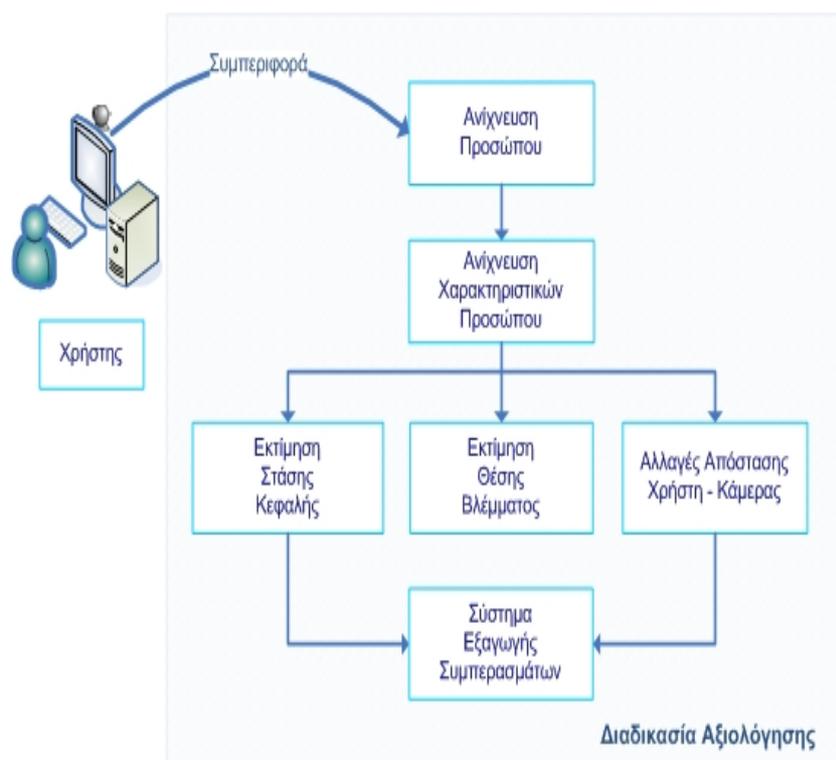
Σε αυτό το πλαίσιο αναπτύχθηκε ένα σύστημα [134], βασισμένο στην προτεινόμενη τεχνική (Ενότητα 2.5) που περιλαμβάνει μία πλατφόρμα δοκιμών (testbed) ανάγνωσης κειμένων 7.3. Η πλατφόρμα αυτή αποτελείται από ένα γραφικό περιβάλλον αλληλεπίδρασης (Graphical User Interface) στο οποίο παρουσιάζεται ένα ηλεκτρονικό έγγραφο σε μορφή pdf (Portable Document Format). Το προτεινόμενο σύστημα αναλαμβάνει την ανάλυση τόσο της κατάστασης στην οποία βρίσκεται ο χρήστης, όσο και της φωνής του, καθώς αυτός διαβάζει το κείμενο, στοχεύοντας στον εντοπισμό προβλημάτων που σχετίζονται με τη δυσλεξία. Η δυσλεξία είναι μια μορφή μαθησιακής δυσκολίας, όπου το άτομο, μικρής ή μεγάλης ηλικίας, παρουσιάζει δυσκολίες στην ανάλυση των λέξεων σε ακουστικές μονάδες συλλαβικής βάσης και στη σύνθεση συλλαβικών ακουστικών μονάδων σε λεξικά σύνολα με εννοιολογικό περιεχόμενο.

Η δημιουργία αυτού του συστήματος βασίζεται στην προτεινόμενη επέκταση της αρχιτεκτονικής συστήματος μαθησιακής τεχνολογίας του IEEE, όπως αυτή περιγράφηκε στην ενότητα 2.5. Η αυτόματη παρακολούθηση τόσο της απόδοσης του χρήστη όσο και της τρέχουσας συναισθηματικής και ψυχολογικής κατάστασης, αποτελεί τα βασικά χαρακτηριστικά του συστήματος. Επιπλέον χαρακτηριστικό του συστήματος θεωρείται η δημιουργία του προσωπικού προφίλ του χρήστη με στόχο την παροχή προσωποποιημένης βοήθειας. Στόχος όλων των προαναφερθέντων χαρακτηριστικών του προηγμένου συστήματος είναι η βελτίωση την αναγνωστικής ικανότητας του χρήστη, κατά τέτοιο βαθμό ώστε αυτός να μπορεί να ενταχθεί στο κανονικό πρόγραμμα της τάξης, χωρίς να έχει την ανάγκη ενισχυτικής διδασκαλίας. Μια αναλυτική παρουσίαση της λειτουργίας του εν λόγω συστήματος δίνεται στον αναγνώστη στην ενότητα 7.3.

Η τεχνική που αναπτύχθηκε στο πλαίσιο αυτής της διατριβής εξάγει πληροφορίες σχετικά με τη κατάσταση του χρήστη βασιζόμενη στα χαρακτηριστικά του προσώπου του και στις κινήσεις των ματιών και της κεφαλής. Για την εξαγωγή αυτών των χαρακτηριστικών χρησιμοποιείται μια απλή δικτυακή κάμερα η οποία καταγράφει το χρήστη καθώς αυτός αλληλεπιδρά με το σύστημα.

Αναλυτικότερα, χρησιμοποιούμε τη θέση και την κίνηση των ευδιάκριτων σημείων γύρω από τα μάτια και τη θέση της ίριδας των ματιών του χρήστη για την κατασκευή διανυσμάτων τα οποία παρέχουν πληροφορίες σχετικά με την κατεύθυνση του βλέμματος. Επιπλέον καταγράφεται και αναλύεται η στάση της κεφαλής του χρήστη, η θέση των χεριών του και η απομάκρυνση ή το πλησίαμά του στην κάμερα. Το προτεινόμενο σύστημα αναλύει περαιτέρω αυτές τις πληροφορίες βασιζόμενο σε γνωστικές θεωρίες [18] και παράλληλα σε αξιολόγηση των ειδικών του συστήματος. Το σύστημα τελικά εξάγει πληροφορίες σχετικές με την προσοχή, το ενδιαφέρον ή όχι του χρήστη (κοιτάζει στην οθόνη ή όχι, το βλέμμα του επικεντρώνεται σε ένα συγκεκριμένο σημείο για μεγάλο χρονικό διάστημα, κ.τ.λ.).

Με την προτεινόμενη μέθοδο εξετάστηκαν, για πρώτη φορά, το πρόσωπο και τα



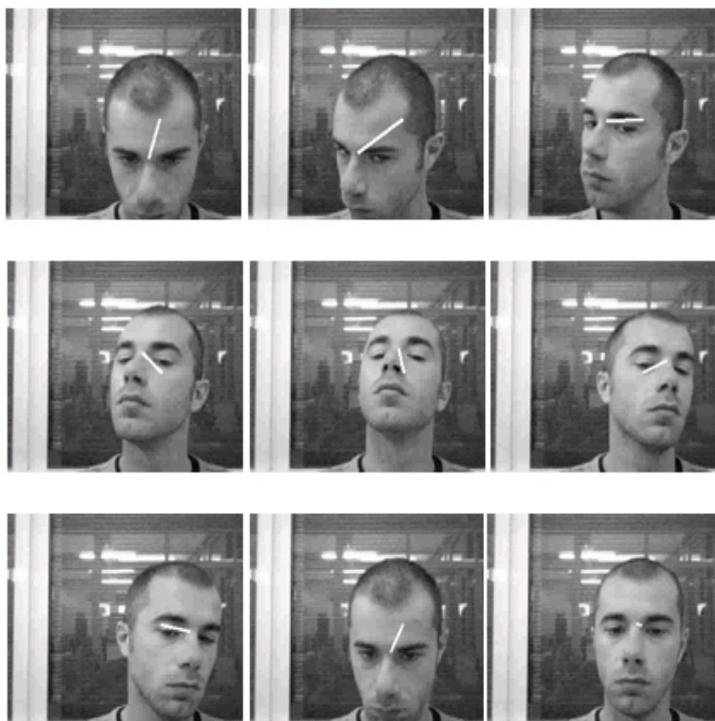
Σχήμα 3.4: Εμπλουτισμός της διαδικασίας Αξιολόγησης

χαρακτηριστικά του προσώπου του χρήστη όταν αυτός βρίσκεται μπροστά σε μια οθόνη υπολογιστή χρησιμοποιώντας μόνο μια απλή δικτυακή κάμερα, χωρίς να είναι αναγκαία η ρύθμιση των παραμέτρων αυτής (calibration- φωτεινότητας, αντίθεσης κτλ.). Με βάση τις κινήσεις τόσο της κεφαλής όσο και της ίριδας των ματιών, ανιχνεύονται η στάση της κεφαλής και εκτιμάται το βλέμμα του χρήστη (eye gaze). Επιπλέον ανιχνεύονται οι κινήσεις του χρήστη καθώς απομακρύνεται ή πλησιάζει κοντά στην κάμερα. Περισσότερα χαρακτηριστικά όπως το άνοιγμα του στόματος, συνοφρύωμα, κινήσεις χεριών μπορούν επίσης να ανιχνευθούν και να διαδραματίσουν το ρόλο τους στον προσδιορισμό της προσοχής που δίνει ο χρήστης σε αυτό που διαβάζει. Τα πειράματα πραγματοποιήθηκαν χρησιμοποιώντας τα ανιχνευμένα βιομετρικά χαρακτηριστικά για την εκτίμηση της κατάστασης του χρήστη. Τα βιομετρικά στοιχεία του χρήστη που ανιχνεύθηκαν στη συνέχεια χρησιμοποιήθηκαν για τον καθορισμό του προφίλ του χρήστη και την προσαρμογή του εκπαιδευτικού υλικού στις τρέχουσες ανάγκες των χρηστών.

3.4.3.2 Ανίχνευση και παρακολούθηση των χαρακτηριστικών του προσώπου-εκτίμηση βλέμματος και στάσης

Η εξαγωγή των χαρακτηριστικών του προσώπου είναι ένα βασικό βήμα σε πλήθος εφαρμογών όπως αναγνώριση προσώπων, συστήματα αλληλεπίδρασης ανθρώπου μηχανής, αναγνώριση εκφράσεων προσώπου, επιτήρηση και ανίχνευση του βλέμματος και της στάσης του σώματος. Η πλειοψηφία αυτών των μεθόδων χρησιμοποιούν την ανίχνευση προσώπου ως βήμα προεπεξεργασίας. Έτσι και το πρώτο στάδιο της προτεινόμενης μεθόδου είναι η ανίχνευση του προσώπου του χρήστη όπως παρουσιάζεται στο Σχήμα 3.4.

Για τον εντοπισμό των προσώπων χρησιμοποιήθηκε η μέθοδος που προτάθηκε από



Σχήμα 3.5: Αποτελέσματα εκτίμησης στάσης κεφαλής

τους Viola και Jones [138]. Η αναλυτική περιγραφή του αλγορίθμου δίνεται στην Ενότητα 4.4.2.1.2.

Στο δεύτερο στάδιο της προτεινόμενης μεθόδου, υιοθετείται μία τεχνική ταιριάσματος προτύπων (template matching) [12] με σκοπό την ανίχνευση των χαρακτηριστικών του προσώπου. Η περιοχή προσώπου που έχει εντοπιστεί στο πρώτο στάδιο χρησιμοποιείται ως είσοδος και με χρήση προτύπων εντοπίζεται τόσο η περιοχή των ματιών όσο και η ίριδας. Αναλυτική περιγραφή της μεθόδου δίνεται στην Ενότητα 4.4.2.2

Το τρίτο στάδιο περιλαμβάνει την εκτίμηση τόσο της στάσης της κεφαλής, όσο και της θέσης του βλέμματος. Επιπλέον στο τρίτο στάδιο εκτιμώνται τυχόν αλλαγές στην απόσταση ανάμεσα στο χρήστη και στην κάμερα.

Όσο αφορά στην εκτίμηση της στάσης της κεφαλής, χρησιμοποιούνται ως είσοδος στον αλγόριθμο η θέση της κεφαλής, όπως αυτή έχει ανιχνευθεί στο πρώτο στάδιο και η θέση των ματιών σύμφωνα με την εκτίμηση που πραγματοποιήθηκε στο δεύτερο στάδιο. Η μέθοδος που ακολουθείται για την εκτίμηση της στάσης της κεφαλής περιγράφεται αναλυτικά στο [13] και τα αποτελέσματά της απεικονίζονται στο Σχήμα 3.5.

Ανά τακτά χρονικά διαστήματα, χρησιμοποιώντας το χρωματικό μοντέλο του δέρματος του κάθε χρήστη, όπως αυτό έχει εκτιμηθεί κατά το πρώτο βήμα του αλγορίθμου, γίνεται εντοπισμός του προσώπου με σκοπό την αντιμετώπιση της περίπτωσης που κάποιο από τα χαρακτηριστικά "χαθεί" λόγω μιας περιστροφής. Έπειτα ένας αλγόριθμος κατάτμησης της περιοχής του προσώπου με βάση αυτήν την πληροφορία μπορεί συνεχώς να παρακολουθεί την περιοχή του προσώπου και να καλεί τον ανιχνευτή προσώπου, όταν αυτή η περιοχή γίνεται μέγιστη, οπότε και θεωρείται ότι ο χρήστης κοιτάζει προς την κάμερα. Ακόμη, η υιοθέτηση της παρακολούθησης της περιοχής του κεφαλιού με βάση το προσωποποιημένο χρωματικό του μοντέλο

βοηθάει στην εύρεση της περιστροφής του προσώπου κατά τον άξονα τον κάθετο στο επίπεδο της εικόνας.

Για την ανίχνευση της θέσης του βλέμματος, ο αλγόριθμος εντοπίζει τέσσερα συγκεκριμένα σημεία (πάνω, κάτω, δεξιά, αριστερά) γύρω από την ανιχνευμένη περιοχή των ματιών. Στη συνέχεια υπολογίζονται οι κανονικοποιημένες διαφορές ανάμεσα στις συσχετισμένες τιμές των σημείων πάνω-κάτω και των σημείων δεξιά-αριστερά. Οι εξισώσεις που χρησιμοποιούνται για την ανίχνευση της θέσης του βλέμματος και η αναλυτική παρουσίαση αυτών παρατίθενται στο [13].

Για τον υπολογισμό της απόστασης μεταξύ χρήστη και κάμερας, ο αλγόριθμος παρακολουθεί τα κέντρα των ματιών σε μια ακολουθία καρέ [28] και υπολογίζει την απόσταση μεταξύ των ματιών. Όταν η απόσταση αυτή αυξάνει, ο χρήστης πλησιάζει στην οθόνη, ενώ στην αντίθετη περίπτωση απομακρύνεται. Συμπερασματικά, στην έξοδο του τρίτου σταδίου, το σύστημα παρέχει χαρακτηριστικά διάνυσματα για κάθε καρέ. Σχετικά με την εκτίμηση της στάσης της κεφαλής εξάγεται το πλάτος της μετατόπισης της κεφαλής από την αρχική θέση και η κατεύθυνση της μετατόπισης. Όσο αφορά στην εκτίμηση της θέσης του βλέμματος, εξάγεται το ζεύγος τιμών που περιγράφει την ακριβή μέση θέση του βλέμματος. Για τις αλλαγές στην απόσταση ανάμεσα στο χρήστη και στην κάμερα εξάγεται μια τιμή που περιγράφει αυτήν την απόσταση, κανονικοποιημένη σε σχέση με τη θέση του χρήστη στο αρχικό καρέ.

Χρησιμοποιώντας τις προαναφερθείσες τεχνικές, τα χαρακτηριστικά που εξήχθησαν κατά τη διάρκεια του τρίτου σταδίου χρησιμοποιούνται ως είσοδος από το σύστημα εξαγωγής συμπερασμάτων. Κατά αυτόν τον τρόπο το αποτέλεσμα της ανάλυσης της τρέχουσας κατάστασης του χρήστη μοντελοποιείται σε μία από τις έξι καλά ορισμένες καταστάσεις χρήστη (Απογοητευμένος/ δυσκολία στην ανάγνωση, Αναστατωμένος, Κουρασμένος/Νυσταγμένος, Με αποσπασμένη προσοχή, Προσηλωμένος/Προσεκτικός, Με μεγάλο ενδιαφέρον), ορισμός που βασίστηκε στις Γνωστικές Θεωρίες.

Η εκτίμηση της τρέχουσας κατάστασης του χρήστη μπορεί να εξαχθεί με τη χρήση ενός ασαφούς συστήματος εξαγωγής συμπερασμάτων και συγκεκριμένα με τη χρήση ενός Sugeno-type συστήματος [131]. Στα πλεονεκτήματα ενός τέτοιου συστήματος συγκαταλέγονται η ακριβής προσέγγιση και η ικανότητα γενίκευσης. Στην περίπτωσή μας χρησιμοποιούμε τα διάνυσματα που εξήχθησαν κατά τη διάρκεια του τρίτου σταδίου, δηλαδή το χαρακτηριστικό διάνυσμα με την εκτίμηση της στάσης της κεφαλής, το διάνυσμα με την εκτίμηση της θέσης του βλέμματος και το διάνυσμα με τις αλλαγές στην απόσταση ανάμεσα στο χρήστη και στην κάμερα.

Πριν από τη φάση της εκπαίδευσης τα δεδομένα μας κατηγοριοποιούνται χρησιμοποιώντας έναν αλγόριθμο υπο-κατηγοριοποίησης όπως αυτός παρουσιάστηκε στην εργασία [154]. Αυτός ο αλγόριθμος πραγματοποιεί αρχικά μία σαφή κατηγοριοποίηση των δεδομένων, κι έπειτα οδηγεί τις ομάδες των κατηγοριοποιημένων δεδομένων σε ένα ασαφές σύστημα με σκοπό την αποφυγή της κατάρτας της διαστασιοποίησης (dimensionality). Το πλήθος των ομάδων που δημιουργήθηκε από τον αλγόριθμο καθορίζει το βέλτιστο αριθμό των ασαφών κανόνων. Οι παράμετροι του ασαφούς συστήματος εξαγωγής συμπερασμάτων (κέντρα και πλάτη της συνάρτησης συμμετοχής) αποκτώνται εφαρμόζοντας τη μέθοδο των ελαχίστων τετραγώνων και τη μέθοδο ανατροφοδότησης μετάδοσης (οπισθοδρομικής διάδοσης) κατάβασης κλίσης [149].

Στον Πίνακα 3.5 παρουσιάζονται τα χαρακτηριστικά του προσώπου του χρήστη που χρησιμοποιούνται με σκοπό τη συλλογή πληροφοριών. Κάθε χαρακτηριστικό

συνδέεται με κάποια συμπεριφορά του χρήστη και με συγκεκριμένη οπτική ένδειξη που ελέγχεται από το σύστημα. Η κάθε οπτική ένδειξη μπορεί να λάβει συγκεκριμένες τιμές, όπως αυτές αναγράφονται στον Πίνακα 3.5.

Για παράδειγμα, ο οπτικός χαρακτηρισμός *Μάτια ορθάνοικτα* συνδέεται με τα χαρακτηριστικά του προσώπου που συλλέγονται από την περιοχή των ματιών. Για τη συλλογή αυτής της πληροφορίας υπολογίζεται η απόσταση των χαρακτηριστικών σημείων γύρω από τα μάτια και οι τιμές που τελικά δίνει το σύστημα είναι οι εξής: Έντονα, Λιγότερο έντονα, Κανονικά, Χαλαρά και Πολύ χαλαρά.

Στον Πίνακα 3.6 παρουσιάζεται η κατηγοριοποίηση των πιθανών καταστάσεων του χρήστη. Ειδικότερα, η κάθε μία από τις έξι καταστάσεις που μπορεί να βρεθεί ο χρήστης χαρακτηρίζεται από ένα σύνολο από οπτικά χαρακτηριστικά καθώς και από σύνολο ενδείξεων συμπεριφοράς του χρήστη.

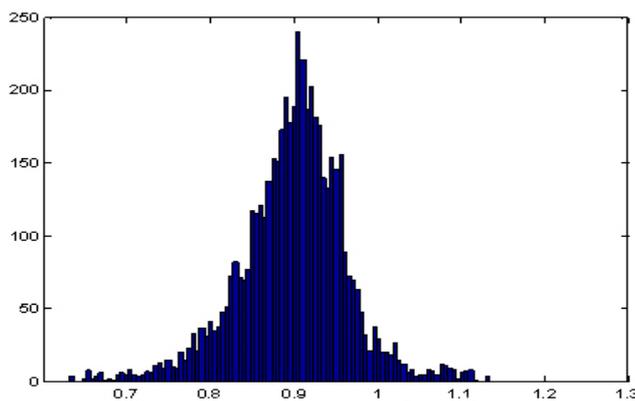
Για παράδειγμα, σε περίπτωση που κάποιος ή κάτι αποσπάσει την προσοχή του χρήστη, καθώς αυτός αλληλεπιδρά με το σύστημα, θα μεταβληθεί η θέση της κεφαλής του, διότι θα υπάρξει περιστροφή προς την κατεύθυνση που ακούστηκε ο θόρυβος ή βρίσκεται ο συνομιλητής. Επιπλέον τα μάτια του χρήστη δεν θα κοιτούν στην οθόνη και γενικά τα χαρακτηριστικά του προσώπου μένουν ακίνητα/παγωμένα, με συνέπεια ο χρήστης να δείχνει αναστατωμένος-σαστισιμένος.

3.4.3.3 Πειραματικά αποτελέσματα

Όπως αναφέρθηκε, η καταγραφή και η ανάλυση του χρήστη ενός ηλεκτρονικού συστήματος εκμάθησης βοηθά στη δημιουργία ενός συστήματος προσαρμοσμένου στις ανάγκες του χρήστη. Αυτή η ανάλυση περιγράφει την τρέχουσα κατάσταση του χρήστη εμπλουτίζοντας το προφίλ του και παρέχοντάς του το κατάλληλο υλικό εκμάθησης με τον κατάλληλο ρυθμό μάθησης. Στο πλαίσιο της εκπαίδευσης, του ελέγχου και της αξιολόγησης της εξαγωγής χαρακτηριστικών και της ανίχνευσης της κατάστασης του χρήστη, διεξήχθη πείραμα. Σε αυτό το πείραμα συμμετείχαν είκοσι χρήστες, μαθητές ηλικίας 8- 10 χρόνων με διαγνωσμένο πρόβλημα μαθησιακής δυσκολίας και συγκεκριμένα με δυσλεξία.

Αρχικά ζητήθηκε από τους χρήστες να διαβάσουν κάποιο κείμενο που εμφανιζόταν στην οθόνη ενός υπολογιστή και παράλληλα πραγματοποιήθηκε η βιντεοσκόπησή τους. Τα βίντεο που παράχθηκαν (20 βίντεο των 800-1200 καρέ) αρχικά χαρακτηρίστηκαν από ομάδα ειδικών ψυχολόγων και ειδικών σε θέματα δυσλεξίας, προσδίδοντας τους μία από τις έξι καταστάσεις των χρηστών σε χαρακτηριστικά καρέ. Κατά αυτόν τον τρόπο δημιουργήθηκαν έξι χαρακτηριστικά σύνολα, ένα για κάθε περίπτωση, που αποτέλεσαν και την υπόθεση αλήθειας (ground truth) για το σύστημα. Οι αποστάσεις και οι μετατοπίσεις των χαρακτηριστικών για κάθε κατάσταση χρησιμοποιήθηκαν για την εκπαίδευση ενός νευροασαφούς δικτύου, το οποίο χρησιμοποιήθηκε για σκοπό ελέγχου.

Ωστόσο, η αντιστοίχιση μεταξύ των ανιχνευμένων χαρακτηριστικών και των καταστάσεων του χρήστη δεν τεκμηριώθηκε από τους ειδικούς. Για παράδειγμα, οι ειδικοί χαρακτήρισαν το χρήστη σε ένα τμήμα ενός βίντεο ως *Προσεκτικός* αλλά δεν έδωσαν λεπτομέρειες σχετικά με τα οπτικά χαρακτηριστικά που παρατήρησαν και τα οποία τους οδήγησαν σε αυτό τον χαρακτηρισμό. Βασιζόμενοι σε προηγούμενες μελέτες αναπαράστασης συναισθήματος και ανάλυσης εκφράσεων προσώπου [33] πραγματοποιήθηκε η αντιστοίχιση μεταξύ των καταστάσεων και των ανιχνευμένων βιομετρικών χαρακτηριστικών.



Σχήμα 3.6: Κατανομή των τιμών που λαμβάνει το κλάσμα της τρέχουσας τιμής της απόστασης προς την τιμή που είχε αυτή η απόσταση στο καρέ αρχικοποίησης για σύνολο 250 διαδοχικών καρέ

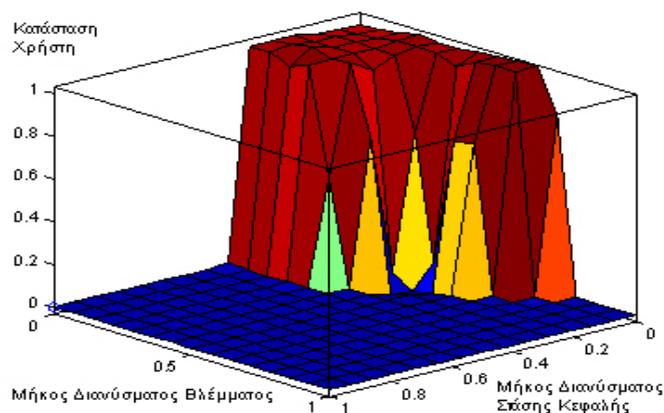
Έπειτα για τα ίδια βίντεο υλοποιήθηκαν τα στάδια 1-3, δηλαδή η ανίχνευση του προσώπου και των χαρακτηριστικών αυτού και η εξαγωγή χαρακτηριστικών σχετικών με το πρόσωπο, το βλέμμα, την κίνηση, τη θέση και τη στάση της κεφαλής του χρήστη. Αυτά τα χαρακτηριστικά που καταγράφονται καθώς ο χρήστης αλληλεπιδρά με τον υπολογιστή, παρέχουν ένα σύνολο στοιχείων σχετικών με τη συμπεριφορά του χρήστη. Ο συνδυασμός αυτών των στοιχείων, όπως φαίνεται στη συνέχεια, αποδίδει την τρέχουσα κατάσταση του χρήστη.

Οι τιμές των διανυσμάτων της στάσης της κεφαλής και του βλέμματος κυμαίνονται μεταξύ 0 και 1, ενώ η τιμή του διανύσματος με τις αλλαγές στην απόσταση ανάμεσα στο χρήστη και στην κάμερα δίνεται ως ένα κλάσμα της τρέχουσας τιμής της απόστασης προς την τιμή που είχε αυτή η απόσταση στο καρέ αρχικοποίησης. Στο Σχήμα 3.6 απεικονίζεται η κατανομή των τιμών που λαμβάνει αυτό το κλάσμα (οριζόντιος άξονας) στο σύνολο των 250 διαδοχικών καρέ (κατακόρυφος άξονας). Αξιοσημείωτο είναι ότι η μέση τιμή είναι μετατοπισμένη προς τα αριστερά. Αυτό συμβαίνει διότι, καθώς το πρόσωπο περιστρέφεται προς τα δεξιά ή αριστερά, η προσπίπτουσα τιμή της τρέχουσας απόστασης μειώνεται.

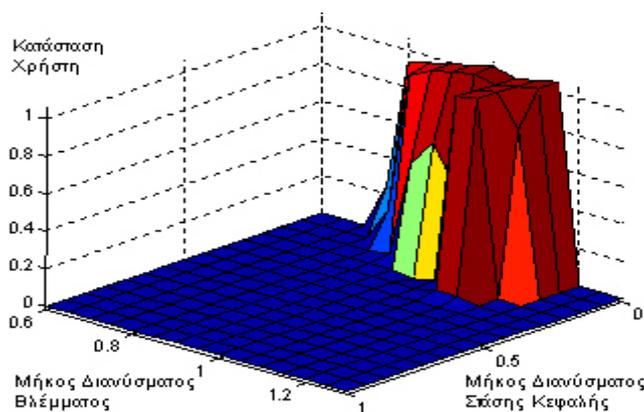
Χρησιμοποιώντας τα βιομετρικά χαρακτηριστικά που λάβαμε κατά τη φάση της εκπαίδευσης του συστήματος, μελετήσαμε τις τιμές τους σε σχέση με τις πιθανές καταστάσεις του χρήστη. Αρχικά έγινε μια αντιστοίχιση των καταστάσεων σε ένα κανονικοποιημένο διάστημα τιμών. Έτσι η κατάσταση *Απογοητευμένος/ δυσκολία στην ανάγνωση* λαμβάνει τιμές κοντά στο 0 ενώ η κατάσταση *Με μεγάλο ενδιαφέρον* λαμβάνει τιμές κοντά στο 1.

Στο Σχήμα 3.7 μπορούμε να παρατηρήσουμε ότι για μικρές τιμές τόσο του διανύσματος της στάσης της κεφαλής όσο και του διανύσματος του βλέμματος, η κατάσταση του χρήστη λαμβάνει τιμές κοντά στο 1. Αντίθετα όταν το πλάτος των διανυσμάτων της στάσης της κεφαλής και του βλέμματος αυξάνεται, η τιμή για την κατάσταση του χρήστη μειώνεται πλησιάζοντας το μηδέν.

Όμοια, στο Σχήμα 3.8 παρατηρούμε ότι για μικρές τιμές του διανύσματος της στάσης της κεφαλής καθώς και για μικρές τιμές του κλάσματος που δημιουργείται από την τρέχουσα απόσταση μεταξύ των ματιών και την αρχική απόστασή τους, ο χρήστης θεωρείται ότι βρίσκεται στην κατάσταση *Με μεγάλο ενδιαφέρον*. Αντίθετα, για τιμές των διανυσμάτων κοντά στο 1, η κατάσταση του χρήστη είναι *Απογοητευμένος*.



Σχήμα 3.7: Τρισδιάστατη απεικόνιση του μήκους του διανύσματος βλέμματος και στάσης της κεφαλής σε σχέση με τις καταστάσεις του χρήστη



Σχήμα 3.8: Τρισδιάστατη απεικόνιση του μήκους του κλάσματος της απόστασης μεταξύ ματιών και στάσης της κεφαλής σε σχέση με τις καταστάσεις του χρήστη

Στον Πίνακα 3.7 παρουσιάζονται τα αποτελέσματα του νευροασαφούς ταξινομητή που χρησιμοποιήσαμε. Στη στήλη *Μέσο σφάλμα* απεικονίζεται η μέση τιμή του απόλυτου σφάλματος ανάμεσα στο χαρακτηρισμό της κατάστασης του χρήστη για κάποιο καρέ που δίνεται στην έξοδο του νευροασαφούς συστήματος και στον καταγεγραμμένο από τους ειδικούς χαρακτηρισμό για το ίδιο καρέ. Στη στήλη *Ποσοστό επιτυχία* καταγράφεται το ποσοστό που ένα καρέ χαρακτηρίστηκε ορθά από το νευροασαφές σύστημα.

Συμπερασματικά μπορούμε να αναφέρουμε ότι η χρήση των βιομετρικών χαρακτηριστικών ενός χρήστη καθώς αυτός αλληλεπιδρά με ένα σύστημα, μπορεί να οδηγήσει σε εκτίμηση του ενδιαφέροντος και της κατάστασης του χρήστη. Αυτές οι πληροφορίες, όπως αποδεικνύεται, μπορεί να αποβούν σημαντικές για την δημιουργία συστημάτων προσαρμοσμένων στις ανάγκες του χρήστη. Στον Πίνακα 3.8 φαίνεται ένας τρόπος σύνδεσης της κατάστασης του χρήστη με την αντίδραση του συστήματος. Στην ενότητα 7.3 παρουσιάζεται αναλυτικά η εφαρμογή όπου με τη βοήθεια των βιομετρικών χαρακτηριστικών του χρήστη που αλληλεπιδρά με τον υπολογιστή, εξάγεται η τρέχουσα κατάσταση του και ανάλογα με αυτήν προσαρμόζεται και το περιβάλλον της εφαρμογής, λαμβάνοντας υπόψη την αντιστοίχιση του Πίνακα 3.8. Συγκεκριμένα, το περιβάλλον της εφαρμογής είναι ένα περιβάλλον μέσα στο οποίο εμφανίζεται ένα κείμενο, με κατάλληλη μορφή (γραμματοσειρά, διάκενο, κ.τ.λ.) και ρυθμό φωτεινής επισήμανσης των λέξεων σύμφωνα με την κατάσταση του χρήστη. Για παράδειγμα, αν η κατάσταση του χρήστη είναι *Με αποσπασμένη προσοχή* τότε η φωτεινή επισήμανση θα σταματήσει στην τρέχουσα λέξη του κειμένου και θα περιμένει έως ότου ο χρήστης εστιάσει το βλέμμα του στο συγκεκριμένο σημείο του κειμένου.

Πίνακας 3.5: Οπτικές Ενδείξεις και Χαρακτηριστικά

Χαρακτηριστικά Προσώπου	Γνώρισμα-Συμπεριφορά	Οπτικά Χαρακτηριστικά	Πιθανές Τιμές
Φρύδια	Κίνηση των φρυδιών προς τα μάτια	Αυστηρότητα	- Ήπια - Μέτρια - Υψηλή
Βλέμμα	Μήκος αντιπροσωπευτικού διανύσματος κίνησης βλέμματος	Θέση ματιών ως προς οθόνη	- Εκτός - Εντός
Περιοχή ματιών	Απόσταση χαρακτηριστικών σημείων γύρω από τα μάτια	Μάτια ορθάνοιχτα	- Έντονα - Λιγότερο έντονα - Κανονικά - Χαλαρά - Πολύ χαλαρά
Οφθαλμικά κέντρα	Μήκος διανύσματος που αντιπροσωπεύει τη στάση της κεφαλής	Κίνηση κεφαλής	- Ναι - Όχι
Οφθαλμικά κέντρα	Απόσταση μεταξύ των ματιών	Διάνυσμα στάσης και κίνησης κεφαλής (ως προς την κατεύθυνση)	- Καμία - Προς τα εμπρός - Προς τα πίσω - Πάνω - Κάτω - Αριστερά - Δεξιά
Οφθαλμικά κέντρα	Μήκος 1ης παραγώγου διανύσματος στάσης κεφαλής	Κίνηση κεφαλής (ως προς την ταχύτητα)	- Καμία - Γρήγορη - Κανονική - Αργή
Βλέφαρα	Κίνηση των βλεφάρων ως προς τα μάτια	Συνοφρυωμένος	- Ναι - Όχι
Βλέμμα	Κατεύθυνση και διακυμάνσεις διανύσματος βλέμματος	Ανοιγόκλεισιμο ματιών	- Ναι - Όχι
Περιοχή προσώπου	Ακίνητα χαρακτηριστικά γνωρίσματα προσώπου	Ακίνητο πρόσωπο	- Ναι - Όχι

Πίνακας 3.6: Κατηγοριοποίηση Πιθανών Καταστάσεων Χρήστη

Οπτικά Χαρακτηριστικά	Ενδείξεις	Κατάσταση
<ul style="list-style-type: none"> - Τα μάτια δεν κοιτούν στην οθόνη - Κίνηση Κεφαλιού (κατεύθυνση / ταχύτητα) - Ανοιγοκλείσιμο Ματιών - Ακίνητο πρόσωπο 	<ul style="list-style-type: none"> - Συνοφρώωση 	Απογοητευμένος / Δυσκολία στην ανάγνωση
<ul style="list-style-type: none"> - Τα μάτια δεν κοιτούν στην οθόνη - Ορθάνοικτα Μάτια - Κίνηση Κεφαλιού (κατεύθυνση / ταχύτητα) 	<ul style="list-style-type: none"> - Κάποιος μιλάει στο χρήστη - Ακούγεται κάποιος θόρυβος 	Αναστατωμένος
<ul style="list-style-type: none"> - Τα μάτια κοιτούν στην οθόνη - Ορθάνοικτα Μάτια 	<ul style="list-style-type: none"> - Χασμουριέται - Προσπαθεί να σταματήσει το διάβασμα μιλώντας σε κάποιον 	Κουρασμένος/Νυσταγμένος
<ul style="list-style-type: none"> - Τα μάτια κοιτούν στην οθόνη - Ορθάνοικτα Μάτια - Κίνηση Κεφαλιού (κατεύθυνση / ταχύτητα) 	<ul style="list-style-type: none"> - Δεν κοιτούν στην οθόνη - Μιλάει σε κάποιον - Σταματά το διάβασμα 	Με αποσπασμένη προσοχή
<ul style="list-style-type: none"> - Τα μάτια κοιτούν στην οθόνη - Ορθάνοικτα Μάτια - Κίνηση Κεφαλιού (κατεύθυνση / ταχύτητα) 	<ul style="list-style-type: none"> - Κοιτάζει επίμονα στην οθόνη - Κάνει μια αξιόλογη προσπάθεια να αλληλεπιδράσει - Αντιμετωπίζει μεγάλες δυσκολίες 	Προσηλωμένος / Προσεκτικός
<ul style="list-style-type: none"> - Αυστηρότητα - Τα μάτια κοιτούν στην οθόνη - Ορθάνοικτα Μάτια - Κίνηση Κεφαλιού (κατεύθυνση / ταχύτητα) 	<ul style="list-style-type: none"> - Κοιτάζει επίμονα στην οθόνη - Βρίσκει κάτι αξιοσημείωτο 	Με μεγάλο ενδιαφέρον

Πίνακας 3.7: Επίδοση κατηγοριοποίησης πιθανών καταστάσεων χρήστη με χρήση νευροασαφούς δικτύου

Κατάσταση χρήστη	Μέσο σφάλμα	% Ποσοστό επιτυχίας
Με μεγάλο ενδιαφέρον	0.04	98
Απογοητευμένος/ δυσκολία στην ανάγνωση	0.24	72

Πίνακας 3.8: Σύνδεση κατάστασης χρήστη με την αντίδραση του συστήματος

Κατάσταση χρήστη	Αντίδραση Συστήματος
Με μεγάλη ενδιαφέρον	καμία μεταβολή
Προσηλωμένος/Προσεκτικός	καμία μεταβολή
Με αποσπασμένη προσοχή	"παγώνει" η φωτεινή υπογράμμιση, ο κέρσορας τοποθετείται στην αρχή της τρέχουσας παραγράφου, δυνατότητα ενεργοποίησης της φωτεινής υπογράμμισης από το χρήστη
Κουρασμένος/Νυσταγμένος	Αύξηση φωτεινότητας υπογράμμισης, μείωση ταχύτητας κίνησης της φωτεινής υπογράμμισης, παράθυρο διαλόγου με χρήστη για να εκφράσει την επιθυμία του αν θέλει ή όχι να συνεχίσει
Αναστατωμένος	"παγώνει" η φωτεινή υπογράμμιση, παράθυρο διαλόγου με χρήστη για να εκφράσει την επιθυμία του αν θέλει ή όχι να συνεχίσει
Απογοητευμένος/ ανάγνωση	δυσκολία στην Αύξηση φωτεινότητας υπογράμμισης, μείωση ταχύτητας κίνησης της φωτεινής υπογράμμισης, αύξηση μεγέθους γραμματοσειράς, αύξηση διάκενο μεταξύ γραμμών

Κεφάλαιο 4

Ανάλυση και τεκμηρίωση πολυμεσικού περιεχομένου

4.1 Εισαγωγή

Ένα σύστημα ηλεκτρονικής μάθησης για να είναι προσαρμοσμένο στις ανάγκες του χρήστη, θα πρέπει να είναι εφοδιασμένο με τεχνικές κατηγοριοποίησης όχι μόνο των χρηστών του αλλά και του εκπαιδευτικού υλικού που αυτό διαθέτει. Επιπλέον θα πρέπει να διαθέτει τεχνικές ταιριάσματος των προφίλ των χρηστών με το κατάλληλο υλικό. Η ανάγκη για αυτόματη κατηγοριοποίηση πολυμεσικού υλικού οδήγησε την έρευνα μας στην περιοχή της ψηφιακής επεξεργασίας εικόνας, έχοντας ως στόχο την εξόρυξη γνώσης ικανής για την αυτόματη κατηγοριοποίηση του υλικού αυτού.

Στη βιβλιογραφία, αρχικά, παρουσιάστηκαν τεχνικές για αυτόματη κατηγοριοποίηση εγγράφων κειμένου [95] και τα τελευταία χρόνια παρατηρείται μια άνθηση τεχνικών αυτόματης κατηγοριοποίησης αρχείων εικόνων, βίντεο και ήχου. Ειδικότερα, για την αυτόματη κατηγοριοποίηση εικόνων χρησιμοποιούνται τεχνικές αυτόματης κατάταξης που βασίζονται σε ανίχνευση χαρακτηριστικών χαμηλού επιπέδου της εικόνας [80].

Εκτός του ότι η κατηγοριοποίηση του πολυμεσικού υλικού μπορεί να συνδυαστεί με τα προφίλ των χρηστών με σκοπό την πρόσβαση του χρήστη σε υλικό κατάλληλο για το προφίλ του, θα μπορούσε να χρησιμοποιηθεί και για σκοπούς μάθησης. Για παράδειγμα, προκειμένου για έγγραφα που χαρακτηρίζονται ως έγγραφα πολιτιστικού περιεχομένου, όπως εικόνες από πίνακες ζωγραφικής, η αυτόματη κατηγοριοποίηση τους μπορεί να γίνεται βάση της τεχνοτροπίας που τα χαρακτηρίζει, με τη βοήθεια μεθόδων ανάλυσης εικόνας. Αυτή η κατηγοριοποίηση μπορεί να αποτελέσει ένα τμήμα της εκπαίδευσης ενός χρήστη πάνω σε θέματα της ιστορίας της τέχνης. Η γνωριμία, για παράδειγμα, του χρήστη με αντιπροσωπευτικά έργα τέχνης του ιμπρεσιονισμού μπορεί να γίνει μέσα από το σύστημα ηλεκτρονικής μάθησης που έχει αυτόματα χαρακτηρισμένο πολιτιστικό υλικό στην κατηγορία *ιμπρεσιονισμός*.

Για να μπορέσει ένα σύστημα αυτόματα να τεκμηριώσει το πολιτιστικό υλικό θα πρέπει να ενσωματώσει τη γνώση του συγκεκριμένου πεδίου. Αναλυτικότερα, για να μπορέσει να θεωρηθεί ένα ψηφιακό έργο δείγμα συγκεκριμένου καλλιτεχνικού ρεύματος π.χ. ιμπρεσιονισμού, θα πρέπει να χαρακτηρίζεται από τις βασικές τεχνικές του ρεύματος, π.χ. μικρές και συχνά εμφανείς πινελιές που δημιουργούν ένα χαρακτηριστικά παχύ στρώμα μπογιάς στον καμβά, χρήση κυρίως των βασικών χρωμάτων, σπάνια χρήση του μαύρου χρώματος, απουσία διαδοχικών επιστρώσεων

χρώματος, έμφαση στον τρόπο που το φως ανακλάται πάνω στα αντικείμενα, φωτεινά και έντονα χρώματα. Με βάση τα παραπάνω, το σύστημα που θα κατασκευασθεί για τον αυτόματο χαρακτηρισμό του υλικού, πρέπει να λαμβάνει υπόψη τη γνώση του συγκεκριμένου πεδίου. Έτσι η ενσωμάτωση της γνώσης συγκεκριμένων πεδίων σε ένα σύστημα ηλεκτρονικής μάθησης μπορεί να συμβάλλει στην ανάλυση του υλικού αλλά και στην περαιτέρω χρήση του για σκοπούς μάθησης.

Μελετώντας διάφορα πολυμεσικά εκπαιδευτικά υλικά, διαπιστώθηκε η συχνή εμφάνιση προσώπων μέσα σε εικόνες και ειδικότερα σε πολιτιστικό εκπαιδευτικό υλικό όπως φωτογραφίες με ιστορικά πρόσωπα, έργα τέχνης, κ.α. Έτσι η έρευνα στράφηκε στην αναζήτηση καλλιτεχνικής ιστορικής περιόδου με ιδιαίτερη έμφαση στην απεικόνιση προσώπων και στην εύκολη πρόσβαση σε πλούσιο ψηφιοποιημένο υλικό. Μέσω αυτής της έρευνας οδηγηθήκαμε στη χρήση ψηφιοποιημένων έργων ζωγραφικής της Βυζαντινής τέχνης. Έτσι οδηγηθήκαμε στη ζωγραφική της Βυζαντινής τέχνης, στην οποία κατά βάση εμφανίζονται απεικονίσεις του Χριστού, της Παναγίας και των Αγίων. Σε αυτό το πλαίσιο αναζητήσαμε πληροφορίες σχετικές με τις τεχνικές που χρησιμοποιούνται στο συγκεκριμένο είδος ζωγραφικής, την Αγιογραφία. Σε αυτό το πλαίσιο ανακαλύψαμε ότι πολλές βυζαντινές εικόνες έχουν ψηφιοποιηθεί δημιουργώντας ψηφιακές βάσεις εικόνων τέχνης και προκαλώντας την ανάγκη για ικανοποιητική διαχείριση και ανάκτησή τους.

Από τη συνεργασία που είχαμε με ειδικούς της βυζαντινής τέχνης, καταγράψαμε τη γνώση του πεδίου με χρήση οντολογίας. Μετά από υπόδειξη των ειδικών, προσανατολιστήκαμε στην ανάλυση της περιοχής του προσώπου των μορφών που απεικονίζονται στις αγιογραφίες, με σκοπό την εξαγωγή πληροφοριών για κάθε μορφή. Στη συνέχεια, χρησιμοποιώντας τεχνικές ανίχνευσης προσώπου και των χαρακτηριστικών αυτού και στηριζόμενοι στη γνώση που καταγράψαμε, αναλύσαμε την περιοχή του προσώπου. Οι πληροφορίες αυτές δόθηκαν σε μία μηχανή συμπερασματολογίας, η οποία εφοδιάστηκε με τη γνώση του πεδίου με αποτέλεσμα την αυτόματη κατηγοριοποίηση της μορφής.

Σε αυτό το κεφάλαιο περιγράφεται η γνώση για τη βυζαντινή τέχνη που χρησιμοποιήθηκε από το σύστημα με σκοπό την ανάλυση και την εξαγωγή πληροφοριών από ψηφιοποιημένες Βυζαντινές εικόνες, καθώς και οι τεχνικές που προτάθηκαν, στο πεδίο της ψηφιακής ανάλυσης εικόνας ενώ παρουσιάζονται και τα πειραματικά αποτελέσματα των προτεινόμενων συστημάτων.

4.2 Ερευνητικό πλαίσιο

Κατά τις τελευταίες δεκαετίες, σημαντικές πρωτοβουλίες ψηφιοποίησης έχουν λάβει χώρα παγκοσμίως με σκοπό τη διατήρηση της πολιτισμικής κληρονομιάς. Λόγω της εξέλιξης της τεχνολογίας, έχει αρχίσει, με ταχύτατους ρυθμούς, η ψηφιοποίηση πολιτιστικού υλικού και η δημιουργία ψηφιακών συλλογών. Εντούτοις, διάφορα θέματα προέκυψαν σχετικά με τη δημιουργία των ψηφιακών πολιτιστικών δεδομένων και την αποθήκευσή τους καθώς και με τη δημιουργία συστημάτων για αναζήτηση και ανάκτηση του ψηφιακού υλικού, την ασφάλεια στη μεταφορά του και τη σωστή χρήση του. Όσον αφορά στη δημιουργία των ψηφιακών πολιτιστικών συλλογών, σε Ευρωπαϊκό επίπεδο, εκδόθηκαν σε κάθε χώρα κανόνες σωστής ψηφιοποίησης των πολιτιστικών δεδομένων ανά κατηγορία καθώς και χρήση κοινών προτύπων για τις διάφορες μορφές μεταδεδωμένων με σκοπό την αποτελεσματική τεκμηρίωση του υλικού.

Διάφορα πρότυπα, τα οποία διασφαλίζουν την διαλειτουργικότητα, έχουν αναπτυχθεί και έχουν χρησιμοποιηθεί κατά τη διαδικασία τεκμηρίωσης υλικού, με παράλληλη χρήση τεχνολογιών σημασιολογικού Ιστού [22], έχοντας ως στόχο τη σημασιολογική περιγραφή των ψηφιακών πολιτιστικών συλλογών. Σε αυτό το πλαίσιο, ο σημασιολογικός σχολιασμός του πολιτιστικού περιεχομένου απαιτεί ειδικούς τεκμηριωτές ανά κατηγορία πολιτιστικού υλικού, γεγονός που συνεπάγεται χρονοβόρες και πολυέξοδες διαδικασίες, αν αναλογιστεί κανείς τον τεράστιο όγκο του υλικού προς τεκμηρίωση. Αυτοί ήταν και οι λόγοι οδήγησαν στη χρήση τεχνικών για την αυτόματη ή ημιαυτόματη ανάλυση και ταξινόμηση του ψηφιοποιημένου πολιτιστικού υλικού.

Η χρήση των τεχνικών ψηφιακής ανάλυσης εικόνας έχει αρχίσει να προτείνεται στην ερευνητική κοινότητα ενιχνύοντας με αυτόν τον τρόπο τις συνεργασίες μεταξύ ερευνητικών ομάδων από διαφορετικά ερευνητικά πεδία (π.χ. ιστορικών τέχνης και μηχανικών υπολογιστών) [73] [82]. Τον Ιούλιο του 2008, στο περιοδικό του IEEE Signal Processing Magazine, παρουσιάστηκαν μερικές ιδιαίτερα ενδιαφέρουσες εφαρμογές της ανάλυσης εικόνας στην περιοχή της πολιτιστικής κληρονομιάς. Στην εργασία [110], περιγράφεται η πολυφασματική ανάλυση του ψηφιοποιημένου έργου Mona Lisa, ενώ στην εργασία [103] προτείνεται μια πρωτότυπη πολυφασματική τεχνική ψηφιακής απεικόνισης για περιοχές εικόνας σε ένα σύνολο εικόνων που περιέχουν παρόμοια χαρακτηριστικά όταν εκτίθενται στην ηλεκτρομαγνητική ακτινοβολία. Αυτή η τεχνική παρέχει τον εντοπισμό του υλικού και τον προσδιορισμό της θέσης του μέσα σε ολόκληρη την ζωγραφισμένη επιφάνεια. Η συνεργασία μεταξύ διαφορετικών ερευνητικών περιοχών είναι εμφανής στην εργασία [73], όπου τρεις ερευνητικές ομάδες ανέπτυξαν σύστημα που προσδιορίζει αυτόματα το ζωγράφο ενός πίνακα βάσει των χαρακτηριστικών πινελιών- της τεχνοτροπίας δηλαδή του έργου. Αυτά τα χαρακτηριστικά προσδιορίστηκαν από τους ιστορικούς τέχνης σύμφωνα με τη γνώση και τις τεχνικές που χρησιμοποιούν όταν καλούνται να επικυρώσουν ένα έργο.

Εντούτοις, αυτές οι μέθοδοι χρησιμοποιούν μόνο τη χαμηλού επιπέδου ανάλυση των εικόνων, χωρίς να χρησιμοποιούν τη γνώση του πεδίου. Δύο θέματα προέκυψαν σχετικά με τη χρήση τεχνικών ανάλυσης εικόνων με σκοπό τον αυτόματο χαρακτηρισμό ψηφιακών πολιτιστικών εικόνων. Το πρώτο είναι η αποτυχία των σημασιολογικών αλγορίθμων κατάτμησης και ανάλυσης εικόνας υπό πραγματικές συνθήκες. Αυτό οφείλεται στην υψηλή μεταβλητότητα του περιεχομένου της εικόνας και σε περιβαλλοντικές παραμέτρους (όπως φωτεινότητα), τα οποία συμβάλλουν στην πολυπλοκότητα του προβλήματος. Η δεύτερη δυσκολία είναι η εκτενής και ασαφής φύση της γνώσης του πολιτιστικού περιεχομένου της εικόνας, το οποίο περιπλέκει την αναπαράσταση της τυπικής γνώσης και της συλλογιστικής.

Ωστόσο, μερικές πολιτιστικές περιοχές είναι κατάλληλες για χρήση αυτόματων μεθόδων ανάλυσης εικόνας και τεκμηρίωσης. Οι βυζαντινές εικόνες ανήκουν σε αυτήν την κατηγορία. Το προκαθορισμένο περιεχόμενο εικόνας και η χαμηλή μεταβλητότητα των χαρακτηριστικών της οδηγεί στην επιτυχή εφαρμογή των μεθόδων ανάλυσης εικόνας. Σε αυτό το πλαίσιο αναπτύξαμε ένα σύστημα που εκμεταλλεύεται τη γνώση του πεδίου για να αναλύσει αυτόματα τις εικόνες. Το σύστημά μας ανιχνεύει ένα σύνολο εννοιών και ιδιοτήτων που απεικονίζονται στην εικόνα, σύμφωνα με το οποίο αναγνωρίζεται το πρόσωπο της Άγιας μορφής, και συνεπώς γίνεται σημασιολογική ταξινόμηση της και ανίχνευση συγκεκριμένων μορφών σε Βυζαντινές εικόνες. Στη συνέχεια παρουσιάζεται αναλυτικά η γνώση του πεδίου, όπως αυτή καταγράφηκε στο [132] καθώς και το σύστημα που αναπτύχθηκε σε αυτό το πλαίσιο.

4.3 Ερμηνεία της Βυζαντινής τέχνης- Διονύσιος εκ Φουρνά

Με τον όρο Βυζαντινή τέχνη αναφερόμαστε γενικά στην καλλιτεχνική παραγωγή και έκφραση που αναπτύχθηκε την περίοδο της Βυζαντινής Αυτοκρατορίας μεταξύ του 4ου αιώνα μ.Χ. και της άλωσης της Κωνσταντινούπολης το 1453. Η βυζαντινή τέχνη θεωρείται ότι αναπτύχθηκε αρχικά στην πρωτεύουσα της αυτοκρατορίας, την Κωνσταντινούπολη, αλλά επεκτάθηκε στο μεγαλύτερο τμήμα του μεσογειακού κόσμου και ανατολικά ως την Αρμενία. Υπήρξε αποτέλεσμα της αλληλεπίδρασης της αρχαίας ελληνικής παράδοσης και της ανατολικής επίδρασης και θρησκευτικότητας. Με την επικράτηση του Χριστιανισμού, η βυζαντινή τεχνοτροπία χαρακτηρίζεται από τα στοιχεία μιας αμιγώς θρησκευτικής τέχνης, που θεωρήθηκε παράδοση για τους ορθόδοξους χριστιανούς, που έζησαν υπό τον οθωμανικό κλοιό, τη μεταβυζαντινή περίοδο (16ος- 19ος αιώνας).

Οι ζωγράφοι της βυζαντινής και της μεταβυζαντινής περιόδου χρησιμοποίησαν εικονογραφικά πρότυπα για την δημιουργία των μορφών και των σκηνών που απεικονίζονται στα έργα τους. Για κάθε πρότυπο εμφανίστηκαν διαφοροποιήσεις, ανάλογες με την εξέλιξη των καλλιτεχνικών τάσεων. Στις αρχές του 18ου αιώνα εμφανίστηκαν στον ελλαδικό χώρο τρεις πραγματείες για τη βυζαντινή ζωγραφική. Τρεις ζωγράφοι με τάσεις θεωρητικής προσέγγισης της αγιογραφικής τέχνης, εκπρόσωποι διαφορετικών ρευμάτων, επιχείρησαν ο καθένας από τη σκοπιά του να δώσουν οδηγίες στους νεότερους ζωγράφους για την κατεύθυνση της αληθινής τέχνης στη βυζαντινή αγιογραφία. Ο Παναγιώτης Δοξαράς, από τη Ζάκυνθο, εκφραστής της ιταλικής αναγεννησιακής σχολής, έγραψε το Περί ζωγραφίας, γύρω στο 1726, έχοντας ως πρότυπο τον Λεονάρντο ντα Βίντσι, τον Αλμπέρτι και τον Βερονέζε. Ο Χριστόφορος Ζεφάροβιτς, ιερομόναχος Σέρβος, έγραψε στο πρώτο μισό του 18ου αιώνα στο Άγιον Όρος την Ερμηνεία Ζωγραφικής, επηρεασμένος κυρίως από τη ζωγραφική παράδοση της Αυστρίας και της Γερμανίας. Τέλος, ο Διονύσιος ο εκ Φουρνά Ευρυτανίας, έχοντας ως πρότυπο την καθαρή βυζαντινή τέχνη, συνέγραψε την Ερμηνεία της Ζωγραφικής Τέχνης [132], μεταξύ 1728-1733, στηριζόμενος στις αρχές της ζωγραφικής του εκ Θεσσαλονίκης Μανουήλ Πανσέληνου, ο οποίος ανιστόρησε περικαλλείς ναούς στο Άγιο Όρος, χωρίς να τους προσδιορίζει. Αυτή η τριπλή θεώρηση της βυζαντινής τέχνης στις αρχές του 18ου αιώνα δείχνει να ορίζει κατά κάποιον τρόπο την κατά περιοχές εξέλιξη της βυζαντινής παράδοσης, όσο και τα όρια μεταξύ βυζαντινής [20], [77],[111] και μεταβυζαντινή τέχνη [99], [35].

Τα περισσότερα και σημαντικότερα δείγματα ζωγραφικής τέχνης ανήκουν στη Μακεδονική σχολή (13ος-14ος αι) και την Κρητική σχολή (16ος αι.). Από τους προηγούμενους αιώνες έχουμε μικρό πλήθος δειγμάτων, λόγω αλληλεπάλληλων καταστροφών (και τα περισσότερα από αυτά προέρχονται από τη Μονή του Όρους Σινά- Αγ. Αικατερίνης). Κύριος εκπρόσωπος της Μακεδονικής σχολής ήταν ο Μανουήλ Πανσέληνος (με αντιπροσωπευτικά του έργα τοιχογραφίες στο Πρωτάτο του Αγίου Όρους) που υπήρξε και κύριος εμπνευστής του έργου του Διονυσίου εκ Φουρνά. Σημεία αναφοράς αυτής της σχολής είναι ρεαλισμός στην απεικόνιση των μορφών, όχι μόνον στα εξωτερικά χαρακτηριστικά, αλλά και στην απόδοση του εσωτερικού κόσμου των μορφών και κυρίως του πάθους. Οι συνθέσεις είναι πολυπρόσωπες, όλες οι μορφές κινούνται μέσα στο χώρο, ο οποίος είναι διευρυμένος και αποδίδεται με αξιοσημείωτο βάθος. Οι τοιχογραφίες του Μανουήλ Πανσέληνου ξεχωρίζουν από τα φωτεινά τους χρώματα, την ολοκληρωμένη απόδοση

των μορφών, την αναδυόμενη πνευματικότητα αλλά και τη μεγαλοπρέπεια των ζωγραφικών συνθέσεων.

Ο 16ος αιώνας σηματοδοτεί την απομάκρυνση από τα χαρακτηριστικά της Μακεδονικής σχολής και την επικράτηση των χαρακτηριστικών της Κρητικής σχολής με κύριο εκπρόσωπό της τον Θεοφάνη Στρυλίζας από το Ηράκλειο Κρήτης, γνωστό ως Θεοφάνης ο Κρης. Αντιπροσωπευτικό δείγμα του Θεοφάνη του Κρήτα είναι οι τοιχογραφίες στο καθολικό της μονής Μεγίστης Λαύρας που χαρακτηρίζονται από μια σταθερή οργάνωση της σύνθεσης με τέλειο στήσιμο των μορφών, πολλές πτυχώσεις στα ενδύματα που χαρίζουν μια φυσικότητα, ευγενικά πρόσωπα και φωτεινά χρώματα. Η πιστή απόδοση των εξωτερικών χαρακτηριστικών και η απεικόνιση της ψυχосύνθεσης δίνει τη θέση της σε μια πιο αυστηρή εικονογραφία. Οι μορφές είναι ψηλόλιγνες, λιπόσαρκες και διαθέτουν μια χαρακτηριστική ασκητική ευγένεια. Πιο πρόσφατα, στα μέσα του 20ου αιώνα, ο Φώτης Κόντογλου βασιζόμενος στη προγενέστερη συγγραφική παράδοση της "Ερμηνείας της Ζωγραφικής Τέχνης" του Διονυσίου εκ Φουρνά καθώς και στη γνώση που είχε αποκτήσει από τη μελέτη έργων της κρητικής σχολής, εξέδωσε το βιβλίο "Εκφρασις της Ορθοδόξου Εικονογραφίας", για το οποίο βραβεύθηκε από την Ακαδημία Αθηνών το 1959.

Η μελέτη της βυζαντινής και της μεταβυζαντινής ζωγραφικής εστιάζεται κυρίως στην αναγνώριση και στην ταξινόμηση των προτύπων και τυχών διακυμάνσεων τους για κάθε μορφή και κάθε παράσταση που απεικονίζεται σε ένα έργο. Αυτή η διεργασία οδηγεί στην εξαγωγή γνώσης σχετικής με τη χρονολόγηση της σχεδίασης μιας μορφής ή παράστασης και την απόδοση της περιόδου στην οποία αυτή δημιουργήθηκε.

Η βυζαντινή εικονογραφία ακολουθεί μια μοναδική σύμβαση της ζωγραφικής. Η καλλιτεχνική γλώσσα των βυζαντινών ζωγράφων χαρακτηρίζεται από την προφανή απλότητα και λιτότητα, από πλασματικά και συμβολικά χρώματα, από την έλλειψη προοπτικής και τις παράξενες αναλογίες. Οι μορφές των Αγίων τίθενται πέρα από πραγματικό χρόνο και χώρο μέσω της χρήσης των χρυσών φόντων. Η πιο σημαντική μορφή της εικόνας απεικονίζεται σε μετωπική θέση, εστιάζοντας στα μάτια, στην έκφραση του προσώπου και τα χέρια. Αυτή η στάση της μορφής βοηθά στην άμεση επαφή με τον θεατή. Η σαφήνεια είναι ο κανόνας, όχι μόνο στην περιγραφή των μορφών αλλά και στην τοποθέτηση πολλών μορφών σε μία εικόνα με σκοπό την δημιουργία ισορροπημένων συνθέσεων που να αποδίδουν αναπαραστάσεις από την θρησκευτική πίστη. Τα πρόσωπα των μορφών συνήθως έχουν μεγάλα μάτια και αυτιά, μακριές και λεπτές μύτες, και μικρά στόματα, και κάθε ένα από αυτά χρωματίζεται βάσει συγκεκριμένων κανόνων [132] προσπαθώντας με αυτό τον τρόπο να δείξουν την διαφοροποίηση τους από τους απλούς ανθρώπους.

Η ζωγραφική βυζαντινών εικόνων σκόπιμα δεν λαμβάνει υπόψη την αρχή της φυσικής προοπτικής προκειμένου να αποφευχθεί με οποιοδήποτε κόστος η παραίσθηση των τριών διαστάσεων. Αντ' αυτού, δίνει την εντύπωση της πλήρους λειότητας και της έλλειψης προοπτικής. Εντούτοις, στη βυζαντινή ζωγραφική χρησιμοποιείται μια προοπτική, που καλείται συνήθως αντιστρέφουσα ή αναστρέφουσα, η οποία εξαρτάται από τις πολλαπλές όψεις. Αυτές οι πολλαπλές όψεις τοποθετούνται μπροστά από τη μορφή, δημιουργώντας μεγεθυμένα ή ακόμη και διαστρεβλωμένα αντικείμενα φόντου.

Κατά την απεικόνιση των διάφορων μορφών, ο καλλιτέχνης ακολουθεί τους κανόνες του εγχειριδίου [132] σχετικά με τη στάση του σώματος, το στυλ των μαλλιών και της γενειάδας, την ενδυμασία και άλλες ιδιότητες. Κατά αυτόν τον τρόπο, συγκεκριμένες μορφές καθορίζονται από συγκεκριμένα χαρακτηριστικά του προσώπου. Παραδείγματος χάριν, ο Ιησούς απεικονίζεται ως νέος με μακριά, σπαστά

και πυκνά και σκούρα μαλλιά και με κοντή, ίσια και σκούρη γενειάδα.

Η συμβολή του έργου του Διονυσίου εκ Φουρνά στη μελέτη βυζαντινών και μεταβυζαντινών έργων τέχνης είναι σημαντικότερη. Το έργο του αποτελείται από δύο μέρη: την "Τεχνολογία", που αναφέρεται στην τεχνική των εικόνων και της τοιχογραφίας και βασίζεται σε παλιότερα ανάλογα κείμενα, και το "Εικονογραφικό μέρος", που περιέχει λεπτομερή περιγραφή των διαφόρων θρησκευτικών σκηνών και των χαρακτηριστικών των αγίων και αποτελεί το πιο πρωτότυπο τμήμα του βιβλίου. Μελετώντας το έργο του Διονυσίου εκ Φουρνά, μπορούμε να αντλήσουμε πληροφορίες για την αναγνώριση και την ταξινόμηση των μορφών βυζαντινών παραστάσεων και μορφών Αγίων.

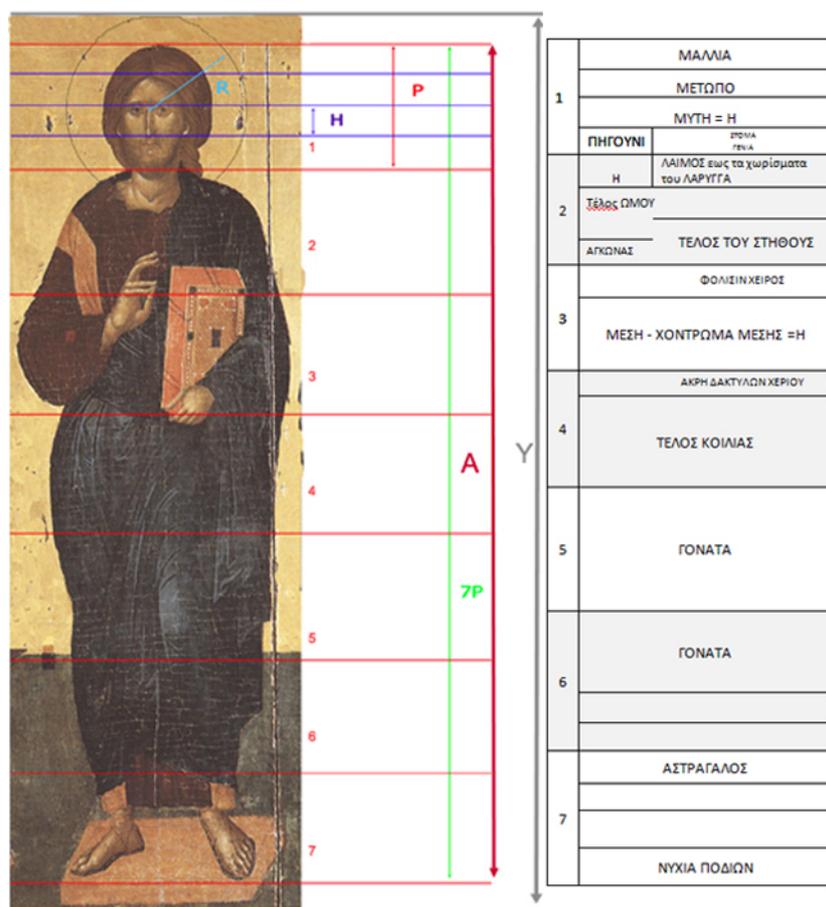
Στο πλαίσιο της έρευνά μας, στηριχθήκαμε στα μέτρα σχεδίασης που αναφέρει, στο εικονογραφικό μέρος του έργου του, ο Διονύσιος εκ Φουρνά, με σκοπό τη χωροχρονική μοντελοποίηση και αναγνώριση των προτύπων και τυχόν διακυμάνσεών τους που απεικονίζονται σε ψηφιακές βυζαντινές εικόνες. Απώτερος σκοπός ήταν η δημιουργία ενός συστήματος που θα μπορεί να "επεξεργάζεται" μια βυζαντινή ή μεταβυζαντινή εικόνα με τρόπο παρόμοιο με αυτόν που χρησιμοποιεί ένας ειδικός βυζαντινής τέχνης.

Σύμφωνα με το εικονογραφικό μέρος του έργου του Διονυσίου εκ Φουρνά, μια εικόνα που περιέχει μια μορφή Αγίου μπορεί να χωρισθεί σε σημασιολογικές περιοχές, καθεμία από τις οποίες έχει συγκεκριμένα χαρακτηριστικά που την κάνουν διαχωρίσιμη. Η μοντελοποίηση αφορά τις αναλογίες των μορφών (λόγοι μεγεθών κορμού - κεφαλής), το σχήμα των μορφών καθώς και τμημάτων αυτών. Το σύνολο αυτών των στοιχείων παρέχει πολύ σημαντική πληροφορία για τη χωροχρονική τοποθέτηση των μορφών, τη συσχέτιση αυτών και τα χαρακτηριστικά στοιχεία τεχνοτροπίας. Αποτέλεσμα της ανάλυσης είναι η κατάτμηση της εικόνας για τον εντοπισμό σημασιολογικών αντικειμένων και η αναγνώριση των αντικειμένων με στόχο τη σημασιολογική περιγραφή του περιεχομένου τους.

Ένας ειδικός βυζαντινής τέχνης όταν θελήσει να αναλύσει μια βυζαντινή εικόνα χρησιμοποιεί τους "κανόνες" που όρισε ο Διονύσιος εκ Φουρνά και τους οποίους οι αγιογράφοι ακολούθησαν κατά τη διάρκεια της δημιουργίας μιας εικόνας. Ουσιαστικά πρόκειται για μία μοντελοποίηση των στοιχείων της εικονογραφίας σε σχέση με το ανθρώπινο σώμα, τις αναλογίες του, τα προσωπογραφικά χαρακτηριστικά και την ενδυματολογία των μορφών. Ειδικότερα τα προσωπογραφικά και τα ενδυματολογικά χαρακτηριστικά είναι καθοριστικά για τον προσδιορισμό της ηλικίας, του φύλου και της ιδιότητας της απεικονιζόμενης μορφής. Γενικότερα, ο κάθε αγιογράφος χρησιμοποιεί τους κανόνες για τη δημιουργία των αρμονικών διαστάσεων προσθέτοντας τη δική του πινελιά στις λεπτομέρειες.

Έτσι, όταν ένας ειδικός βυζαντινής τέχνης αναλύει μια μορφή σε μια βυζαντινή εικόνα, ουσιαστικά ακολουθεί τους κανόνες του Διονυσίου εκ Φουρνά. Δηλαδή χωρίζει την εικόνα σε τμήματα όμοια με αυτά που ο αγιογράφος χωρίζει την επιφάνεια όπου πρόκειται να ζωγραφίσει μια μορφή. Σύμφωνα με τους κανόνες, αρχικά ένας αγιογράφος χωρίζει την περιοχή στην οποία θέλει να ζωγραφίσει σε ίσα μέρη, για παράδειγμα για την απεικόνιση μιας ιστάμενης μορφής απαιτείται ο χωρισμός της περιοχής σε επτά τμήματα ίσου ύψους, έστω P . Σε κάθε ένα από τα τμήματα αυτά θα απεικονίζεται ένα σημασιολογικό κομμάτι της μορφής (Σχήμα 4.1).

Αναλυτικότερα, το πρώτο τμήμα περιέχει την κεφαλή της μορφής, σε ελλειπτικό σχήμα, εστιάζοντας στα μάτια και στην έκφραση του προσώπου. Αυτό το τμήμα περιέχει τα πιο χαρακτηριστικά στοιχεία της μορφής, για παράδειγμα τα μαλλιά και



Σχήμα 4.1: Αναπαράσταση μέτρων για ιστάμενες Μορφές σύμφωνα με τα μέτρα του Διονυσίου εκ Φουρνά

τη γενειάδα τα οποία μπορούν να τη χαρακτηρίσουν. Τα εναπομείναντα τμήματα περιέχουν το σώμα της μορφής, εστιάζοντας στα χέρια, στο ένδυμα και στα αντικείμενα που τυχόν κρατά η μορφή. Το σώμα της μορφής μπορεί να απεικονισθεί σε διαφορετικές πόζες όπως ιστάμενη, καθίμενη, 3/4, έως τη μέση και έως το θώρακα.

Σε κάθε πρόσωπο που σχεδιάζεται υπάρχει ένας κεντρικός άξονας, που ονομάζεται άξονας αντικίνησης (contrapposso), και το χωρίζει σε δύο άνισα μέρη. Αυτός ο άξονας δίνει στη μορφή μια δυναμική αντιστάθμιση η οποία την κάνει να φαίνεται πιο "ζωντανή" προς τον θεατή. Όταν υπάρχουν πολυπρόσωπες εικόνες η στροφή του προσώπου είναι πιο έντονη (μετατόπιση του άξονα κατά 3/4). Τα μέρη του προσώπου που βρίσκονται στη στενότερη πλευρά σχεδιάζονται με το ίδιο ύψος, αλλά με μικρότερο πλάτος από τα αντίστοιχα μέρη της φαρδύτερης πλευράς. Για παράδειγμα, αν η μορφή έχει δεξιά κίνηση, τα στοιχεία του προσώπου (μάτια, αυτιά παρειές) που βρίσκονται στη δεξιά πλευρά της μορφής θα απεικονίζονται πιο έντονα σε σχέση με τα στοιχεία της αριστερής πλευράς.

Στον πίνακα που απεικονίζεται στο Σχήμα 4.1 περιγράφονται τα τμήματα στα οποία χωρίζεται μια περιοχή μέσα στην οποία πρόκειται να ζωγραφιστεί μια μορφή καθώς και τι περιέχουν τα τεμάχια που απαρτίζουν το κάθε τμήμα. Για παράδειγμα, στο πρώτο τμήμα ύψους P θα ζωγραφιστεί το κεφάλι της μορφής ως εξής: στο πρώτο τεμάχιο θα ζωγραφιστούν τα μαλλιά, στο δεύτερο το μέτωπο, στο τρίτο η περιοχή της μύτης και στο τέταρτο η περιοχή που περιέχει το στόμα, το πηγούνι, το μουστάκι (αν υπάρχει) και τμήμα του γενιού (αν υπάρχει). Το δεύτερο τμήμα, επίσης

ύψος P , περιέχει την απεικόνιση από το λαιμό έως το θώρακα, ενώ το τρίτο τμήμα περιέχει την απεικόνιση από το θώρακα έως τον αγκώνα και καταλαμβάνει ύψος P . Το τέταρτο τμήμα θα περιέχει την κοιλιακή χώρα, ενώ το πέμπτο τμήμα θα απεικονίζει την περιοχή έως το γόνατο της μορφής. Το έκτο τμήμα θα απεικονίζει την περιοχή έως τον αστράγαλο, ενώ το έβδομο έως την άκρη των ποδιών της μορφής. Ανάλογα με ποια πόζα θέλει να αποδώσει στη μορφή ο αγιογράφος, χρησιμοποιεί τα κατάλληλα τμήματα. Αντιπαραθέτοντας τον πίνακα με την εικόνα μπορεί κανείς να διαπιστώσει τους "κανόνες" που ακολούθησε ο αγιογράφος. Τους ίδιους "κανόνες" προσπαθεί να χρησιμοποιήσει και ένας ειδικός βυζαντινής τέχνης, ώστε να εξάγει σημασιολογική πληροφορία από μία βυζαντινή εικόνα.

Σύμφωνα με τον Διονύσιο εκ Φουρνά, η μύτη αποτελεί τη μετρική για την παραγωγή αρμονικών αναλογιών του προσώπου και του σώματος. Για παράδειγμα το ύψος του τμήματος της κεφαλής είναι ίσο με τέσσερις μύτες ενώ το πλάτος του είναι ίσο με τρεις μύτες. Όλο το ύψος της μορφής θα βρίσκεται σε απλή αναλογία με το ύψος της μύτης. Επιπλέον, η απόσταση μεταξύ της ίριδας των ματιών είναι πάλι μία μύτη. Τα αυτιά ζωγραφίζονται παράλληλα με τη μύτη και λίγο μικρότερα σε ύψος από αυτήν, ενώ το πλάτος τους είναι μισή μύτη. Το ύψος των μαλλιών είναι ίσο με τη μύτη. Ο άξονα αντικίνησης περνά από τον κάθετο άξονα της μύτης και από το κέντρο του στόματος και διαιρεί το πρόσωπο σε δύο άνισα μέρη, οπότε το ένα μέρος του προσώπου είναι μικρότερο από το άλλο. Το στενό μέρος του προσώπου θεωρείται ότι βρίσκεται πλησιέστερα στην πηγή του φωτός. Άρα οι σκιές της μύτης και των παρειών είναι εντονότερες στο πλατύτερο μέρος. Γενικότερα, στο πλατύτερο μέρος σχεδιάζονται τα τμήματα του προσώπου μεγαλύτερα από ότι στο στενό μέρος.

Όσον αφορά στη σχεδίαση των οφθαλμών, το σχήμα τους είναι ωσειδές, με πλάτος διπλάσιο από το ύψος τους. Η θέση της ίριδας μέσα στον οφθαλμό καθορίζει και την έκφραση της μορφής: όσο απομακρύνεται από το επάνω βλέφαρο και κινείται προς τα κάτω τόσο η έκφραση από γαλήνια γίνεται πιο έντονη και πιο άγρια. Τα φρύδια επίσης παίζουν σημαντικό ρόλο στην έκφραση. Το γαλήνιο φρύδι ξεκινά από τη ρίζα της μύτης και διαγράφεται πάνω από το άνω βλέφαρο ως το τέλος του. Όσο το κέντρο του τόξου του φρυδιού απομακρύνεται από το άνω βλέφαρο, τόσο πιο αυστηρή είναι η έκφραση. Γενικότερα, οι κάθετες γραμμές στο πρόσωπο προσδίδουν αυστηρότητα.

Στη συνέχεια παρουσιάζουμε τις αναλογίες μεταξύ ύψος μύτης και του ύψους της μορφής. Ας θεωρήσουμε ως P το ύψος κάθε τμήματος. Αρχίζοντας με φορά από την κεφαλή προς τα πόδια, χωρίζουμε το πρώτο τμήμα σε τέσσερα μικρότερα μέρη, τα τεμάχια. Αν συμβολίσουμε με H το ύψος καθενός από αυτά τα τεμάχια, θα ισχύει: $H = P/4$, όπου H είναι το ύψος της μύτης.

Αν ένας αγιογράφος θελήσει να ζωγραφίσει μια μορφή ως την μέση τότε αρκεί να χωρίσει την περιοχή που πρόκειται να ζωγραφίσει σε 3 ίσα τμήματα (δες Σχήμα 4.1). Αν πάλι θελήσει ο αγιογράφος να ζωγραφίσει μια ιστάμενη μορφή αρκεί να χωρίσει την περιοχή σχεδίασης σε 7 ίσα μέρη. Ας θεωρήσουμε ως A την περιοχή στην οποία ένας αγιογράφος σκοπεύει να ζωγραφίσει μια μορφή, τότε θα ισχύει το εξής: $3 \cdot P \leq A \leq 7 \cdot P \Rightarrow 12 \cdot H \leq A \leq 28 \cdot H$

Συνήθως το ύψος Y μιας εικόνας είναι: $Y = A + 2 \cdot P \Rightarrow Y = A + 8 \cdot H$ άρα $12 \cdot H + 8 \cdot H \leq A + 8 \cdot H \leq 28 \cdot H + 8 \cdot H \Rightarrow 20 \cdot H \leq Y \leq 36 \cdot H$

Η τελευταία ανισότητα συσχετίζει την τιμή του ύψους των κομματιών με το ύψος όλης της εικόνας. Κατά αυτόν τον τρόπο, γνωρίζοντας το ύψος της μύτης καθώς και το ύψος της εικόνας μπορούμε να εξάγουμε με τη βοήθεια των μέτρων του Διονυσίου εκ Φουρνά και με τη βοήθεια τεχνικών εξαγωγή φόντου, τόσο την κεφαλή όσο και το

σώμα της μορφής.

Σύμφωνα με τον Διονύσιο εκ Φουρνά, το κεφάλι μιας μορφής πρέπει να περιβάλλεται από ένα φωτοστέφανο. Για τη σχεδίαση του φωτοστεφάνου ο αγιογράφος σχεδιάζει έναν κύκλο έχοντας ως κέντρο του τη μέση της μύτης της μορφή και με ακτίνα $R = 2.5 \cdot H$.

Στο πλαίσιο της παρούσας διδακτορικής διατριβής ασχοληθήκαμε με την ανάλυση της περιοχής του προσώπου βασιζόμενοι στο εγχειρίδιο του Διονυσίου. Σύμφωνα με το εγχειρίδιο αυτό, τα χρώματα που χρησιμοποιούνται στην περιοχή του προσώπου μιας μορφής είναι γήινα χρώματα, διαλυμένα με κρόκο αυγού, νερό και ξύδι, εξασφαλίζοντας με αυτόν τον τρόπο τη διατήρησή τους στο πέρασ των χρόνων. Γενικά η περιοχή των σαρκωμάτων του προσώπου ζωγραφίζεται με επάλληλα στρώματα χρωμάτων, ξεκινώντας πάντα από τις σκούρες αποχρώσεις και φτάνοντας σταδιακά στις ανοιχτότερες.

Τα χρώματα των σαρκωμάτων του προσώπου καθορίζονται μοναδικά από τον αγιογράφο. Αρχικά, ο αγιογράφος δημιουργεί το βασικό χρώμα για τα σαρκώματα που καλείται *σάρκωμα*. Αναμειγνύοντας το σάρκωμα με μαύρο χρώμα δημιουργεί το σκούρο χρώμα που καλείται *σκούρεμα* απλώνεται ως βάση σ' όλη την επιφάνεια των σαρκωμάτων και είναι ο τόνος που θα μείνει στο τέλος στο σκιαζόμενο μέρος. Με την ανάμειξη των δύο χρωμάτων δημιουργείται ένα ενδιάμεσο χρώμα που καλείται *προπλασμός* και χρησιμοποιείται για να ενώσει τα σημεία του σαρκώματος με τα σημεία σκουρέματος, δημιουργώντας μια ομαλή μετάβαση. Έπειτα αναμειγνύοντας το σάρκωμα με λευκό χρώμα δημιουργείται το χρώμα που χρησιμοποιείται για τα φωτίσματα του προσώπου και καλείται *φώτισμα*. Το φώτισμα ουσιαστικά συμβάλλει στο να είναι ικανοποιητικός ο όγκος που θα πάρει το πρόσωπο. Στο τελικό στάδιο, ο αγιογράφος για να τονίσει τα φωτίσματα βάζει τις ψιμμυθίες, πολύ λεπτές πινελιές σε καθαρό, σχεδόν άσπρο χρώμα, με τις οποίες επεμβαίνει καθοριστικά στην έκφραση του προσώπου της μορφής. Με ανάλογο τρόπο σχεδιάζονται τα μαλλιά, η γενειάδα και το μουστάκι.

Το εγχειρίδιο του Διονυσίου μας παρέχει συγκεκριμένες πληροφορίες για το κάθε σημασιολογικό τμήμα του προσώπου. Για παράδειγμα, ο Άγιος Δημήτριος απεικονίζεται με συγκεκριμένο τρόπο ως προς τα χαρακτηριστικά του προσώπου, δηλαδή εμφανίζεται πάντα, σε όλες τις εικόνες, χωρίς ρυτίδες, αγένειος, με παρειές νέου, μαυρομάλλης, με σπαστά μαλλιά, μήκους κάτω από τα αυτιά. Τα σημασιολογικά τμήματα στα οποία θεωρούμε ότι μπορεί να αναλυθεί το πρόσωπο και να εξάγουμε πληροφορίες για τη μορφή που απεικονίζεται είναι η περιοχή των μαλλιών, η περιοχή του μετώπου, η περιοχή των παρειών, η περιοχή του μουστακιού και η περιοχή της γενειάδας. Έχοντας ως πηγή το εγχειρίδιο καταγράψαμε στους πίνακες 4.1, 4.2, 4.3, 4.4, 4.5 τις ιδιότητες που μπορεί να έχουν αυτά τα τμήματα.

4.4 Σύστημα αυτόματης κατηγοριοποίησης Βυζαντινών εικόνων

Στο πλαίσιο της αυτόματης τεκμηρίωσης κατηγοριοποίησης Βυζαντινών εικόνων προτάθηκαν δύο συστήματα. Το πρώτο σύστημα στηρίχθηκε στον εντοπισμό του φωτοστέφανου θέλοντας να κατηγοριοποιήσει τις εικόνες, ενώ το δεύτερο στη χρήση όλης της γνώσης, που προσφέρει το εγχειρίδιο του Διονυσίου εκ Φουρνά, με σκοπό την τεκμηρίωση αυτών. Στη συνέχεια παραθέτονται τα δύο συστήματα.

Πίνακας 4.1: Κατηγοριοποίηση Μαλλιών

Ύπαρξη Κώμης	Μήκος	Χρώμα	Σχήμα	Πυκνότητα
- Φαλακρός				
- Με Κώμη	- Τας τρίχας έχων άνωθεν των ωτίων - Τας τρίχας έχων κάτωθεν των ωτίων - Τας τρίχας έχων έως τον αριστερό ώμο με ένα πλοχμό - Τας τρίχας έχων έως το δεξιό ώμο με ένα πλοχμό - Με δύο πλοχμούς	- Μαυρομάλλης - Περικαπνισμένος - Ασπρομάλλης	- Βουρλομάλλης - Με τρίχας ηγριωμένας - Σγουροκέφαλος - Φουντομάλλης	- Δασυμάλλης - Ολιγομάλλης

Πίνακας 4.2: Κατηγοριοποίηση Μετώπου

Χρώμα	Αριθμός Ρυτίδων
-Φώτισμα	- έως 4
-Σκούρεμα	-πάνω από 4

4.4.1 Σύστημα αυτόματης κατηγοριοποίησης Βυζαντινών εικόνων με βάση εντοπισμό του φωτοστέφανου

Κατά τις πρώτες απόπειρες αυτόματης αναγνώρισης του πολιτισμικού υλικού και συγκεκριμένα των Βυζαντινών εικόνων υπήρξαν διάφορες ιδέες. Αρχικά, προτάθηκε μια ιδέα [109] που όμως δεν εκμεταλλευόταν πλήρως τη γνώση που παρέχει το εγχειρίδιο του Διονυσίου εκ Φουρνά. Ουσιαστικά προσπαθήσαμε να εντοπίσουμε το φωτοστέφανο που περιβάλλει την κεφαλή κάθε Άγιας μορφής που παρουσιάζεται σε μια Βυζαντινή εικόνα.

Για τον εντοπισμό του φωτοστεφάνου, χρησιμοποιήσαμε το μετασχηματισμό Hough για εντοπισμό κύκλων [17], προσδιορίζοντας τον ακριβή κύκλο. Έχοντας την ακτίνα του φωτοστέφανου R μπορούμε να εξάγουμε τη μετρική - το ύψος της μύτης - $H = R/2.5$, όπως αναφέρθηκε στην Ενότητα 4.3. Έπειτα εγγράφουμε δύο ομόκεντρους, με το φωτοστέφανο, κύκλους, ο ένας εκ των οποίων έχει ακτίνα $1.5 \cdot H$ και ο άλλος ακτίνα $2 \cdot H$. Ο μικρός κύκλος, σύμφωνα με τον Διονύσιο, περιέχει την κεφαλή, ενώ ο δακτύλιος που σχηματίζεται από τους εγγεγραμένους κύκλους περιλαμβάνει την περιοχή μεταξύ φωτοστέφανου και κεφαλής (συνήθως περιέχει το χρώμα του φόντου της εικόνας). Για αυτές τις δύο περιοχές υπολογίζουμε τη μέση τιμή της φωτεινότητά τους. Επιλέγοντας κάποιο κατώφλι μεταξύ αυτών των μέσων τιμών, γίνεται η κατάτμηση των δύο περιοχών, του φωτοστέφανου και της κεφαλής.

Τέλος, εφαρμόζουμε ένα φίλτρο ενδιάμεσης τιμής (median filter) κατάλληλου παραθύρου στην κατατμημένη εικόνα προκειμένου να παραχθούν οι μάσκες που

Πίνακας 4.3: Κατηγοριοποίηση Παρειάς

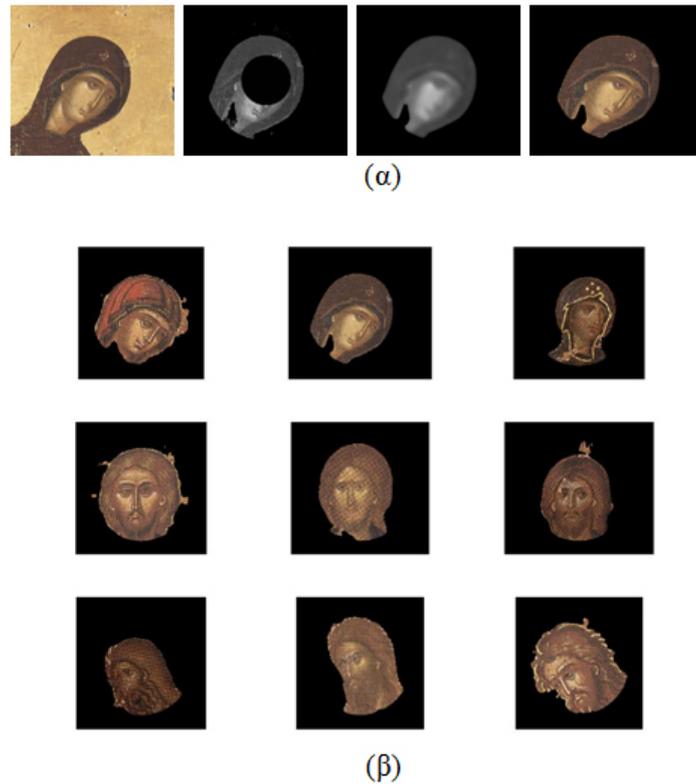
Άνω τμήμα	Κάτω τμήμα
-Φώτισμα	-Φώτισμα
-Σκούρεμα	-Σκούρεμα

Πίνακας 4.4: Κατηγοριοποίηση Μουστακίου

Ύπαρξη	Μορφή
-Μουστακίων	-Με τρίχας Σγουράς
-Χωρίς μουστάκι	-Με τρίχας ίσιας

Πίνακας 4.5: Κατηγοριοποίηση Γενειάδας

Ύπαρξη	Χρώμα	Μορφή	Πυκνότητα	Σχήμα	Μήκος	Κυματισμός
-σπανός	-Μαυρογένης	-μη χωρίζων το γένυ	-δασυγένης	-οξυγένης	- κοντογένης	-με τρίχας ηγριωμένας
-αγένειος	- Περικαπνισμένο	-διχалоγένης	- ολιγογένης	- στρογγυλογένης ολίγον	-μακρυγένης ου πολλά	- κατζαρογένης
- αρχιγένης	-Ασπρογένης	- τριχалоγένης		- στρογγυλογένης	-μακρυγένης	
-με γένυ		-εις τέσσερα χωρίζων το γένειον		-πλατυγένης	- μακρυγένης έως την μέσην	
		-εις πέντε χωρίζων το γένειον		- φουντογένης	- μακρυγένης έως την ζώνην	
					- μακρυγένης έως τα γόνατα	



Σχήμα 4.2: Εντοπισμός κεφαλής με χρήση φωτιστέφανου: (α) Εξαγωγή κεφαλής από εικόνα (β) Εξαχθείσες κεφαλές από διάφορες μορφές

απομονώνουν καλύτερα την κατατμημένη κεφαλή. Με την εφαρμογή αυτών των μασκών στις αρχικές εικόνες παράγονται οι τελικές εικόνες των κατατμημένων κεφαλών που χρησιμοποιούνται στο επόμενο τμήμα της ταξινόμησης (Σχήμα 4.2).

Για την κατηγοριοποίηση των εξαγόμενων κεφαλών (π.χ. Σχήμα 4.2.β) χρησιμοποιήσαμε κλασσικούς ταξινομητές, τους α) Naive Bayes (NB) [43], β) Flexible Naive Bayes (FNB) [90], γ) Support Vector Machines (SVM) [32] και δ) πολυστρωματικό Perceptron με μάθηση οπίσθιας διάδοσης (NNBP) [59]. Ως είσοδο σε αυτούς τους ταξινομητές δίνουμε ένα διάνυσμα αναπαράστασης της κάθε εξαγόμενης κεφαλής.

Η περιγραφή μιας περιοχής μιας διδιάστατης εικόνας μπορεί να βασιστεί σε χαρακτηριστικά μεγέθη που μπορούν να εξαχθούν από αυτήν. Αυτά τα χαρακτηριστικά μπορεί να προέρχονται είτε από το περίγραμμα ή από το εσωτερικό της περιοχής. Τα χαρακτηριστικά που εξάγονται από το περίγραμμα της περιοχής (ακμές, όρια) χρησιμοποιούνται όταν μας ενδιαφέρουν τα χαρακτηριστικά του σχήματος. Τα χαρακτηριστικά μεγέθη που εξάγονται από το εσωτερικό μιας περιοχής (το σύνολο των pixels από τα οποία αποτελείται) συνήθως χρησιμοποιούνται όταν ενδιαφερόμαστε για χαρακτηριστικά όπως φωτεινότητα, χρώμα, υφή, εμβαδόν κ.α. Τα χωρικά αυτά χαρακτηριστικά χωρίζονται περαιτέρω σε γεωμετρικά χαρακτηριστικά, στατιστικά χαρακτηριστικά και σε χαρακτηριστικά υφής και μετασχηματισμού.

Για τη δημιουργία του διανύσματος αναπαράστασης, το οποίο θα περιέχει την οπτική πληροφορία χαμηλού επιπέδου, χρησιμοποιούμε στο προτεινόμενο σύστημα κάποιους από τους περιγραφείς του προτύπου MPEG-7 ([34],[113]) καθώς και τις ροπές Hu [66]. Για να μειώσουμε τη διάσταση των συσχετισμένων MPEG-

7 περιγραφών, χρησιμοποιούμε την Ανάλυση Κυρίων Συνιστωσών (PCA). Τα αποτελέσματα των πειραμάτων για τον εντοπισμό εννοιών και την ταξινόμηση εικόνων έρχονται να ενισχύσουν ακόμη περισσότερο τη θέση ότι οι MPEG-7 περιγραφείς είναι ικανοί για αυτή την περιγραφή, ενώ οι περιγραφείς γεωμετρικών ροπών ενισχύουν τις διαδικασίες αποτελεσματικού εντοπισμού και ταξινόμησης σε περιπτώσεις περιστροφών, μετατόπισης και αλλαγής κλίμακας.

Η ανάκτηση εικόνων μέσω ταξινόμησης αποτελεί ένα ανοικτό και δύσκολο ερευνητικό πεδίο. Η ποικιλομορφία του συνόλου ελέγχου είναι συνήθως πολύ μεγαλύτερη από εκείνη του συνόλου εκπαίδευσης και τις περισσότερες φορές οι περιγραφείς που χρησιμοποιούνται για αναπαράσταση του περιεχομένου δεν μπορούν να χαρακτηρίσουν μονοσήμαντα το πρότυπο προς αναγνώριση. Για το λόγο αυτό τα σφάλματα αναγνώρισης ποικίλουν και μαζί με αυτά ποικίλει και η απόδοση κάθε αλγορίθμου ταξινόμησης.

4.4.1.1 Οπτικοί περιγραφείς του προτύπου MPEG-7

Οι οπτικοί περιγραφείς (visual descriptors) του προτύπου MPEG-7 που χρησιμοποιήσαμε είναι οι εξής: 1) περιγραφέας κύριων χρωμάτων (Dominant Color Descriptor), 2) κλιμακωτός περιγραφέας χρώματος (Scalable Color Descriptor), 3) περιγραφέας διάταξης χρώματος (Color Layout Descriptor), 4) περιγραφέας ομοιογενούς υφής (Homogeneous Texture Descriptor) και 5) περιγραφέας ιστογράμματος ακμών (Edge Histogram Descriptor).

Ο περιγραφέας κύριων χρωμάτων (Dominant Color Descriptor) αποτελείται από τα κύρια χρώματα (μέχρι 8 χρώματα), τα ποσοστά τους στην εικόνα ή σε μια περιοχή ενδιαφέροντος της, τη χωρική τους συνοχή (spatial coherency), και τη διακύμανση τους (variance) και ορίζεται από τη σχέση: $DCD = [(c_i, p_i, v_i), s], i = 1, 2 \dots N$ όπου c_i είναι το i -οστό κύριο χρώμα, p_i το ποσοστό του, v_i η διακύμανσή του και s η χωρική συνοχή.

Ο κλιμακωτός περιγραφέας χρώματος (Scalable Color Descriptor) ουσιαστικά είναι μια κωδικοποίηση του ιστογράμματος του χρώματος της εικόνας στον HSV χώρο. Η κωδικοποίηση πραγματοποιείται με τη βοήθεια ενός υψιπερατού και ενός βαθυπερατού φίλτρου (μετατροπή Haar). Για τη μείωση του μεγέθους του περιγραφέα, πραγματοποιείται στη συνέχεια κβάντιση του ιστογράμματος και τελικά το διάλυσμα έχει τη μορφή: $SCD = [c_1, c_2 \dots c_N]$ όπου N είναι το μέγεθος του ιστογράμματος.

Ο περιγραφέας διάταξης χρώματος (Color Layout Descriptor) αναπαριστά τη χωρική κατανομή του χρώματος της εικόνας ή μιας περιοχής ενδιαφέροντός της. Για την εξαγωγή του περιγραφέα, αρχικά διαιρείται η εικόνα σε 64 μπλοκ και εξάγεται, για κάθε ένα από αυτά, η μέση τιμή του χρώματος. Έτσι παράγεται μια εικόνα 8x8 που κάθε εικονοστοιχείο της έχει τη μέση τιμή του κάθε μπλοκ. Στην 8x8 εικόνα εφαρμόζουμε διακριτό μετασχηματισμό συνημιτόνου (ΔΜΣ- Discrete Cosine Transform - DCT) ξεχωριστά σε κάθε ένα από τα τρία χρωματικά κανάλια. Ο περιγραφέας εξάγεται χρησιμοποιώντας zigzag scanning από τους πρώτους 12 μη ομοιόμορφα κβαντισμένους συντελεστές του μετασχηματισμού για το κάθε κανάλι και ορίζεται από τη σχέση: $CLD = \{ \{DY_{DC} DY_{AC_i}\} \{DC_{rDC} DC_{rAC_j}\} \{DC_{bDC} DC_{bAC_k}\} \}$ όπου τα i, j, k δηλώνουν τον αριθμό των AC συντελεστών και μπορούν να πάρουν τις τιμές 3, 6, 10, 15, 21, 28, 64.

Ο περιγραφέας ομοιογενούς υφής (Homogeneous Texture Descriptor) χαρακτηρίζει την υφή της εικόνας/περιοχής χρησιμοποιώντας τη μέση ενέργεια

και την απόκλιση ενέργειας από ένα σύνολο καναλιών συχνότητας. Το διάστημα επίπεδο συχνότητας διαμερίζεται ομοιόμορφα κατά τη γωνιακή κατεύθυνση και μη ομοιόμορφα κατά την ακτινική κατεύθυνση σε 30 κανάλια και υπολογίζεται η μέση ενέργεια και η απόκλιση της συχνότητας κάθε καναλιού. Το διάστημα του περιγραφέα ορίζεται από τη σχέση: $HTD = [f_{DC}, f_{SD}, e_1, e_2, \dots, e_N, d_1, d_2, \dots, d_N]$, όπου f_{DC} είναι η μέση και f_{SD} η τυπική απόκλιση των τιμών της εικόνας, N ο αριθμός των καναλιών στα οποία χωρίζει ο περιγραφέας το χώρο συχνοτήτων και e_i και d_i η μέση ενέργεια και η απόκλιση ενέργειας του καθενός από αυτά.

Ο περιγραφέας ιστογράμματος ακμών (Edge Histogram Descriptor) καταγράφει τη χωρική κατανομή των ακμών εντοπίζοντας κατευθυντικές (οριζόντιες, κάθετες, διαγώνιες με προσανατολισμό 45° , διαγώνιες με προσανατολισμό 135°) και μη ακμές. Αρχικά η εικόνα διαιρείται σε 4×4 υποεικόνες κι έπειτα η κάθε μία από αυτές διαιρείται σε προκαθορισμένο συνολικό αριθμό μη επικαλυπτόμενων τετραγωνικών μπλοκ. Κάθε ένα από τα μπλοκ ταξινομείται σε μία από τις πέντε κατηγορίες ακμών ή σαν μπλοκ χωρίς ακμές και έπειτα υπολογίζονται τα ιστογράμματα ακμών με 5 κορυφές, μία για κάθε είδος ακμής, για τις 16 υποεικόνες. Ο περιγραφέας ιστογράμματος ακμών ορίζεται από τη σχέση: $EHD = [e_1, e_2, \dots, e_80]$

4.4.1.2 Περιγραφείς γεωμετρικών ροπών

Οι ροπές (moments) θεωρούνται στατιστικά χαρακτηριστικά που μπορούν να εξαχθούν από μια εικόνα ή μια περιοχή αυτής και παρέχουν πληροφορίες σχετικές με τις συμμετρίες και τις ασυμμετρίες αυτής της περιοχής και έχουν χρησιμοποιηθεί ως χαρακτηριστικά γνωρίσματα σε διάφορες εφαρμογές [4], [7], [21], [144], [66]. Οι ροπές είναι αμετάβλητες σε γεωμετρικούς μετασχηματισμούς, οπότε μπορούν να αποτελέσουν έναν εποικοδομητικό τρόπο για την εξαγωγή χαρακτηριστικών γνωρισμάτων από τις διδιάστατες εικόνες, οι οποίες μπορούν να παρέχουν τις σφαιρικές πληροφορίες για την εικόνα. Οι ροπές μπορούν να χωριστούν σε δύο κατηγορίες, στις ροπές που είναι αμετάβλητες σε αλλαγές μεγέθους, μετατόψεις και περιστροφές και στις ροπές που είναι αμετάβλητες κάτω από οποιοσδήποτε αλλαγές, συμπεριλαμβανομένης της αντανάκλασης.

Αν θεωρήσουμε $f(x, y)$ τις τιμές των εικονοστοιχείων της εικόνας, οι γεωμετρικές ροπές αυτής της περιοχής είναι θετικές ακέραιες τιμές που δίνονται από τη σχέση:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (4.1)$$

όπου $p, q = 0, 1, 2, \dots$ είναι φυσικοί αριθμοί που δηλώνουν την τάξη της ροπής.

Για μια περιοχή $R(k, l) \subseteq f(x, y)$ η σχέση 4.1 γίνεται:

$$m_{pq} = \sum_x \sum_y x^p y^q R(x, y) \quad (4.2)$$

όπου τα αθροίσματα λαμβάνονται μόνο για τα εικονοστοιχεία που ανήκουν στην περιοχή R .

Ο ορισμός των ροπών ουσιαστικά ορίζει την προβολή της $f(x, y)$ σε ένα σύνολο ορθογωνίων βάσεων $\{x^p y^q\}$ (σύμφωνα με το θεώρημα προσέγγισης Weierstrass [5]). Σύμφωνα με τον ορισμό οι ροπές εξαρτώνται από τις συντεταγμένες της υπό εξέταση περιοχής της εικόνας, για αυτό και στερούνται από αμετάβλητες ιδιότητες (invariance property). Για να πετύχουμε σταθερότητα κατά τη μετατόπιση ($x' = x+a, y' = y+b$),

περιστροφή $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$ και αλλαγή κλίμακας ($x' = ax, y' = ay$), έχουν προταθεί οι κεντρικές ροπές, οι κανονικοποιημένες κεντρικές ροπές και οι ροπές Hu.

Οι γεωμετρικές ροπές μηδενικής και πρώτης τάξης, όπως έχει αποδειχθεί στην [4], αντιστοιχούν στις συντεταγμένες του κέντρου βάρους της περιοχής και δίνονται από τη σχέση: $\bar{x} = m_{10}/m_{00}$ και $\bar{y} = m_{01}/m_{00}$.

Με τη βοήθεια του κέντρου βάρους υπολογίζονται από τον ακόλουθο τύπο οι κεντρικές ροπές της εικόνας ή περιοχής αυτής:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (4.3)$$

Ουσιαστικά, οι κεντρικές ροπές ορίζονται με βάση τη σχετική απόσταση των εικονοστοιχείων της περιοχής από το κέντρο της. Οι κεντρικές ροπές παραμένουν αναλλοίωτες στους μετασχηματισμούς μετατόπισης, αφού κατά την παράλληλη μεταφορά του συστήματος συντεταγμένων οι αποστάσεις των σημείων από το κέντρο παραμένουν σταθερές.

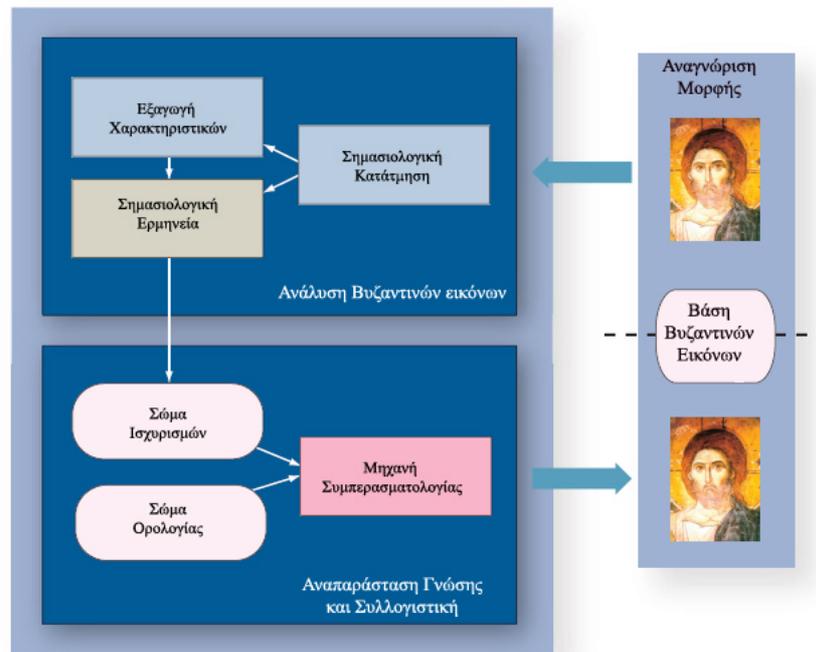
Η αμετάβλητη κλίμακα μπορεί να επιτευχθεί κανονικοποιώντας τις κεντρικές ροπές της κλιμακούμενης εικόνας ή περιοχής της. Οι κεντρικές κανονικοποιημένες ροπές παραμένουν αναλλοίωτες σε παράλληλη μεταφορά και σε αλλαγή κλίμακας και ορίζονται ως εξής :

$$n_{pq} = \mu_{pq} / \mu_{00}^\gamma \quad (4.4)$$

όπου $\gamma = (p + q/2) + 1, p + q > 2$ ο παράγοντας κανονικοποίησης. Ωστόσο οι κανονικοποιημένες κεντρικές ροπές είναι μεταβλητές στην περιστροφή.

4.4.1.2.1 Ροπές Hu Η αναζήτηση χαρακτηριστικών μιας περιοχής, τα οποία παραμένουν αναλλοίωτα σε μετασχηματισμούς περιστροφής οδήγησε στη δημιουργία των αμετάβλητων ροπών [66].

Το 1962, ο Hu παρουσίασε το μαθηματικό ορισμό των διδιάστατων ροπών, βασιζόμενος σε αμετάβλητες αλγεβρικές σχέσεις και έδειξε πως αυτές μπορούν να χρησιμοποιηθούν για την αναγνώριση σχημάτων σε εικόνες. Η μέθοδος που πρότεινε ο Hu, βασίζεται στο μη γραμμικό συνδυασμό της 2ης και 3ης τάξης κανονικοποιημένων κεντρικών ροπών, παρέχοντας ένα σύνολο συναρτήσεων που είναι αμετάβλητες σε μετασχηματισμούς μετατόπισης, αλλαγής κλίμακας και περιστροφής, αλλά και σε κατοπτρικούς μετασχηματισμούς. Έτσι, όρισε ένα σύνολο από επτά απόλυτες κανονικοποιημένες ροπές, από τις οποίες οι phi_1, \dots, phi_6 παραμένουν αμετάβλητες σε περιστροφή, μεταφορά, αλλαγή κλίμακας και κατοπτρισμό, ενώ η phi_7 είναι χρήσιμη στο διαχωρισμό κατοπτρικών εικόνων, αφού είναι αμετάβλητη σε μετασχηματισμούς διαστρέβλωσης (skew orthogonal invariant) και αλλάζει πρόσημο σε κατοπτρικούς μετασχηματισμούς. Για την αποφυγή προβλημάτων ακριβείας στους υπολογισμούς, συνήθως χρησιμοποιούμε τους αλγορίθμους των απολύτων τιμών.



Σχήμα 4.3: Αρχιτεκτονική του συστήματος τεκμηρίωσης κατηγοριοποίησης Βυζαντινών εικόνων. Η μονάδα Ανάλυση Βυζαντινών εικόνων εξάγει πληροφορίες για την εικόνα και η μονάδα Αναπαράσταση Γνώσης και Συλλογιστική χρησιμοποιεί αυτές τις πληροφορίες για την εξαγωγή της υπονοούμενης γνώσης, κατηγοριοποιώντας την εικόνα

$$\begin{aligned}
 \phi_1 &= n_{20} + n_{02} \\
 \phi_2 &= (n_{20} - n_{02})^2 + 4n_{11}^2 \\
 \phi_3 &= (n_{30} - 3n_{12})^2 + (n_{03} - 3n_{21})^2 \\
 \phi_4 &= (n_{30} - n_{12})^2 + (n_{03} + n_{21})^2 \\
 \phi_5 &= (3n_{30} - 3n_{12})^2 + (n_{03} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] \\
 &\quad + (3n_{21} - n_{03}) + (n_{21} + n_{03})[3(n_{30} - n_{12})^2 - (n_{21} + n_{03})^2] \\
 \phi_6 &= (n_{20} - n_{02})[(n_{30} - n_{12})^2 + (n_{21} + n_{03})^2] \\
 &\quad + 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03}) \\
 \phi_7 &= (3n_{21} - n_{03})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] \\
 &\quad + (3n_{12} - n_{03})(n_{21} + n_{03})[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2]
 \end{aligned} \tag{4.5}$$

4.4.2 Σύστημα αυτόματης κατηγοριοποίησης και τεκμηρίωσης Βυζαντινών εικόνων με βάση τη γνώση

Στο Σχήμα 4.3 παρουσιάζεται η γενική αρχιτεκτονική του συστήματος κατηγοριοποίησης των Βυζαντινών εικόνων που προτάθηκε στο [135]. Αποτελείται από τη μονάδα ανάλυσης Βυζαντινών εικόνων που έχει σκοπό την επεξεργασία και ανάλυση εικόνων για την ανίχνευση ενός συνόλου πρωτόγονων εννοιών και ιδιοτήτων που αυτή περιέχει και τη μονάδα Αναπαράστασης Γνώσης και Συλλογιστικής που έχει ως σκοπό την κατηγοριοποίηση της εικόνας.

Η πρώτη μονάδα, η ανάλυση Βυζαντινών εικόνων, περιλαμβάνει τη σημασιολογική κατάτμηση, την εξαγωγή χαρακτηριστικών και τη σημασιολογική ερμηνεία των

εξαγόμενων τμημάτων 4.3.

Αναλυτικότερα, κατά τη σημασιολογική κατάτμηση αρχικά πραγματοποιείται η ανίχνευση του προσώπου, όπως αυτή περιγράφεται στην Ενότητα 4.4.2.1. Το αποτέλεσμα της μεθόδου απεικονίζεται με πράσινο πλαίσιο στην πρώτη εικόνα του πλαισίου Εντοπισμός προσώπου του Σχήματος 4.4. Έπειτα σύμφωνα με την τεχνική που περιγράφεται στην Ενότητα 4.4.2.2, εντοπίζονται τα μάτια της μορφής όπως αυτά απεικονίζονται με κόκκινους κύκλους στη δεύτερη εικόνα του πλαισίου Εντοπισμός προσώπου του Σχήματος 4.4.

Σύμφωνα με τον Διονύσιο εκ Φουρνά, η απόσταση μεταξύ των ματιών είναι ίση με το ύψος της μύτης. Κατά συνέπεια εντοπίζοντας τις ακριβείς θέσεις των ματιών και μετρώντας την μεταξύ τους απόσταση, μπορούμε να έχουμε μια αρχική προσέγγιση για το ύψος της μύτης. Το ακριβές ύψος της μύτης υπολογίζεται με τη χρήση της τεχνικής που παρουσιάζεται στην Ενότητα 4.4.2.2.

Έπειτα εντοπίζεται ο άξονας συμμετρίας (Ενότητα 4.4.2.2), ο οποίος είναι καθοριστικός επειδή μας φανερώνει την πλευρά του προσώπου που έχει τα πιο έντονα χαρακτηριστικά από τα οποία θα εξάγουμε τις πληροφορίες.

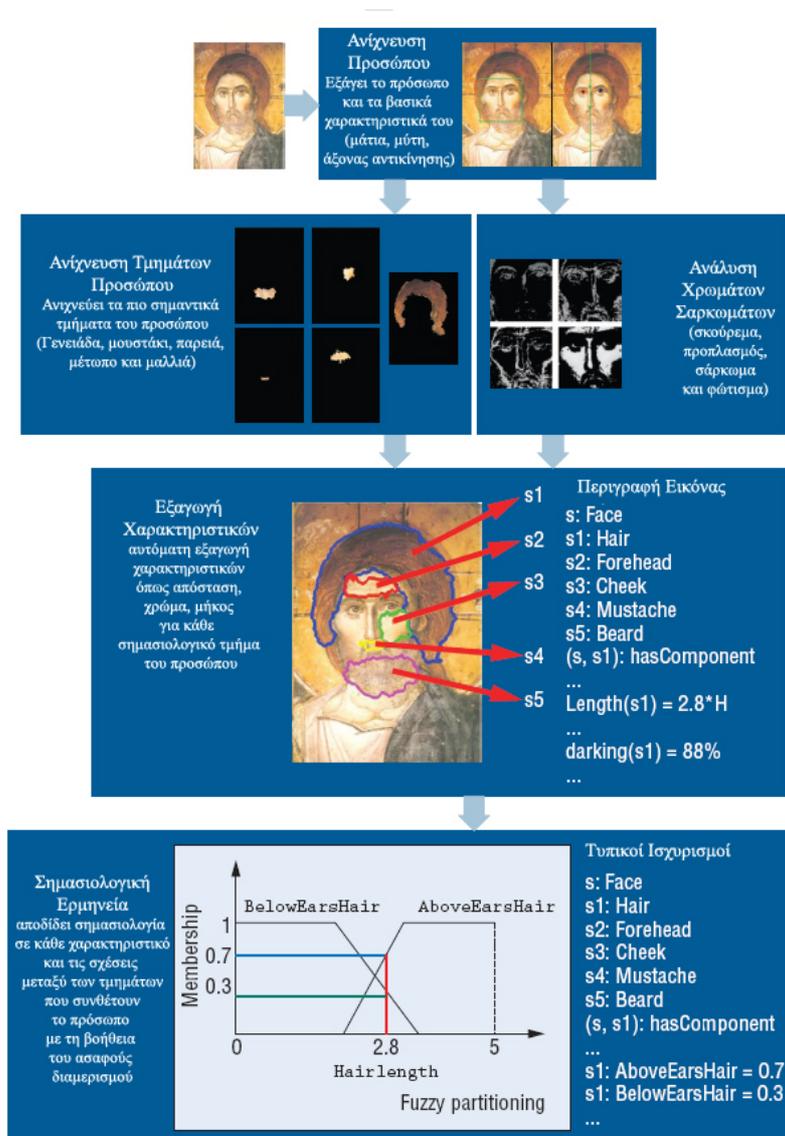
Έχοντας υπολογίσει το ύψος καθώς και τη θέση της μύτης και τον άξονα συμμετρίας, γνωρίζουμε χονδρικά όλα τα τμήματα στα οποία χωρίζεται η μορφή 4.2 καθώς και την πλατύτερη πλευρά της. Επόμενο στάδιο λοιπόν είναι η ανίχνευση και η εξαγωγή των σημασιολογικών τμημάτων της μορφής. Για την υλοποίηση αυτού του βήματος έχει υιοθετηθεί ο αλγόριθμος που προτείνεται στην [30] και το αποτέλεσμα της εφαρμογής του φαίνεται στις εικόνες του πλαισίου ανίχνευσης τμημάτων προσώπου (Σχήμα 4.4).

Στη συνέχεια γίνεται ο προσδιορισμός των χρωμάτων που ο αγιογράφος έχει χρησιμοποιήσει στην περιοχή του προσώπου. Γι'αυτό το σκοπό έχει χρησιμοποιηθεί ο αλγόριθμος κατάτμησης που προτάθηκε στη [101] εξάγοντας τα τέσσερα βασικά χρώματα, σκούρεμα, προπλασμό, σάρκωμα και φώτισμα της μορφής (Σχήμα 4.4).

Η εξαγωγή των ιδιοτήτων του κάθε σημασιολογικού τμήματος μας οδηγεί στην κατηγοριοποίηση ανά ιδιότητα (π.χ. μήκος, χρώμα, μορφή) όπως ορίστηκε στους πίνακες 4.1, 4.2, 4.3, 4.4, 4.5. Για την εξαγωγή των χαρακτηριστικών του κάθε τμήματος χρησιμοποιήθηκαν οι τεχνικές που περιγράφονται στην Ενότητα 4.4.2.3.

Έχοντας εξάγει τα χαρακτηριστικά από κάθε σημασιολογικό τμήμα μπορούμε να προχωρήσουμε στην καταγραφή αυτών με τη χρήση ασαφών ισχυρισμών όπως φαίνεται στη δεξιά πλευρά του πλαισίου Εξαγωγή Χαρακτηριστικών. Εντούτοις, οι ιδιότητες κάθε τμήματος είναι ασαφείς. Παραδείγματος χάριν, δεν μπορούμε με σαφήνεια να καθορίσουμε βάσει του μήκους, εάν η μορφή έχει την ιδιότητα μακρυμάλλης. Αυτή η δυσκολία οφείλεται στην ασάφεια του ορισμού του μήκους και στην ανακρίβεια που εισάγεται από την εξαγωγή χαρακτηριστικών γνωρισμάτων, όπως φαίνεται και στην εικόνα με τις κατατμήσεις των τμημάτων (εικόνα πλαισίου Εξαγωγή Χαρακτηριστικών). Αυτό μας οδήγησε στην υιοθέτηση της ασαφούς λογικής [151] και στην έκφραση των ισχυρισμών που εξάγονται για κάθε τμήμα με χρήση αυτής της λογικής. Αυτή η διεργασία είναι πολύ σημαντική και τα αποτελέσματα της, δηλαδή οι ασαφείς ισχυρισμοί για τα τμήματα, βοηθούν ουσιαστικά την επόμενη μονάδα του συστήματος αφού αποτελούν το σώμα ασαφών ρόλων.

Η λειτουργία της μονάδας αναπαράστασης γνώσης και συλλογιστική (Σχήμα 4.5) περιγράφεται αναλυτικά στο Κεφάλαιο 5. Η μονάδα αυτή αξιοποιεί τις εξαγόμενες πληροφορίες από τη μονάδα ανάλυσης εικόνας και συνδυάζοντάς αυτές με την καταγεγραμμένη γνώση του πεδίου είναι ικανή να παράγει νέα γνώση.



Σχήμα 4.4: Ανάλυση Βυζαντινής εικόνας. Κατ' αρχάς, ο αλγόριθμος ανιχνεύει την περιοχή του προσώπου καθώς και τα μάτια, τη μύτη του και τον άξονα αντικίνησης. Κατόπιν, εξάγει τις σημασιολογικές περιοχές των μαλλιών, του μετώπου, των παρειών, της γενειάδας και του μουστακίου καθώς και τα βασικά χρώματα που χρησιμοποιεί ο αγιογράφος για τα σαρκώματα του προσώπου. Η περαιτέρω ανάλυση των εξαγόμενων τμημάτων του προσώπου παρέχει πληροφορίες για τα χαρακτηριστικά του προσώπου. Τέλος, ο αλγόριθμος παράγει μια σημασιολογική ερμηνεία για κάθε ένα από αυτά τα χαρακτηριστικά γνωρίσματα.



Σχήμα 4.5: Ένα παράδειγμα συλλογιστικής. Οι εξαγόμενες πληροφορίες από τη μονάδα ανάλυσης εικόνας αποτελούν το ασαφές σώμα ρόλων (ABox) της βάσης γνώσεων. Το ασαφές σώμα ορολογίας (TBox) καθορίζεται βάσει του εγχειριδίου του Διονυσίου εκ Φουρνά. Αυτά τα σώματα αποτελούν την είσοδο για τη μηχανή συλλογισμού, η οποία εξάγει πληροφορίες για την εικόνα.

Στη συνέχεια του κεφαλαίου θα περιγραφούν αναλυτικά οι τεχνικές που προτάθηκαν για τη μονάδα ανάλυσης Βυζαντινής εικόνας ενώ στο κεφάλαιο 5 θα παρουσιαστούν αναλυτικά οι τεχνικές που χρησιμοποιήθηκαν στη μονάδα αναπαράστασης γνώσης και συλλογιστική.

4.4.2.1 Ανίχνευση προσώπου σε πολιτιστικό υλικό

Η ανάγκη για δημιουργία κατάλληλου αλγορίθμου ο οποίος θα αναλύει πολιτιστικό υλικό μιας καλλιτεχνικής περιόδου με έμφαση στον εντοπισμό προσώπων, μας οδήγησε στην ερευνητική περιοχή της επεξεργασίας εικόνων με πρόσωπα. Στην αντίστοιχη βιβλιογραφία υπάρχει πλήθος μελετών [138], [71], [70], πολλές από τις οποίες έχουν καθιερωθεί έπειτα από δοκιμές.

Η έρευνα για τον εντοπισμό του προσώπου και των χαρακτηριστικών του σε μια εικόνα οδήγησε στην ανάπτυξη ευρείας γκάμας εφαρμογών με τεράστια εμπορική ζήτηση. Τα σύγχρονα συστήματα αναγνώρισης έχουν φθάσει σε μεγάλο επίπεδο ωριμότητας προσπαθώντας να προσομοιάσουν την ικανότητα του ανθρώπινου συστήματος αντίληψης. Τα συστήματα αυτόματης αναγνώρισης προσώπων έχουν ήδη ενσωματωθεί σε διάφορες συσκευές μέσω εφαρμογών με σκοπό τον εντοπισμό και την ομαδοποίηση προσώπων, την πιστοποίηση της ταυτότητας του χρήστη του από τα χαρακτηριστικά του προσώπου του.

Ο εντοπισμός προσώπου βρίσκει εφαρμογή σε πολλά συστήματα που έχουν ως στόχο την εξαγωγή πληροφοριών από τα πρόσωπα. Ο στόχος των μεθόδων που έχουν προταθεί στη βιβλιογραφία για ανίχνευση προσώπου είναι η εύρεση της παρουσίας οποιουδήποτε προσώπου σε μία εικόνα και ο προσδιορισμός της θέσης του και του

περιεχομένου του μέσα σε αυτήν την εικόνα. Οι τεχνικές ανίχνευσης προσώπου που έχουν αναπτυχθεί ασχολήθηκαν κυρίως με τα όρθια μετωπικά πρόσωπα.

Ωστόσο έχουν προταθεί τεχνικές ανίχνευσης προσώπου οι οποίες λαμβάνουν υπόψη τους διάφορες παραμέτρους όπως α) τη στάση του προσώπου (μετωπική, πλάγια 45°, προφίλ, ανεστραμμένη), β) την ευκρίνεια του προσώπου (π.χ εικόνα με πολλά πρόσωπα που κάποια από αυτά επικαλύπτουν μερικώς ή ολικώς τα υπόλοιπα) αλλά και των χαρακτηριστικών του (π.χ μερική ή πλήρης απόκρυψη του ματιού ή του χεριού), γ) την παρουσία ή απουσία δομικών συστατικών (γένια, μουστάκι, γυαλιά) καθώς και τα ιδιαίτερα χαρακτηριστικά τους (π.χ.σχήμα, χρώμα, και μέγεθος), δ) την έκφραση του προσώπου και ε) τον προσανατολισμό της εικόνας καθώς και τις συνθήκες φωτισμού και τα χαρακτηριστικά της κάμερας.

Παρόλες αυτές τις δυσκολίες, τεράστια πρόοδος έχει σημειωθεί την τελευταία δεκαετία και πολλά νέα συστήματα έχουν παρουσιάσει εντυπωσιακή απόδοση σε πραγματικό χρόνο. Η πρόσφατη πρόοδος αυτών των αλγορίθμων έχει επίσης συνεισφέρει και στην ανίχνευση άλλων αντικειμένων (π.χ. άνθρωποι/πεζοί, αυτοκίνητα).

Η λειτουργία των περισσοτέρων συστημάτων ανίχνευσης προσώπων βασίζεται στην εξαγωγή ορισμένων ιδιοτήτων (π.χ., τοπικά χαρακτηριστικά γνωρίσματα ή ολιστικά πρότυπα που βασίζονται στη φωτεινότητα) από ένα σύνολο εικόνων που χρησιμοποιούνται για εκπαίδευση αυτών των συστημάτων. Αρχικά οι εικόνες υποβάλλονται σε επεξεργασία (εξισορρόπηση ιστογραμμάτων [127],[112] ή υπολογισμός διακύμανσης από τη μηδενική μέση τιμή [138]) με σκοπό τη μείωση των επιδράσεων των αλλαγών φωτισμού. Βασιζόμενα στα εξαγόμενα χαρακτηριστικά, τα συστήματα αυτά εξετάζουν λεπτομερώς ολόκληρη την εικόνα σε κάθε πιθανή θέση και κλίμακα προκειμένου να ανιχνεύουν παρόμοια χαρακτηριστικά και τελικά να εντοπίσουν τα πρόσωπα.

Στη βιβλιογραφία έχουν προταθεί πολλές μέθοδοι για ανίχνευση προσώπου [149], βασισμένες σε: α) εικονοστοιχεία ([127], [112], [150]), β) τμήματα της εικόνας ([94], [117], [61],[98]), γ) τοπικά χαρακτηριστικά γνωρίσματα των ακμών της εικόνας ([8], [49]), δ) κυματίδια Haar ([102], [117]) και ε) χαρακτηριστικά γνωρίσματα τύπου Haar ([42], [138]). Οι μέθοδοι που χρησιμοποιούν χαρακτηριστικά γνωρίσματα τύπου Haar είναι σε θέση να ανιχνεύσουν πρόσωπα ακόμη και κάτω από συνθήκες επικάλυψης των προσώπων.

Για την ανίχνευση προσώπων σε διαφορετική κλίμακα, η διαδικασία ανίχνευσης επαναλαμβάνεται συνήθως χρησιμοποιώντας μία πυραμίδα εικόνων των οποίων η ανάλυση μειώνεται κατά ένα συγκεκριμένο παράγοντα ([112], [127]) σε σύγκριση με την αρχική εικόνα. Με τη χρήση άλλων οπτικών χαρακτηριστικών, όπως το χρώμα και η κίνηση, ως βήματα προεπεξεργασίας είναι δυνατή η μείωση του διαστήματος αναζήτησης ([150]) και κατά συνέπεια η επιτάχυνση της όλης διαδικασίας ανίχνευσης.

Η ανίχνευση προσώπου ουσιαστικά μπορεί να τυποποιηθεί κυρίως ως πρόβλημα αναγνώρισης προτύπων. Στη βιβλιογραφία έχουν προταθεί πολυάριθμοι αλγόριθμοι που είτε κάνουν εκμάθηση των γενικών προτύπων τους (π.χ., eigenfaces και στατιστική κατανομή) είτε χρησιμοποιούν ταξινομητές (π.χ., τα νευρωνικά δίκτυα(NN), Fisher linear discriminant, sparse network of Winnows, δέντρα απόφασης, ταξινομητές Bayes, διανυσματικές μηχανές υποστήριξης (SVM) και Adaboost) με σκοπό την ανίχνευση προσώπων.

4.4.2.1.1 Περιγραφή μεθόδου βασισμένη στον εντοπισμό δέρματος
 Ανάμεσα στις πιο δημοφιλείς μεθόδους εντοπισμού προσώπου είναι αυτές που λαμβάνουν υπόψη τους το χρώμα του δέρματος. Οι μέθοδοι αυτές χαρακτηρίζονται από την απλότητα και την αποδοτικότητα τους ακόμη και υπό δυσμενείς συνθήκες (περιστροφές, μερική κάλυψη προσώπου, κ.τ.λ.). Σύμφωνα με κλασσικές μελέτες [142],[58], οι αποχρώσεις του ανθρωπίνου δέρματος εκτείνονται σε μια μικρή περιοχή του χρωματικού επιπέδου $C_b - C_r$ του μοντέλου YC_bC_r . Η εργασία [146] προτείνει έναν αλγόριθμο ανίχνευσης δέρματος με γρήγορες επιδόσεις, όπου η κατανομή του χρώματος του δέρματος προσεγγίζεται από μία διδιάστατη κανονική (Gaussian) συνάρτηση πυκνότητας πιθανότητας.

Αρχικά, σύμφωνα με αυτή την εργασία, οι δύο χρωματικές συνιστώσες $C_b - C_r$ μίας έγχρωμης εικόνας χρησιμοποιούνται για τον εντοπισμό του ανθρωπίνου προσώπου, θεωρώντας δεδομένο το μικρό εύρος κατανομής των χρωματικών τιμών που αντιστοιχούν στο ανθρώπινο δέρμα [142]. Τα τμήματα της εικόνας με χρωματικές τιμές σε αυτή τη μικρή περιοχή του χρωματικού χώρου, μπορούν να θεωρηθούν ως υποψήφια τμήματα προσώπου, το σύνολο των οποίων θεωρούμε ότι απαρτίζει μια κλάση προσώπου Ω_f .

Στο προτεινόμενο σύστημα, το ιστόγραμμα των χρωματικών τιμών που αντιστοιχούν στην κλάση προσώπου Ω_f αρχικά μοντελοποιείται με χρήση μίας Γκαουσιανής συνάρτησης πυκνότητας πιθανότητας (Gaussian pdf). Οπότε ισχύει:

$$P(x | \Omega_f) = \frac{\exp(-\frac{1}{2} \cdot (x - \mu_f)^T \cdot \Sigma_f^{-1} (x - \mu_f))}{2\pi \cdot |\Sigma|^{1/2}} \quad (4.6)$$

όπου $x = [u \ v]^T$ είναι ένα διάνυσμα μεγέθους 2×1 που περιέχει τις μέσες χρωματικές συνιστώσες u και v του υπό εξέταση τμήματος, μ_f είναι το μέσο διάνυσμα μιας περιοχής προσώπου διαστάσεων 2×1 και Σ είναι ο 2×2 πίνακας μεταβλητότητας (variance matrix) της συνάρτησης πυκνότητας πιθανότητας που έχει τη μορφή:

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{u,v} \\ \sigma_{u,v} & \sigma_v^2 \end{bmatrix} \quad (4.7)$$

όπου σ_u^2 είναι η μεταβλητότητα (variance) της χρωματικής συνιστώσας u , σ_v^2 είναι η μεταβλητότητα (variance) της χρωματικής συνιστώσας v και $\sigma_{u,v}$ αντιστοιχεί στη συμμεταβλητότητα (covariance) των u και v . Οι παράμετροι μ_f και Σ υπολογίζονται με βάση ένα σύνολο από πολλές εικόνες που περιέχουν περιοχές ανθρωπίνου προσώπου με χρήση του αλγόριθμου μέγιστης πιθανοφάνειας (maximum likelihood algorithm).

Μετά τη μοντελοποίηση της κατανομής των χρωματικών συνιστωσών περιοχών προσώπου (αφού δηλαδή εκτιμηθούν οι παράμετροι μ_f και Σ της εξίσωσης 4.7, μπορούμε να υπολογίσουμε για κάθε τμήμα B_i μίας εικόνας, την πιθανότητα το τμήμα αυτό να ανήκει σε δερματική περιοχή προσώπου. Πιο συγκεκριμένα για κάθε τμήμα B_i εξάγονται οι δύο χρωματικές του συνιστώσες δηλαδή τα $\mathbf{u}(B_i)$ και $\mathbf{v}(B_i)$ και έτσι σχηματίζεται το διάνυσμα $\mathbf{x}(B_i) = [\mathbf{u}(B_i) \ \mathbf{v}(B_i)]^T$. Έπειτα χρησιμοποιείται η εξίσωση 4.7 προκειμένου να υπολογισθεί η πιθανότητα το εξεταζόμενο τμήμα να ανήκει στην κλάση προσώπου.

Η εξίσωση 4.7 υποδεικνύει ότι ένα τμήμα της εικόνας B_i ανήκει σε περιοχή προσώπου, εφόσον η αντίστοιχη πιθανότητα των χρωματικών του τιμών $P(\mathbf{x}(B_i)|\Omega_f)$ είναι υψηλή. Αντίθετα τμήματα με χαμηλή τιμή πιθανότητας $P(\mathbf{x}(B_i)|\Omega_f)$ ταξινομούνται ως τμήματα που δεν ανήκουν σε περιοχή προσώπου.

Με σκοπό τη λήψη πιο αξιόπιστων αποτελεσμάτων, χρησιμοποιούμε μόνο μία υποπεριοχή της Γκαουσιανής συνάρτησης πυκνότητας πιθανότητας, σύμβαση απαραίτητη, δεδομένου ότι όλα τα τμήματα μίας εικόνας παρουσιάζουν μία πιθανότητα να ανήκουν σε περιοχή προσώπου (ακόμα και αν η πιθανότητα αυτή είναι πολύ μικρή).

Πιο συγκεκριμένα, επιλέγεται ένα διάστημα εμπιστοσύνης το οποίο αντιστοιχεί στο 80% της περιοχής που καλύπτει το Γκαουσιανό μοντέλο, έτσι ώστε μόνο τα τμήματα που θα παρουσιάσουν πιθανότητα μέσα σε αυτή την περιοχή να θεωρούνται ως υποψήφια τμήματα περιοχής προσώπου. Τα τμήματα που ανήκουν στο υπόλοιπο 20% θεωρούνται ως τμήματα που δεν ανήκουν σε περιοχή προσώπου.

Έτσι για κάθε εικόνα, για την ταξινόμηση των τμημάτων της, μεγέθους $N_1 \times N_2$, παράγεται μία δυαδική μάσκα M μεγέθους $N_1/8 \times N_2/8$ εικονοστοιχείων (καθώς η ταξινόμηση γίνεται σε ανάλυση επιπέδου τμημάτων). Μία μονάδα εικονοστοιχείου (τμήμα μεγέθους 8×8), με τιμή ίση με τη μονάδα, υποδεικνύει ένα υποψήφιο τμήμα περιοχής προσώπου, ενώ τιμή ίση με το μηδέν υποδεικνύει τμήματα τα οποία δεν ανήκουν σε περιοχή προσώπου.

Εντούτοις, καθώς η προηγούμενη διαδικασία λαμβάνει υπόψη μόνο την πληροφορία χρώματος, στην τελική δυαδική μάσκα M είναι πιθανό να έχουν ταξινομηθεί τμήματα ως υποψήφιας περιοχής προσώπου, ενώ στην πραγματικότητα είναι περιοχές που παρουσιάζουν παρόμοιες χρωματικές ιδιότητες με τις ζητούμενες, όπως για παράδειγμα τα ανθρώπινα χέρια.

Για να αντιμετωπιστεί αυτός ο παράγοντας, λαμβάνεται υπόψη και η πληροφορία σχήματος που έχει το ανθρώπινο πρόσωπο. Πιο συγκεκριμένα, το σχήμα του ανθρώπινου προσώπου είναι μοναδικό και παρουσιάζει συνέπεια ως προς τις παραμέτρους του.

Στο παρελθόν έχουν προταθεί συστήματα τα οποία προσεγγίζουν το σχήμα του προσώπου με έλλειψη, με ενωμένα τόξα ή με χρήση ορθογωνίου παραλληλογράμμου [142]. Στο προτεινόμενο σύστημα υιοθετείται η μέθοδος που προτάθηκε στην εργασία [142], όπου ορθογώνια παραλληλόγραμμα με συγκεκριμένους λόγους πλευρών χρησιμοποιούνται για να προσεγγίσουν το σχήμα της περιοχής προσώπου. Αυτό οφείλεται στο γεγονός ότι επιδιώκουμε την εύρεση των συνδεδεμένων τμημάτων της μάσκας M που αντιστοιχούν σε πραγματικό ανθρώπινο πρόσωπο και όχι την ακριβή μοντελοποίηση του σχήματος της περιοχής του ανθρώπινου προσώπου.

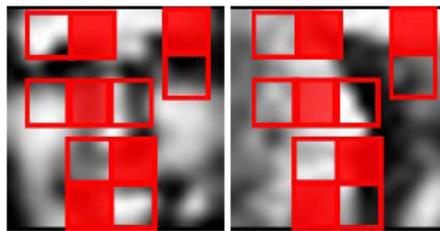
Πιο συγκεκριμένα στο προτεινόμενο σύστημα ο λόγος πλευρών για μία περιοχή προσώπου ορίζεται ως εξής:

$$R = \frac{H_f}{W_f} \quad (4.8)$$

όπου H_f είναι το ύψος του κεφαλιού, ενώ το W_f αντιστοιχεί στο πλάτος του προσώπου. Μετά από αρκετά πειράματα ο λόγος R βρέθηκε να ανήκει στο κλειστό διάστημα $[1.4, \dots, 1.6]$. Επιπρόσθετα η γωνία περιστροφής του προσώπου (face tilt) λαμβάνει τιμές που ανήκουν στο διάστημα $[-30^\circ, \dots, +30^\circ]$.

Συγχρόνως, τμήματα μεγέθους μικρότερου από 8×8 εικονοστοιχεία αγνοούνται καθώς αντιστοιχούν σε θόρυβο. Τελικά κατασκευάζεται μία δυαδική μάσκα M_f , μεγέθους $N_1/8 \times N_2/8$ και τα εικονοστοιχεία της, με τιμή ίση με τη μονάδα, ανήκουν στην περιοχή προσώπου, ενώ τα εικονοστοιχεία με τιμή ίση με το μηδέν δεν ανήκουν.

4.4.2.1.2 Περιγραφή μεθόδου βασισμένη στην Adaboost τεχνική Ένας από τους πιο διαδεδομένους αλγόριθμους ανίχνευσης προσώπου είναι αυτός που προτάθηκε από τους Viola και Jones [138] και βασίζεται σε AdaBoost τεχνικές. Ο



Σχήμα 4.6: Τέσσερις τύποι Haar-ομοειδών χαρακτηριστικών γνωρισμάτων. Αυτά τα χαρακτηριστικά γνωρίσματα εμφανίζονται σε διαφορετικές θέσεις και κλίμακες. Τα Haar-ομοειδή χαρακτηριστικά γνωρίσματα υπολογίζονται ως διαφορά των σκουρόχρωμων και ανοιχτόχρωμων περιοχών. Μπορούν να θεωρηθούν ως χαρακτηριστικά γνωρίσματα που συλλέγουν τις τοπικές πληροφορίες ακμών με διαφορετικό προσανατολισμό και κλίμακα. Το σύνολο των Haar-ομοειδών χαρακτηριστικών γνωρισμάτων είναι μεγάλο, και μόνο μια μικρή ποσότητα αυτών μαθαίνεται από τα θετικά και αρνητικά παραδείγματα για την ανίχνευση προσώπου

αλγόριθμος αυτός εντοπίζει πρόσωπα με μεγάλη ευρωστία και σε πραγματικό χρόνο (δηλ. περισσότερα από 15 καρέ ανά δευτερόλεπτο σε εικόνες με ανάλυση 320x240 σε ένα κοινό υπολογιστή) ακόμη και υπό μερική κάλυψη του προσώπου. Οι Viola και Jones πρότειναν τη χρήση των χαρακτηριστικών γνωρισμάτων τύπου Haar για τον αποτελεσματικό υπολογισμό του ολοκληρώματος της εικόνα (integral image)[138].

Στο Σχήμα 4.6 παρουσιάζονται οι τέσσερις τύποι χαρακτηριστικών γνωρισμάτων Haar που χρησιμοποιούνται για να κωδικοποιήσουν τις οριζόντιες, κάθετες και διαγώνιες πληροφορίες φωτεινότητας των εικόνων που περιέχουν πρόσωπα σε διαφορετικές θέσεις και κλίμακες. Λαμβάνοντας υπόψη μια εικόνα-δείγμα 24x24 εικονοστοιχείων, το πλήρες σύνολο των παραμετροποιημένων χαρακτηριστικών γνωρισμάτων τύπου Haar (στις διαφορετικές θέσεις και κλίμακες) αυτού του δείγματος είναι πολύ μεγάλο (περίπου 160.000). Αντίθετα με τους περισσότερους από τους προγενέστερους αλγόριθμους που χρησιμοποιούν έναν ενιαίο ισχυρό ταξινομητή (π.χ., νευρωνικά δίκτυα και διανυσματικές μηχανές υποστήριξης), ο συγκεκριμένος αλγόριθμος χρησιμοποιεί ένα σύνολο ασθενέστερων ταξινομητών όπου καθένας κατασκευάζεται από την κατωφλίωση ενός χαρακτηριστικού γνωρίσματος τύπου Haar.

Χρησιμοποιώντας τον αλγόριθμο AdaBoost [50] επιλέγονται και βαθμονομούνται οι ασθενέστεροι ταξινομητές. Η ταξινόμηση αυτών των ασθενέστερων ταξινομητών γίνεται με τη χρήση διαδοχικά συνδεδεμένων ταξινομητών, χρησιμοποιώντας ένα σύνολο κριτηρίων βελτιστοποίησης. Σε κάθε επίπεδο των διαδοχικά συνδεδεμένων ταξινομητών, ένα σύνολο διάφορων ασθενέστερων ταξινομητών εκπαιδεύεται χρησιμοποιώντας τον αλγόριθμο AdaBoost. Το πλεονέκτημα της χρήσης των διαδοχικά συνδεδεμένων ταξινομητών είναι ότι οι απλοί ταξινομητές στο πρώτο στάδιο μπορούν να αφαιρέσουν τα περισσότερα αρνητικά δεδομένα (όχι πρόσωπα) αποτελεσματικά και οι ισχυρότεροι ταξινομητές στο προχωρημένο στάδιο να εξετάζουν τις περιπτώσεις που μοιάζουν με τα πρόσωπα. Ο τελικός ανιχνευτής που αποτελείται από 38 διαδοχικά επίπεδα συνδεδεμένων ταξινομητών με 6.060 Haar χαρακτηριστικά γνωρίσματα, κατέδειξε εντυπωσιακή απόδοση σε πραγματικό χρόνο με αρκετά υψηλά ποσοστά ανίχνευσης και με χαμηλά ψευδώς θετικά (false positive) ποσοστά.

Παρά την άριστη απόδοση χρόνου εκτέλεσης του αλγορίθμου των Viola και Jones, απαιτείται μεγάλος χρόνος εκπαίδευσης για ένα σύστημα. Επιπλέον, ο αλγόριθμος

αυτός είναι ένα παράδειγμα εκφυλισμένου δέντρου απόφασης με ένα μη ισορροπημένο σύνολο δεδομένων (δηλ., ένα μικρό σύνολο θετικών δειγμάτων και ένα τεράστιο σύνολο αρνητικών δειγμάτων). Για να χειριστούν αυτήν την ασυμμετρία μεταξύ των θετικών και αρνητικών συνόλων δεδομένων προτάθηκε από τους ίδιους ερευνητές ο ασύμμετρος αλγόριθμος AdaBoost [137], ο οποίος επιλέγει τα θετικά δείγματα.

Στην εργασία [138], ο αλγόριθμος AdaBoost χρησιμοποιείται για να επιλέξει ένα συγκεκριμένο αριθμό ασθενέστερων ταξινομητών με τα χαμηλότερα ποσοστά λάθους για κάθε διαδοχικά συνδεδεμένο ταξινομητή και η διαδικασία επαναλαμβάνεται έως ότου ένα σύνολο κριτηρίων βελτιστοποίησης ικανοποιηθεί (ο αριθμός των επιπέδων, ο αριθμός των χαρακτηριστικών γνωρισμάτων κάθε επιπέδου και η ανίχνευση/ψευδώς θετικά ποσοστά). Δεδομένου ότι κάθε αδύνατος ταξινομητής αποτελείται από ένα ενιαίο χαρακτηριστικό γνώρισμα τύπου Haar, η διαδικασία μέσα σε κάθε επίπεδο μπορεί να θεωρηθεί ως πρόβλημα επιλογής χαρακτηριστικών γνωρισμάτων.

Αντί της επανάληψης της διαδικασίας επιλογής χαρακτηριστικών γνωρισμάτων σε κάθε επίπεδο, στην εργασία [146] παρουσιάστηκε ένας άπληστος αλγόριθμος για τον καθορισμό ενός συνόλου χαρακτηριστικών γνωρισμάτων για όλα τα επίπεδα πρώτα πριν την εκπαίδευση των διαδοχικά συνδεδεμένων ταξινομητών. Χρησιμοποιώντας τον άπληστο αλγόριθμο επιλογής χαρακτηριστικών γνωρισμάτων για την προεργασία των δεδομένων, ο χρόνος εκπαίδευσης των διαδοχικά συνδεδεμένων ταξινομητών με AdaBoost, όπως αναφέρεται στη [146] μειώθηκε από 50 έως 100 φορές. Επιπλέον, ο αλγόριθμος που προτάθηκε στη [146] εκμεταλλεύτηκε επίσης την ασυμμετρία μεταξύ των θετικών και αρνητικών δεδομένων χρησιμοποιώντας έναν γραμμικό ταξινομητή, υποθέτοντας ότι αυτά τα δεδομένα μπορούν να μοντελοποιηθούν με τη χρήση γκαουσιανών κατανομών [146].

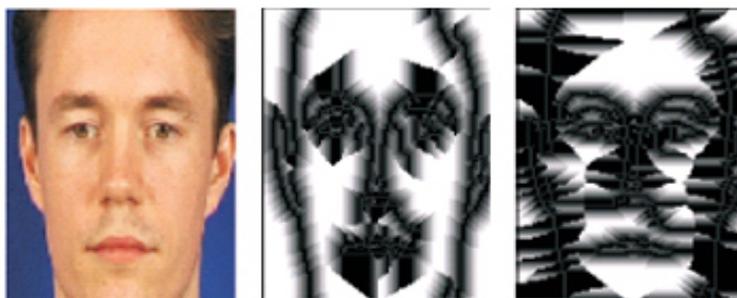
Οι αρχικοί τέσσερις τύποι χαρακτηριστικών γνωρισμάτων τύπου Haar είναι επαρκείς για να κωδικοποιήσουν τις μετωπικές εικόνες προσώπου, ενώ άλλοι τύποι χαρακτηριστικών γνωρισμάτων είναι ουσιαστικοί για να αντιπροσωπεύσουν πιο σύνθετα σχέδια (π.χ., τα πρόσωπα σε διαφορετικό στάση) ([42], [153], [67], [74], [105]).

Η κατασκευή και η επιλογή των σημαντικότερων χαρακτηριστικών γνωρισμάτων τύπου Haar είναι τα πιο βασικά προβλήματα που πρέπει να επιλυθούν πριν την εφαρμογή του αλγορίθμου σε εικόνες. Πιο γενικευμένα χαρακτηριστικά γνωρίσματα τύπου Haar περιγράφονται στις εργασίες ([42], [153]) στις οποίες οι ορθογώνιες περιοχές της εικόνας δεν είναι απαραίτητως γειτονικές, και επιπλέον ο αριθμός τέτοιων ορθογώνιων περιοχών είναι τυχαίος [42]. έχουν Για την αποτελεσματική επιλογή χαρακτηριστικών γνωρισμάτων έχουν προταθεί διάφοροι άπληστοι αλγόριθμοι οι οποίοι εκμεταλλεύονταν τη στατιστική ανάλυση των χαρακτηριστικών γνωρισμάτων πριν την εκπαίδευση των διαδοχικά συνδεδεμένων ταξινομητών ([105], [146]).

Έχει παρατηρηθεί ότι, αν και οι μέθοδοι ανίχνευσης προσώπου έχουν σήμερα εντυπωσιακή σε πραγματικό χρόνο, απόδοση, υπάρχει ακόμα πολύ περιθώριο για βελτίωση από την άποψη της ακρίβειας. Τα ανιχνευμένα πρόσωπα από τους σύγχρονους αλγορίθμους είναι συχνά μερικά εικονοστοιχεία (περίπου 10) εντοπισμένα στις "ακρίβεις" θέσεις, γεγονός ιδιαίτερα σημαντικό αφού ένα πρόσωπο καταλαμβάνει 21x21 εικονοστοιχεία σε μία εικόνα.

4.4.2.2 Ανίχνευση χαρακτηριστικών προσώπου - Εντοπισμός ματιών, μύτης και άξονα αντικίνησης

Για τον εντοπισμό των χαρακτηριστικών ενός προσώπου σε μία εικόνα υπάρχουν δύο κατηγορίες τεχνικών, σύμφωνα με τη βιβλιογραφία. Η πρώτη τεχνική είναι βασισμένη



Σχήμα 4.7: α) Περιοχή προσώπου, β) οριζόντιες συντεταγμένες διανυσματικού πεδίου, γ) κάθετες συντεταγμένες περιοχής προσώπου

στον υπολογισμό ενός συνόλου γεωμετρικών χαρακτηριστικών γνωρισμάτων της εικόνας ενός προσώπου, ενώ η δεύτερη κατηγορία τεχνικών είναι βασισμένη στο ταιρίασμα προτύπων. Στην πρώτη κατηγορία ανήκουν οι τεχνικές ταιριάσματος που βασίζονται σε χαρακτηριστικά γνωρίσματα εκμεταλλεύομενες το ότι ένα πρόσωπο μπορεί να αναγνωριστεί από τις λεπτομέρειες των μεμονωμένων χαρακτηριστικών γνωρισμάτων. Η ιδέα είναι να εξαχθεί η σχετική θέση και άλλες παράμετροι των διακριτών χαρακτηριστικών (μάτια, στόμα, μύτη, πηγούνι). Στη δεύτερη κατηγορία ανήκουν οι τεχνικές ταιριάσματος προτύπου. Σε αυτή την κατηγορία ανήκει η τεχνική που χρησιμοποιήθηκε στο πλαίσιο αυτής της διατριβής για την εύρεση των ματιών σε εικόνες πολιτισμικού περιεχομένου.

Αναλυτικότερα, θέλοντας να εντοπίσουμε τα χαρακτηριστικά του προσώπου υιοθετούμε μία τεχνική ταιριάσματος προτύπων (template matching) με σκοπό την ανίχνευση των χαρακτηριστικών του προσώπου [12]. Η περιοχή προσώπου που έχει εντοπιστεί στο πρώτο στάδιο χρησιμοποιείται ως είσοδος και εξάγεται από αυτήν ο χάρτης ακμών (corresponding edge map). Για κάθε εικονοστοιχείο του χάρτη ακμών, αποθηκεύονται οι συντεταγμένες του διανύσματος το οποίο δείχνει προς το κοντινότερο εικονοστοιχείο ακμών. Το τελικό αποτέλεσμα είναι ένα διανυσματικό πεδίο που κωδικοποιεί τη γεωμετρία του προσώπου (Σχήμα 4.7).

Για την εύρεση πρότυπων διανυσματικών πεδίων της περιοχής του ματιού και του φρυδιού, χρησιμοποιήθηκαν περιοχές ματιών που να εμπεριέχουν το μάτι και το φρύδι και υπολογίστηκαν τα διανυσματικά τους πεδία και τα μέσα διανυσματικά πεδία, χωριστά για το δεξί και το αριστερό μάτι. Αυτά τα διανυσματικά πεδία χρησιμοποιούνται ως πρότυπα. Συνεπώς, στο διανυσματικό πεδίο του προσώπου εντοπίζονται διανυσματικά πεδία παρόμοια με αυτά των μέσων πρότυπων περιοχών ματιών και έτσι γίνεται μια αρχική εκτίμηση της θέσης των ματιών στην εικόνα.

Για την εύρεση των κέντρων των ματιών χρησιμοποιείται η περιοχή που βρέθηκε προηγουμένως, αφού πρώτα αφαιρεθούν οι αντανακλάσεις του φωτός, αυξάνοντας την αποδοτικότητα του αλγορίθμου. Για την αφαίρεση αυτών των αντανακλάσεων, αρχικά η περιοχή του προσώπου μετατρέπεται σε δυαδική μεγιστοποιώντας τη διασπορά μεταξύ κλάσεων (σκούρων και ανοιχτόχρωμων περιοχών) και χρησιμοποιούνται διαδοχικά μορφολογικά ανοίγματα-κλεισίματα. Το αποτέλεσμα αυτής της επεξεργασίας είναι μικρές λευκές περιοχές της δυαδικής εικόνας να "κλείσουν" και οι αντίστοιχές τους στην εικόνα φωτεινότητας αντικαθίσταται από το μέσο όρο γειτονικών εικονοστοιχείων. Τελικό αποτέλεσμα είναι η περιοχή του ματιού χωρίς αντανακλάσεις φωτός. Ακολούθως, με χρήση προβολών διαφορικών εξάγεται το περίγραμμα της ίριδας και, συνεπώς, το κέντρο αντιστοιχεί στο μεσαίο σημείο (κόκκινοι κύκλοι που απεικονίζονται στη δεύτερη εικόνα).

Έχοντας εντοπίσει τα μάτια που αντιστοιχούν στο πρόσωπο μιας μορφής, μπορούμε να προχωρήσουμε στον υπολογισμό της απόστασης μεταξύ των ματιών. Αν αυτή η απόσταση ισούται με H , σύμφωνα με το εγχειρίδιο του Διονυσίου εκ Φουρνά (βλ. Ενότητα 4.3) τόσο θα είναι και το μήκος της μύτης. Αυτό το μέτρο λαμβάνεται ως μια πρώτη εκτίμηση του μήκους της μύτης.

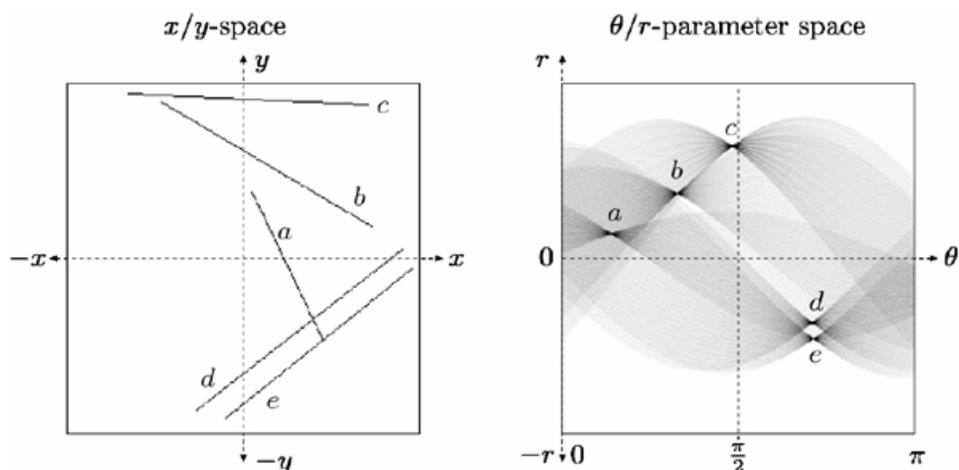
Άρα, για να αναζητήσουμε τη μύτη της μορφής πρέπει να εφαρμόσουμε ένα αλγόριθμο που θα εντοπίζει οριζόντιες και κατακόρυφες γραμμές σε περιοχή που πιθανώς να έχει ζωγραφιστεί η μύτη. Ορίζοντας την πιθανή περιοχή όπου θα πρέπει να αναζητούμε τη μύτη, βελτιώνουμε την ταχύτητα του συστήματος μιας και μικραίνει η περιοχή αναζήτησης. Σύμφωνα με τον Διονύσιο, ο αγιογράφος αρχίζει να σχεδιάζει τη μύτη σε απόσταση $H/5$ πάνω από τον άξονα που ενώνει τα κέντρα των ματιών, δημιουργώντας μία μικρή οριζόντια γραμμή. Το ύψος της μύτης είναι H και ο αγιογράφος χρησιμοποιεί κάθετες γραμμές για να τη σχεδιάσει. Το τέλος της μύτης ορίζεται πάλι με τη σχεδίαση μιας μικρής οριζόντιας γραμμής. Ο άξονας αντικίνησης είναι ο κάθετος άξονας που διατρέχει όλη την εικόνα και ταυτίζεται με τον άξονα που περνά από το μέσο της οριζόντιας γραμμής που ορίζει την αρχή της μύτης.

Άρα η πιθανή περιοχή όπου θα πρέπει να αναζητούμε τη μύτη θα έχει πλάτος ίσο με H ενώ το ύψος της θα είναι $1.8 \cdot H$. Η περιοχή της μύτης, σύμφωνα με τον Διονύσιο, θα βρίσκεται σε απόσταση $H/2$ από το μικρό οφθαλμό της μορφής και σε απόσταση H από το μεγάλο οφθαλμό της μορφής. Δεδομένου ότι κατά τον εντοπισμό των ματιών, ασχοληθήκαμε με την ανίχνευση της ίριδας και όχι με το μέγεθος του οφθαλμού, θα αναζητήσουμε την περιοχή της μύτης σε απόσταση $H/2$ δεξιά από τη θέση εντοπισμού της ίριδας του αριστερού ματιού και ανάλογα πράττουμε και για το δεξί μάτι.

Έχοντας ορίσει την υποπεριοχή της εικόνας που πιθανά θα έχει σχεδιαστεί η μύτη της μορφής, θέλουμε να εντοπίσουμε την ευθεία από την οποία αρχίζει η μύτη, το ύψος αυτής και την ευθεία στην οποία τελειώνει. Ο αλγόριθμος που κατασκευάσαμε, προσπαθεί να προσομοιώσει τις κινήσεις που κάνει το χέρι του αγιογράφου. Έτσι για να εντοπίζουμε την ακριβή θέση των γραμμών στην περιοχή που σχεδιάζει ο αγιογράφος υιοθετούμε μια αντιπροσωπευτική μέθοδος ανίχνευσης ευθειών, το μετασχηματισμός Hough [47] που περιγράφεται στην Ενότητα 4.4.3.

Ο αλγόριθμος αυτός λαμβάνει ως είσοδο την υποπεριοχή του προσώπου στην οποία πιθανολογείται ότι έχει σχεδιάσει ο αγιογράφος τη μύτη της μορφής. Ο αλγόριθμος ανιχνεύει τις οριζόντιες και κάθετες γραμμές, εξάγοντας μια ακριβή μέτρηση του ύψους της μύτης καθώς και τα όρια της περιοχής μέσα στην οποία αυτή έχει εντοπιστεί. Επιπλέον ο αλγόριθμος εντοπίζει τον άξονα αντικίνησης, ο οποίος περνά από το μέσο της μύτης και είναι καθοριστικός, αφού μας παρέχει την πληροφορία σε ποια πλευρά του προσώπου ο αγιογράφος έχει σχεδιάσει τα μεγαλύτερα χαρακτηριστικά της μορφής. Αυτή η πλευρά είναι αυτή στην οποία θα αναζητήσουμε στη συνέχεια την παρειά του προσώπου της μορφής.

4.4.2.2.1 Μετασχηματισμός Hough Ο μετασχηματισμός Hough [47] είναι μια μέθοδος ανίχνευσης ευθειών γραμμών και καμπυλών σε εικόνες. Η βασική ιδέα της μεθόδου μπορεί να αναδειχτεί από το απλό πρόβλημα της ανίχνευσης μιας ευθείας γραμμής σε μια εικόνα. Η εξίσωση με την οποία ορίζεται μια ευθεία γραμμή ϵ είναι η $y = ax + b$ η οποία και ικανοποιεί τα σημεία που ανήκουν στην ευθεία. Χρησιμοποιώντας πολικές συντεταγμένες, η εξίσωση της ευθείας γίνεται $r = x \cdot \cos\theta + y \cdot \sin\theta$, η οποία περιγράφει μία γραμμή με προσανατολισμό θ και σε απόσταση r από την αρχή των αξόνων.



Σχήμα 4.8: Μετασχηματισμός Hough α) Αρχική εικόνα β) Εικόνα μετασχηματισμού Hough

Μια ευθεία γραμμή που διέρχεται από ένα σημείο (x_1, y_1) παριστά μια ημιτονοειδή καμπύλη $r = x_1 \cdot \cos\theta + y_1 \cdot \sin\theta$ στον παραμετρικό χώρο (r, θ) (Σχήμα 4.8). Τα συγγραμικά σημεία (x_i, y_i) στο χώρο της δυαδικής εικόνας θα αντιστοιχούν στην τομή των ημιτονοειδών στον παραμετρικό χώρο. Άρα μία ευθεία στον καρτεσιανό χώρο αντιστοιχεί σε ένα σημείο του παραμετρικού χώρου (Σχήμα 4.8).

Με αυτήν την αναπαράσταση προβάλλεται κανονικοποιημένος ο χώρος των πολικών συντεταγμένων πάνω σε έναν τετραγωνικό χώρο. Η αρχή των αξόνων βρίσκεται στην πάνω αριστερή γωνία, ο άξονας των θ εκτείνεται προς τα δεξιά και ο άξονας των r εκτείνεται προς τα κάτω. Με αυτό τον τρόπο ο παραμετρικός χώρος διακριτοποιείται και σχηματίζεται μία μήτρα $P(r, \theta)$, η μήτρα συσσώρευσης. Για κάθε τιμή της παραμέτρου θ υπολογίζεται η τιμή της παραμέτρου r και το αντίστοιχο στοιχείο της παραμετρικής μήτρας $P(r, \theta)$ αυξάνεται κατά 1. Θέτοντας ένα κατώφλι στην παραμετρική μήτρα προσδιορίζονται οι ευθείες γραμμές που αντιστοιχούν στα σημεία συσσώρευσης με τιμές μεγαλύτερες από την τιμή του κατωφλίου.

4.4.2.3 Κατάτμηση σημασιολογικών περιοχών προσώπου

Η ψηφιοποίηση υλικού, διαδικασία που ανθεί στις μέρες μας, οδήγησε στη δημιουργία μεγάλων βάσεων δεδομένων με πολιτιστικά δεδομένα. Πολλές πολιτιστικές βάσεις έχουν χρησιμοποιηθεί από συστήματα μάθησης στο πλαίσιο της εξ αποστάσεως εκπαίδευσης. Βασική απαίτηση από τέτοιου είδους συστήματα είναι η δυνατότητα αυτόματης τεκμηρίωσης/κατηγοριοποίησης του υλικού βασισμένη στα σημασιολογικά χαρακτηριστικά του.

Η αγιογραφία, όπως προαναφέρθηκε, στηρίζεται σε καθορισμένους κανόνες σχεδίασης που ουσιαστικά αφορούν στον τρόπο σχεδίασης των σημασιολογικών τμημάτων μιας μορφής που απεικονίζεται στην εικόνα. Αξιοποιώντας αυτήν τη γνώση, προχωρήσαμε στη δημιουργία ενός συστήματος κατάτμησης των σημασιολογικών περιοχών της μορφής και συγκεκριμένα της περιοχής του προσώπου.

Έχοντας καθορίσει με ακρίβεια το ύψος της μύτης (Ενότητα 4.4.3) και χρησιμοποιώντας τις αναλογίες των περιοχών του προσώπου της μορφής (Ενότητα 4.3) μπορούμε να οδηγηθούμε στον εντοπισμό τμήματος των σημασιολογικών περιοχών (μαλλιών, μετώπου, παρειάς, μουστακίου και γενειάδας). Για την ανίχνευση

των σημασιολογικών τμημάτων του προσώπου υιοθετήθηκε η μέθοδος κατάτμησης με τομές γράφων [29]. Αυτή η τεχνική λαμβάνει ως είσοδο κάποια εικονοστοιχεία από το φόντο της εικόνας και κάποια από τα εικονοστοιχεία του αντικειμένου που θέλουμε να ξεχωρίσουμε από την εικόνα και βρίσκει την ολική βέλτιστη κατάτμηση της εικόνας, όπως περιγράφεται αναλυτικά στην Ενότητα 4.4.2.3.1.

Για την ανάλυση χρωμάτων του σαρκώματος του προσώπου (Ενότητα 4.3) χρησιμοποιήθηκε η τεχνική κατάτμησης με χρήση κατωφλίων η οποία παρουσιάζεται στην Ενότητα 4.4.3.

4.4.2.3.1 Ανίχνευση τμημάτων προσώπου Η ανίχνευση τμημάτων του προσώπου σε μια Βυζαντινή εικόνα μας οδήγησε στην ανάγκη να κατατμήσουμε την εικόνα σε σημασιολογικές ομοιογενείς περιοχές. Σε αυτό το πλαίσιο αναζητήσαμε τη βέλτιστη λύση για το διαχωρισμό της εικόνας σε ομοιογενείς περιοχές. Οι αλγόριθμοι που έχουν προταθεί για αυτό το πρόβλημα στη βιβλιογραφία χρησιμοποιούν κριτήρια που βασίζονται στην ομοιογένεια των χαρακτηριστικών χαμηλού επιπέδου στο εσωτερικό των περιοχών ή/και στην ασυνέχεια τους εκατέρωθεν των ορίων των περιοχών.

Το πρόβλημα της πλήρως αυτόματης κατάτμησης δείχνει προς το παρόν να μην μπορεί να αντιμετωπιστεί με βέλτιστο τρόπο (NP-hard) πρόβλημα. Στη βιβλιογραφία έχουν προταθεί αλγόριθμοι που στηρίζονται στη διαλογική κατάτμηση. Στόχος της διαλογικής κατάτμησης είναι η διαίρεση μιας εικόνας σε δύο τμήματα, το "αντικείμενο" και το "φόντο". Ο χρήστης του συστήματος διαλογικής κατάτμησης ουσιαστικά επιβάλλει αυστηρούς περιορισμούς για την κατάτμηση με την ένδειξη ορισμένων εικονοστοιχείων (σπόροι-seeds) που πρέπει να περιέχονται στο αντικείμενο και ορισμένων εικονοστοιχείων που πρέπει να είναι μέρος του φόντου. Αυτοί οι αυστηροί περιορισμοί είναι οι ενδείξεις για το τι θεωρεί ο χρήστης αντικείμενο και τι φόντο.

Ένα πρόβλημα κατάτμησης εκφράζεται από ένα σύνολο εικονοστοιχείων \mathcal{P} και από ένα σύνολο ετικετών \mathcal{L} και η λύση του προβλήματος έγκειται στο να αναθέσουμε μία ετικέτα από το σύνολο \mathcal{L} σε κάθε ένα από τα σημεία στο \mathcal{P} . Το σύνολο $f = \text{emph}f_p | p \in \mathcal{P}$ ονομάζεται αντιστοίχιση ετικετών (labelling) και $f: \mathcal{P} \mapsto \mathcal{L}$. Ο χώρος των δυνατών αντιστοιχίσεων (labelling space) είναι το καρτεσιανό γινόμενο $f = \mathcal{L}^{|\mathcal{P}|}$.

Η τεχνική κατάτμησης [29] που υιοθετήθηκε, ανήκει στην κατηγορία των μεθόδων ολικής βελτιστοποίησης, οι οποίες στηρίζονται σε συναρτήσεις ενέργειας.

4.4.2.3.2 Τυχαία μαρκοβιανά μοντέλα και μέθοδοι ελαχιστοποίησης ενέργειας Αναλυτικότερα, η μέθοδος που μελετήθηκε υιοθετεί η συνάρτηση κόστους $E(f)$ που υπολογίζεται με τη χρήση της μεθόδου εκτίμησης της μέγιστης εκ των υστέρων πιθανότητας σε ένα τυχαίο πεδίο Markov (MAP-MRF) [55]. Η χρήση Markov πεδίων είναι πολύτιμη αφού μας δίνει τη δυνατότητα να κωδικοποιήσουμε τις αλληλεπιδράσεις μεταξύ των εικονοστοιχείων καθώς και να θέσουμε περιορισμούς για τη συνέχεια των ποσοτήτων που εκτιμούμε.

Η εύρεση του ελαχίστου της συνάρτησης ενέργειας που έχει τη γενική μορφή:

$$E(f) = \sum_{p \in \mathcal{P}} D_p(f_p) + \sum_{\{p,q\} \in \mathcal{P}} V_{(p,q)} f_p f_q \quad (4.9)$$

ουσιαστικά μας οδηγεί στην κατάτμηση της εικόνας.

Για να λάβουμε υπόψη αυτές τις αλληλεπιδράσεις θεωρούμε ένα αυθαίρετο σύνολο στοιχείων \mathcal{P} (π.χ. Τα εικονοστοιχεία μιας εικόνας) και ορίζουμε ένα σύστημα γειτονιών N τέτοιο ώστε $N = N_p | \forall p \in \mathcal{P}$ όπου N_p είναι τα γειτονικά σημεία του p για τα οποία ισχύει: α) ένα στοιχείο δεν είναι γειτονικό με τον εαυτό του: $p \notin N_p$ και β) η σχέση γειτνίασης είναι αμοιβαία $p \in N_q \Leftrightarrow q \in N_p$. Η κλίκα (clique) είναι ένα σύνολο σημείων όπου για κάθε σημείο της τα υπόλοιπα σημεία ανήκουν στη γειτονιά του.

Για να καθορίζουμε ένα τυχαίο πεδίο Markov μπορούμε να καθορίσουμε το δυναμικό της κάθε κλίκας (σύνολο σημείων όπου για κάθε σημείο του, τα υπόλοιπα σημεία ανήκουν στη γειτονιά του). Έστω μια οικογένεια τυχαίων μεταβλητών $F = F_1, F_2, \dots, F_n$ ορισμένες στο \mathcal{P} , όπου κάθε F_p παίρνει τιμές από το σύνολο των ετικετών \mathcal{L} . Η F θα ονομάζεται τυχαίο πεδίο Markov αν ισχύουν οι συνθήκες:

1. $P(f) > 0, \forall f \in F$
2. $P(f_p | f_{\mathcal{P}-p}) = P(f_p | f_{N_p})$

Με την πρώτη συνθήκη εξασφαλίζεται ότι η από κοινού συνάρτηση πυκνότητας πιθανότητας μπορεί να καθοριστεί μοναδικά από τις τοπικές δεσμευμένες πιθανότητες, ενώ η δεύτερη δηλώνει ότι η ετικέτα για κάθε σημείο p εξαρτάται μόνο από τα γειτονικά σημεία του και όχι από όλο το πεδίο. Στην εργασία [57], αποδείχθη ότι ένα τυχαίο πεδίο Markov είναι ισοδύναμο με ένα τυχαίο πεδίο Gibbs που καθορίζεται πλήρως από την κατανομή Gibbs:

$$P(f) = Z^{-1} \cdot \exp\left(-\sum_{c \in C} V_c(f)\right) \quad (4.10)$$

όπου C το σύνολο των κλικών στο πεδίο, Z παράγοντας κανονικοποίησης (καλείται partition function) και V_c συναρτήσεις με πεδίο ορισμού μια αντιστοίχιση ετικετών και πεδίο τιμών πραγματικούς αριθμούς που καλούνται δυναμικά των κλικών (clique potentials).

Για κλίκες μεγέθους 2 το δυναμικό καθορίζεται από τη σχέση: $V_c(F) = V_{\{p,q\}}(f_p, f_q)$ όπου $V_{\{p,q\}}(f_p, f_q)$ αντιστοιχούν σε πιθανές υποθέσεις για τη συνέχεια της αντιστοίχισης. Για παράδειγμα, αν υποθέσουμε ότι η f είναι κατά τμήματα σταθερή, θα μπορούσαμε να αναθέσουμε ποινή αν δύο γειτονικά σημεία έχουν διαφορετική ετικέτα ορίζοντας τη συνάρτηση $V_c(F) = V(f_p, f_q)$ από την οποία με χρήση κατανομής (π.χ. Gibbs) υπολογίζεται η $P(f) = Z^{-1} \cdot \exp\left(-\sum_{p,q \in \mathcal{N}} V_c(f)\right)$.

Η F δεν είναι γνωστή, άρα δε μπορούμε να υπολογίσουμε άμεσα την αντιστοίχιση $\hat{f} = \operatorname{argmin}_{f \in F} E(f)$. Με τη βοήθεια της συνάρτησης πιθανοφάνειας (likelihood) $p(d|f)$ μπορούμε να εκτιμούμε την f , μεγιστοποιώντας την εκ των υστέρων πιθανότητα (maximum a posteriori probability) $p(f|d)$ ή ισοδύναμα, σύμφωνα με το θεώρημα Bayes, το γινόμενο $p(d|f)p(f)$, όπου d είναι το χαρακτηριστικό γνώρισμα που χρησιμοποιούμε για την κατάτμηση. Η μέγιστη εκ των υστέρων εκτίμηση f^* ισούται με την αντιστοίχιση, η οποία μεγιστοποιεί το γινόμενο $p(d|f)p(f)$:

$$f^* = \operatorname{argmax}_{f \in F} p(d|f)p(f) \quad (4.11)$$

Ο όρος $p(f)$ δίνεται από την εξίσωση 4.10, άρα χρειάζεται να υπολογισθεί ο όρος $p(d|f)$. Υποθέτοντας ότι σε κάθε σημείο p έχουμε μια μέτρηση d_p και ότι $p(d|f) = \prod_{p \in \mathcal{P}} p(d_p | f_p)$ ορίζουμε:

$$p(d_p | l) = C_p \cdot \exp(-D_p(l)), l \in \mathcal{L}_p(d_p | l) \propto \exp\left(-\sum_{p \in \mathcal{P}} D_p(f_p p)\right) \quad (4.12)$$

όπου C_p είναι μια σταθερά κανονικοποίησης και $D_p(l)$ μια συνάρτηση που μετρά την ακρίβεια της αντιστοίχισης του εικονοστοιχείου p με την ετικέτα l .

Από τις εξισώσεις 4.10, 4.11 και 4.12 προκύπτει :

$$f^* = \operatorname{argmax}_{f \in F} \left(- \sum_{\{p,q\} \in \mathcal{N}} V(f_p, f_q) - \sum D_p(f_p) \right) \quad (4.13)$$

το οποίο είναι ισοδύναμο με την ελαχιστοποίηση της Εξίσωσης 4.9. Άρα η ελαχιστοποίηση της ενέργειας ισοδυναμεί με τη μεγιστοποίηση της εκ των υστέρων πιθανότητας σε ένα τυχαίο πεδίο Markov.

4.4.2.3.3 Γράφοι και μέγιστη ροή - ελάχιστη τομή Οι αλγόριθμοι που έχουν προταθεί στη βιβλιογραφία αναπαριστούν το πεδίο που πρόκειται να καταταμηθεί με τη βοήθεια βεβαρημένων γράφων (γράφους-δίκτυα). Θεωρώντας κάθε εικονοστοιχείο της εικόνας ως έναν κόμβο του γράφου, κάθε γειτονική σχέση μεταξύ των εικονοστοιχείων ως μία ακμή και την ομοιότητα μεταξύ δύο εικονοστοιχείων ως τα βάρη των ακμών του γράφου δημιουργούμε ένα βεβαρημένων γράφο (Σχήμα 4.4.2.3.4.β).

Έστω $G = (V, E, w_e, s, t)$ ο γράφος με σύνολο κόμβων V , σύνολο ακμών E , συνάρτηση βάρους w_e , κόμβος πηγής s και κόμβος δεξαμενής t . Θεωρούμε ότι ο γράφος G είναι συνδεδεμένος δηλαδή ότι κάθε κόμβος βρίσκεται σε ένα μονοπάτι από την πηγή προς τη δεξαμενή $s \rightsquigarrow u \rightsquigarrow t$, και ότι η συνάρτηση βάρους $w_e : E \rightarrow \mathbb{R}_+ \cup \{0\}$ είναι μη αρνητική και αναπαριστά τη χωρητικότητα του γράφου.

Μια ροή στο γράφο G είναι μια πραγματική συνάρτηση $f : V \times V \rightarrow \mathbb{R}$ για την οποία ισχύουν οι εξής ιδιότητες:

1. $\forall u, v \in V : f(u, v) \leq w_e(u, v)$ (Περιορισμός χωρητικότητας: η ροή από τον κόμβο u στον άλλον κόμβο v είναι το πολύ ίση με τη χωρητικότητα της ακμής w_e),
2. $\forall u, v \in V : f(u, v) = -f(v, u)$ (Περιττή Συμμετρία: η ροή προς μια κατεύθυνση είναι η αντίθετη της αντίστροφης κατεύθυνσης),
3. $\forall u \in V \setminus \{s, t\} : \sum_{v \in V} f(u, v) = 0$ (Διατήρηση της ροής: η συνολική ροή σε έναν κόμβο είναι 0, εκτός της πηγής και της δεξαμενής).

Η τιμή της ροής $f(u, v)$ ουσιαστικά είναι η συνολική ροή που εξέρχεται από την πηγή και δίνεται από τον τύπο $|f| = \sum_{u \in V} f(u, v)$. Η τομή $C \cup E$ σε ένα δίκτυο $G = (V, E, w_e, s, t)$ είναι μια διαμέριση του συνόλου των κόμβων V σε δύο σύνολα S και $T = V \setminus S$ τέτοια ώστε $s \in S$ και $t \in T$.

Το κόστος (ή βάρος ή χωρητικότητα) κάποιας τομής C ορίζεται ως το άθροισμα των βαρών των ακμών που αυτό διαμερίζει:

$$|C| = \sum_{e \in E} w_e$$

Μια τομή ενός δικτύου της οποίας το κόστος είναι ελάχιστο ονομάζεται ελάχιστη τομή.

Σύμφωνα με το θεώρημα μέγιστης ροής ελάχιστης τομής [51], το κόστος της ελάχιστης τομής του δικτύου είναι ίσο με το άνω φράγμα της τιμής της μέγιστης ροής πάνω στο δίκτυο G . Το κόστος της ελάχιστης τομής του δικτύου δημιουργεί μια κατάτμηση η οποία είναι βέλτιστη σύμφωνα με τις ιδιότητες των βαρών των ακμών.

4.4.2.3.4 Τεχνική Κατάτμησης Η τεχνική με χρήση τομών γράφων που παρουσιάστηκε στην [29] ανάγει το πρόβλημα της κατάτμησης στην εύρεση των τομών που ελαχιστοποιούν την ενέργεια [30] όπως αυτή ορίστηκε στην εξίσωση 4.9.

Η εξέταση των αυστηρών περιορισμών οδηγεί στον καθορισμό των περιοχών κατάτμησης και όχι στο όριο αυτών. Ας υποθέσουμε ότι μερικά εικονοστοιχεία είναι χαρακτηρισμένα ως εσωτερικά και μερικά ως εξωτερικά για το αντικείμενο ενδιαφέροντος. Τα υποσύνολα των χαρακτηρισμένων εικονοστοιχείων θα θεωρηθούν σπόροι αντικειμένου και σπόροι φόντου αντίστοιχα. Το όριο κατάτμησης μπορεί να είναι οπουδήποτε, αλλά πρέπει να ξεχωρίζει τους σπόρους του αντικειμένου από τους σπόρους του φόντου. Επιπλέον οι σπόροι μπορούν να τοποθετηθούν αόριστα μέσα στις περιοχές αντικειμένου και φόντου. Προφανώς οι αυστηροί περιορισμοί από μόνοι τους δεν είναι αρκετοί για να πετύχουν μια καλή κατάτμηση. Αν τα όρια μεταξύ των περιοχών δεν είναι βέλτιστα, η κατάτμηση δε θα είναι ακριβής όταν τα όρια είναι θολά.

Στην εργασία [29], η κατάτμηση της εικόνας γίνεται με τον υπολογισμό ενός ολικού βέλτιστου μεταξύ όλων των κατατμήσεων που ικανοποιούν τους αυστηρούς περιορισμούς. Για τον υπολογισμό του βέλτιστου, καθορίζεται η συνάρτηση κόστους λαμβάνοντας υπόψη τόσο το όριο όσο και τα χαρακτηριστικά της περιοχής κατάτμησης, τα οποία μπορούν να αντιμετωπισθούν ως χαλαροί περιορισμοί για την κατάτμηση.

Έστω $A = \{A_1, A_2, \dots, A_n\}$ είναι ένα διάνυσμα του οποίου τα στοιχεία A_p δείχνουν την παραχώρηση του p εικονοστοιχείου στο σύνολο \mathcal{P} . Κάθε στοιχείο A_p μπορεί να είναι είτε 'obj' (object), δηλαδή αντικείμενο, ή 'bkg' (background), δηλαδή φόντο. Το διάνυσμα A ορίζει μια κατάτμηση.

Ας υποθέσουμε ότι τα \mathcal{O} και \mathcal{B} είναι τα υποσύνολα με τα εικονοστοιχεία του αντικειμένου και του φόντου αντίστοιχα. Θα ισχύει ότι: $\mathcal{O} \subset \mathcal{P}$ και $\mathcal{B} \subset \mathcal{P}$ με $\mathcal{O} \cap \mathcal{B} = \emptyset$. Έτσι για κάθε εικονοστοιχείο θα έχουμε :

$$\forall p \in \mathcal{O}, f_p = 'obj' \text{ και } \forall p \in \mathcal{B}, f_p = 'bkg'$$

Οι χαλαροί περιορισμοί που επιβλήθηκαν στο όριο και στα χαρακτηριστικά της περιοχής κατάτμησης περιγράφονται από τη συνάρτηση κόστους $E(A)$ σε αντιστοιχία με την Εξίσωση 4.9:

$$E(A) = \lambda \cdot R(A) + B(A) \quad (4.14)$$

όπου

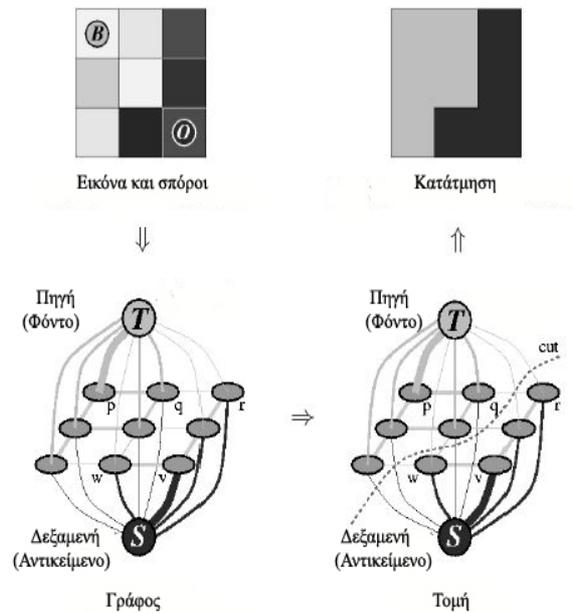
$$R(A) = \sum_{p \in \mathcal{P}} R_p(A_p), \quad (4.15)$$

$$B(A) = \sum_{\{p,q\} \in \mathcal{N}} B_{p,q} \cdot \delta(A_p, A_q) \quad (4.16)$$

$$\delta(A_p, A_q) = \begin{cases} 1, & \text{αν } A_p \neq A_q \\ 0, & \text{αλλιου} \end{cases} \quad (4.17)$$

Ο συντελεστής λ στην εξίσωση 4.14 καθορίζει μια σχετική βαρύτητα των χαρακτηριστικών της περιοχής $R(A)$ έναντι στο όριο της περιοχής $B(A)$.

Ο όρος $R(A)$ υποθέτει ότι οι μεμονωμένες ποινές για την ανάθεση του εικονοστοιχείου p στο "αντικείμενο" R_p ("obj") ή στο φόντο R_p ("bkg") είναι γνωστές.



Σχήμα 4.9: Παράδειγμα κατάτμησης εικόνας 3×3 όπως εμφανίζεται στην εργασία [29]

Για παράδειγμα, $R_p(\cdot)$ μπορεί να απεικονίζει τον τρόπο που η φωτεινότητα του εικονοστοιχείου p ταιριάζει σε ένα γνωστό μοντέλο φωτεινότητας (π.χ. ιστόγραμμα) για το αντικείμενο και για το φόντο.

Ο όρος $B(A)$ περικλείει το χαρακτηριστικό "όριο" της κατάτμησης της περιοχής A . Ο συντελεστής $B_{p,q} \geq 0$ θεωρείται η ποινή για πιθανή ασυνέχεια μεταξύ των εικονοστοιχείων p και q . Άρα το $B_{p,q}$ αυξάνει όταν τα εικονοστοιχεία p και q είναι όμοια, ενώ τείνει στο μηδέν στην αντίθετη περίπτωση, δηλαδή η ποινή $B_{p,q}$ αυξάνεται συναρτήσει της απόστασης των εικονοστοιχείων p και q . Η ποινή $B_{p,q}$ μπορεί να βασιστεί σε κριτήρια όπως τοπικά μέγιστα της φωτεινότητας, Laplacian zero-crossing, κατεύθυνση κλίσης και άλλα.

Στόχος της τεχνικής είναι ο υπολογισμός της βέλτιστης ελάχιστης τομής του δικτύου η οποία θα είναι ικανή να χωρίσει την πηγή από τη δεξαμενή.

Δεδομένης της εικόνας (Σχήμα) δημιουργείται ένας γράφος με δύο τερματικά σημεία, ένα για τα εικονοστοιχεία του αντικειμένου (σπόροι O) και ένα για τα εικονοστοιχεία του φόντου (σπόροι B). Τα βάρη των ακμών αντιστοιχούν στις παραμέτρους των όρων $R(A)$ (Εξ. 4.15) και $B(A)$ (Εξ. 4.16) της συνάρτησης κόστους καθώς και στις θέσεις των σπόρων της εικόνας.

Κάθε σύνολο ακμών E αποτελείται από δύο τύπους μη κατευθυνόμενων ακμών: των n-links (neighborhood links) και των t-links (terminal links). Κάθε εικονοστοιχείο έχει δύο t-links τα $\{p, S\}$ και $\{p, T\}$ που το συνδέουν με την πηγή και τη δεξαμενή αντίστοιχα. Κάθε ζεύγος γειτονικών εικονοστοιχείων $\{p, q\}$ στο \mathcal{N} συνδέονται με ένα n-link. Ο πίνακας 4.6 δίνει τα βάρη των ακμών του συνόλου $E = \cup_{\{p\} \in \mathcal{P}} \{\{p, S\}, \{p, T\}\}$.

Ας θεωρήσουμε ως \mathcal{F} το σύνολο όλων των δυνατών τομών C στο γράφο G έτσι ώστε να ισχύουν τα εξής:

- το C διαχωρίζει ακριβώς ένα t-link για κάθε p
- $p, q \in C$, αν τα p, q είναι συνδεδεμένα με t-link σε διαφορετικά τερματικά (πηγή

Πίνακας 4.6: Βάρη Ακμών Συνόλου E

Ακμή	Βάρος	
$\{p, q\}$	$B(p, q)$	$\{p, q\} \in \mathcal{N}$
$\{p, S\}$	$\lambda \cdot R_p('bkg')$	$p \in \mathcal{P}, p \ni \mathcal{O} \cup \mathcal{B}$
	0	$p \in \mathcal{O}$
	$1 + \max_{p \in \mathcal{P}} \sum_{q: \{p, q\} \in \mathcal{N}} B(\{p, q\})$	$p \in \mathcal{B}$
$\{p, T\}$	$\lambda \cdot R_p('obj')$	$p \in \mathcal{P}, p \ni \mathcal{O} \cup \mathcal{B}$
	0	$p \in \mathcal{O}$
	$1 + \max_{p \in \mathcal{P}} \sum_{q: \{p, q\} \in \mathcal{N}} B(\{p, q\})$	$p \in \mathcal{B}$

και δεξαμενή)

- Αν $p \in \mathcal{O}$, τότε $p, T \in C$
- Αν $p \in \mathcal{B}$, τότε $p, S \in C$

Υπό αυτές τις συνθήκες μπορεί να εφαρμοστεί η ελάχιστη τομή στο γράφο G . Η απόδειξη παρατίθεται στην εργασία [29]. Επιπλέον μπορεί να ορισθεί μία ανάλογη κατάτμηση $\hat{A} = A(\hat{C})$, τέτοια ώστε

$$A_p(C) = \begin{cases} 'obj', & \text{αν } \{p, T\} \in C \\ 'bkg', & \text{αν } \{p, S\} \in C \end{cases} \quad (4.18)$$

Η τεχνική των τομών γράφων [30],[29] έχει βρει εφαρμογή σε πλήθος προβλημάτων κατάτμησης. Αυτή η τεχνική εφαρμόστηκε και για την κατάτμηση των τμημάτων του προσώπου μιας μορφής που εμφανίζεται σε βυζαντινή εικόνα. Για την κατάτμηση κάθε σημασιολογικού τμήματος του προσώπου μιας μορφής σε μία αγιογραφία, χρησιμοποιούμε το ύψος της μύτης H και μαζί με το εγχειρίδιο της αγιογραφίας, ορίζουμε τους "σπόρους" που απαιτεί η τεχνική κατάτμησης με τομές γράφων. Η τεχνική αυτή δέχεται ως είσοδο σπόρους οι οποίοι είναι εικονοστοιχεία που ανήκουν είτε στην περιοχή του αντικειμένου είτε στην περιοχή του φόντου. Αυτοί οι σπόροι θεωρούνται ως αυστηροί περιορισμοί που καταδεικνύουν την πρόθεση μας για το ποιο αντικείμενο θέλουμε να εξάγουμε από την εικόνα. Τα υπόλοιπα εικονοστοιχεία της εικόνας κατηγοριοποιούνται υπολογίζοντας ένα ολικό βέλτιστο ανάμεσα σε όλες τις κατατμήσεις με τέτοιο τρόπο ώστε να ικανοποιούνται οι αυστηροί περιορισμοί.

Για παράδειγμα, γνωρίζοντας το ύψος της μύτης και τη θέση του άξονα αντικίνησης, θέλουμε να εξάγουμε την περιοχή του μαλλιού της μορφής. Ο αλγόριθμος που υλοποιήσαμε ορίζει ως σπόρους αντικειμένου, εικονοστοιχεία που βρίσκονται γύρω από τον άξονα αντικίνησης και μέσα στην περιοχή του τμήματος του μαλλιού και ως σπόρους φόντου εικονοστοιχεία τόσο από το μέτωπο της μορφής, όσο και από το φόντο της εικόνας (συνήθως χρυσού χρώματος). Η χρήση του αλγορίθμου κατάτμησης με τομή γράφων μας εξάγει την περιοχή μαλλιού, όπως αυτή φαίνεται στη μονάδα Ανίχνευση Τμημάτων Προσώπου της εικόνας 4.4. Ομοίως πραγματοποιείται και η κατάτμηση των υπολοίπων περιοχών.

4.4.2.4 Ανάλυση χρωμάτων σαρκωμάτων

Το κριτήριο βέλτιστου κατωφλίου που χρησιμοποιεί η μέθοδος της διασποράς [101] είναι η μεγιστοποίηση της διαχωρισιμότητας μεταξύ φωτεινών και σκοτεινών



Σχήμα 4.10: Εφαρμογή του αλγορίθμου με επιλογή διαφορετικού πλήθους κατωφλίων

περιοχών. Έστω μία εικόνα με L επίπεδα φωτεινότητας και N εικονοστοιχεία όπου $N = \sum_{i=1}^L n_i$. Το ιστόγραμμα της εικόνας μπορούμε να το κανονικοποιήσουμε χρησιμοποιώντας ως μεταβλητές τις πιθανότητες εμφάνισης κάθε φωτεινότητας:

$$p_i = \frac{n_i}{N}, p_i \geq 0, \sum_{i=1}^L p_i = 1$$

αντιστοιχούν στο φόντο και στο αντικείμενο. Θεωρώντας μια αρχική τιμή κατωφλίου, έστω k , δημιουργούνται δύο κλάσεις φωτεινότητας $C_0 = \{g_1, \dots, g_k\}$ και $C_1 = \{g_{k+1}, \dots, g_{L-1}\}$. Για κάθε εικονοστοιχείο υπολογίζουμε την πιθανότητα να ανήκει ή όχι σε μία από τις δύο κλάσεις. Επιπλέον, υπολογίζουμε τις μηδενικού, πρώτου βαθμού και δεύτερου αθροιστικές ροπές του ιστογράμματος έως το επίπεδο k καθώς και το συνολικό μέσο επίπεδο της εικόνας.

Το βέλτιστο κατώφλι υπολογίζεται μεγιστοποιώντας την ακόλουθη συνάρτηση του k :

$$\eta(k) = \sigma_B^2 / \sigma_L^2 \quad (4.19)$$

όπου $\sigma_B^2 = \omega_0 \cdot \omega_1 (\mu_1 - \mu_0)^2$ και $\sigma_L^2 = \sum_{i=1}^L (i - \mu_L)^2 p_i$

Η βέλτιστη τιμή κατωφλίου k_T , η οποία μεγιστοποιεί το $\eta(k)$ ή ισοδύναμα μεγιστοποιεί το σ_B^2 δίνεται από τη σχέση

$$\sigma_B^2(k_T) = \max_{2 \leq k \leq L} \sigma_B^2(k) = \frac{[\mu_L \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (4.20)$$

Η τεχνική αυτή μπορεί να εφαρμοστεί χρησιμοποιώντας πολλαπλά κατώφλια, εξασφαλίζοντας με αυτό τον τρόπο την τμηματοποίηση της εικόνας. Τα πολλαπλά κατώφλια μπορούν να υπολογιστούν με τη μεγιστοποίηση του $\sigma_B^2(k_T)$ σύμφωνα με τη σχέση 4.20.

Θέλοντας να υπολογίσουμε τα τέσσερα βασικά επίπεδα φωτεινότητας (σκούρεμα, προπλασμός, σάρκωμα, φώτισμα), που χρησιμοποιούνται από τον αγιογράφο ως σαρκώματα στην περιοχή του προσώπου 4.2 χρησιμοποιούμε την προτεινόμενη τεχνική. Διαλέγοντας τέσσερα κατώφλια μπορούμε να εξάγουμε περιοχές με παραπλήσια φωτεινότητα που αντιστοιχούν στις περιοχές των σαρκωμάτων του προσώπου (Σχήμα 4.10).

4.4.3 Εξαγωγή χαρακτηριστικών από τα σημασιολογικά τμήματα και σημασιολογική ερμηνεία

Η ανάλυση blob είναι μια από τις βασικές μεθόδους που χρησιμοποιούνται ευρέως στη αναγνώριση προτύπων. Ένα blob (σταγόνα) ορίζεται ως περιοχή από συνεκτικά εικονοστοιχεία. Ο αλγόριθμος ανάλυσης blob διακρίνει τα εικονοστοιχεία βάσει τις τιμής της φωτεινότητας τους. Έπειτα τα ταξινομεί σε μία από τις δύο κατηγορίες που διαθέτει. Όσα εικονοστοιχεία έχουν μη μηδενικές τιμές τα θεωρεί στοιχεία του αντικειμένου ενώ τα υπόλοιπα στοιχεία του φόντου. Στο επόμενο βήμα του αλγορίθμου χρησιμοποιούνται οι κανόνες συνδετικότητας γειτονιάς (4 ή 8 στοιχείων) με σκοπό την εξαγωγή των διακριτών αντικειμένων. Για τα εξαγόμενα αντικείμενα μπορούμε να υπολογίσουμε διάφορα χαρακτηριστικά τους, τα οποία παρέχουν σημαντικές πληροφορίες για το αντικείμενο όπως θέση, μέγεθος, προσανατολισμός, εκκεντρότητα του αντικειμένου κ.α. Η εκκεντρότητα προσδιορίζει πόσο μία κωνική τομή "απέχει" από το να είναι τέλειος κύκλος.

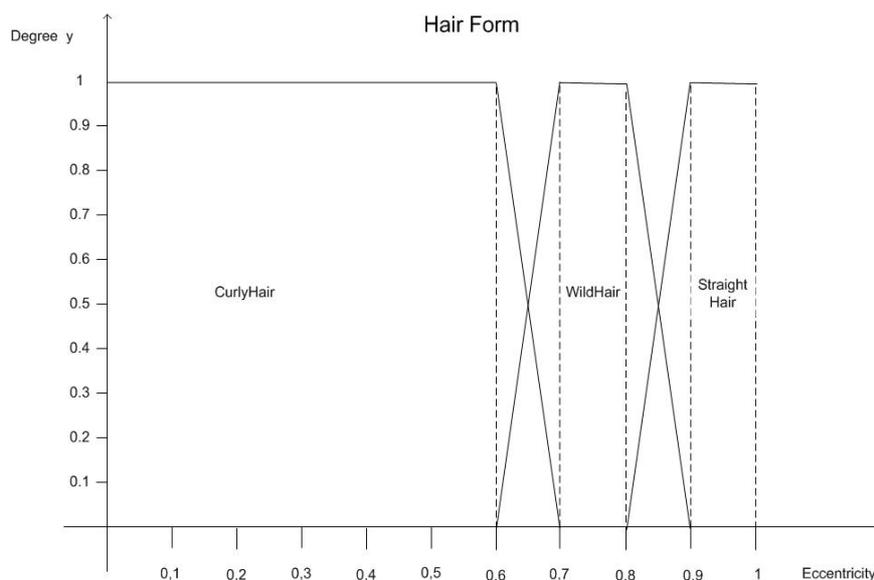
Στα πλεονεκτήματα της blob ανάλυσης συγκαταλέγονται η υψηλή ταχύτητα, η ακρίβεια και η ανεκτικότητα σε μεταβολές στον προσανατολισμό και στο μέγεθος. Τα μειονεκτήματα περιλαμβάνουν την ανικανότητα να ανταποκρίνεται στα επικαλυπτόμενα αντικείμενα.

Έχοντας εξάγει κάθε σημασιολογική περιοχή του προσώπου που απεικονίζεται σε μία αγιογραφία, προχωράμε στην εξαγωγή των χαρακτηριστικών της κάθε περιοχής, όπως αυτά ορίζονται στους Πίνακες 4.1, 4.2, 4.3, 4.4 και 4.5.

Όσον αφορά στη σημασιολογική περιοχή των μαλλιών, μπορούμε να εξάγουμε πληροφορίες σχετικά με το μήκος, το χρώμα, το σχήμα και την πυκνότητα (Πίνακας 4.1). Ως μονάδα μέτρησης ορίζεται το ύψος της μύτης, ο υπολογισμός του οποίου περιγράφηκε στην Ενότητα 4.4.3. Για να υπολογίσουμε το μήκος των μαλλιών απλά μετράμε το ύψος της εξαγόμενης περιοχής των μαλλιών. Χρησιμοποιώντας τη γνώση από το εγχειρίδιο του Διονυσίου εκ Φουρνά, μπορούμε να προβούμε σε ακριβή εκτίμηση του μήκους των μαλλιών.

Δεν μπορούμε με μεγάλη βεβαιότητα να αποφανθούμε για την ακριβή μέτρηση του μήκους των μαλλιών ή οποιασδήποτε άλλης μέτρησης μιας και υπεισέρχονται πολλοί παράγοντες κυρίως από την ορθότητα των αποτελεσμάτων των προηγούμενων σταδίων της επεξεργασίας της εικόνας. Σύμφωνα με την αρχή της απροσδιοριστίας του Heisenberg, όταν η πολυπλοκότητα ενός συστήματος αυξάνει, η ικανότητα μας να προβαίνουμε σε ακριβείς και σημαντικές δηλώσεις για τη συμπεριφορά του μειώνεται μέχρι που φθάνουμε σε ένα σημείο πέραν του οποίου ακρίβεια και σημαντικότητα καθίστανται σχεδόν αμοιβαίως αποκλειόμενα χαρακτηριστικά. Η ασαφής λογική λοιπόν υπεισέρχεται σε αυτό το σημείο για να δώσει μια ικανοποιητική λύση. Για να αντιμετωπιστούν εκ των προτέρων οι λανθασμένες αποφάσεις για τον ακριβή χαρακτηρισμό της κάθε περιοχής, υιοθετούμε τον ασαφή διαμερισμό του πεδίου ορισμού του κάθε χαρακτηριστικού της.

Ας υποθέσουμε ότι οι τιμές ενός συγκεκριμένου χαρακτηριστικού γνωρίσματος βρίσκονται στο διάστημα X . Ένα ασαφές σύνολο A του X ορίζεται ως το σύνολο διατεταγμένων ζευγών $A = (x, \mu_A(x)) | x \in X, \mu_A(x) : X \rightarrow [0, 1]$ όπου x είναι ένα στοιχείο του X και $\mu_A(x)$ η συνάρτηση συμμετοχής του στοιχείου αυτού στο υποσύνολο A . Οι τιμές του $\mu_A(x)$ καλούνται βαθμοί συμμετοχής του x στο σύνολο A . Μια οικογένεια ασαφών υποσυνόλων του X λέγεται ασαφής διαμέριση $P^n(X)$ του X , τάξης $n, n \in N$ και συμβολίζεται με $\mathcal{A}^n = A_1, A_2, \dots, A_n$ αν και μόνο αν



Σχήμα 4.11: Ασαφής Διαμέριση Σχήματος Μαλλιών

$A_i \neq A_j, \forall i, j \in N_n (i \neq j)$ και $0 < \sum_{k=1}^m A_i(x_k) < m, \forall i \in N_n$. Τα στοιχεία $A_i, i \in N_n$ της \mathcal{A}^n θα λέγονται κλάσεις της ασαφούς διαμέρισης.

Έτσι, καθορίζουμε μια κατάλληλη ασαφή διαμέριση (ένα σύνολο ασαφών υποσυνόλων) του X και συνδέουμε κάθε ασαφές σύνολο της διαμέρισης με μια υποέννοια της ορολογίας - κατά αναλογία με τις έννοιες του Πίνακα 4.1. Παραδείγματος χάριν, σύμφωνα με τον Διονύσιο εκ Φουρνά, το μήκος της σημασιολογικής περιοχής των μαλλιών είναι σε αναλογία με το μήκος της μετρικής που χρησιμοποιεί ο αγιογράφος δηλαδή της μύτης. Έστω L_h το μήκος που προκύπτει για μια συγκεκριμένη περιοχή μαλλιού. Αν το $L_h \in [1 \cdot H \dots 3 \cdot H]$, η περιοχή του μαλλιού θα χαρακτηρίζεται ως "Τας τρίχας έχων άνωθεν των ωτίων" ενώ αν το $L_h \in [2.5 \cdot H \dots 5 \cdot H]$, η περιοχή του μαλλιού θα χαρακτηρίζεται ως "Τας τρίχας έχων κάτωθεν των ωτίων". Το γράφημα που απεικονίζεται στη μονάδα Σημασιολογική Ερμηνεία του Σχήματος 4.4 αναπαριστά τη διαμέριση του συνόλου που αφορά στην έννοια Μήκος μαλλιών.

Όσον αφορά στην περιοχή των μαλλιών, μπορούμε να εξάγουμε επιπλέον πληροφορία για το αν υπάρχει αριστερά ή δεξιά ή και από τις δύο πλευρές πλοχμός. Έτσι, αν το μήκος της δεξιάς πλευράς είναι μεγαλύτερο από $3 \cdot H$, μπορούμε να αποφανθούμε με βεβαιότητα ότι η μορφή έχει δεξί πλοχμό. Το ίδιο θα ισχύει και για την αριστερή πλευρά της περιοχής των μαλλιών. (Σημείωση: Ως αριστερή και δεξιά πλευρά της μορφής αναφέρεται η πλευρά της απεικονιζόμενης μορφής της εικόνας)

Για το χρώμα της περιοχής των μαλλιών χρησιμοποιούμε το ιστόγραμμα της φωτεινότητας της περιοχής και σε συνδυασμό με την κατάλληλη ασαφή διαμέρισή του μπορούμε να αποφανθούμε σε ποια από τις τρεις πιθανές κατηγορίες - Μαυρομάλλης, Περικαπνισμένος, Ασπρομάλλης - ανήκει η περιοχή. Η ίδια τεχνική χρησιμοποιείται και για την περιοχή της γενειάδας, δεδομένης της ύπαρξης αυτής.

Όσον αφορά στο σχήμα των μαλλιών, χρησιμοποιείται η τεχνική της ανάλυσης Blob και συγκεκριμένα μετρείται η εκκεντρότητα των εξαγόμενων αντικειμένων. Αν ο αγιογράφος θέλει να απεικονίσει ίσια μαλλιά, η εκκεντρότητα των γραμμών θα τείνει στη μονάδα, ενώ αν τα μαλλιά είναι σγουρά, η εκκεντρότητα των εξαγόμενων γραμμών θα τείνει στο μηδέν. (Σχήμα). Υπολογίζοντας, λοιπόν, το ποσοστό των

γραμμών που ανήκουν σε κάθε κατηγορία, μπορούμε να αποφανθούμε για το σχήμα των μαλλιών. Η ίδια τεχνική χρησιμοποιείται και για την εξαγωγή του κυματισμού της γενειάδας και της μορφής του μουστακίου (Πίνακες 4.5 και 4.4).

Η πυκνότητα των μαλλιών μπορεί να εξαχθεί με συνδυασμό του ανιχνευτή γραμμών Canny και ενός μορφολογικού φίλτρου, ενώ το ίδιο μπορεί να εφαρμοστεί και για την πυκνότητα της γενειάδας.

Χρησιμοποιώντας την τεχνική της ανάλυσης χρωμάτων των σαρκωμάτων (Ενότητα), μπορούμε να αποφανθούμε για την ύπαρξη ή μη της κόμης, της γενειάδας και του μουστακίου. Οι περιοχές που περιέχουν ως βασικά επίπεδα φωτεινότητας τιμές από το πεδίο τιμών του σαρκώματος είναι περιοχές που απεικονίζουν σάρκα. Άρα, έχοντας εξάγει τις περιοχές που περικλείουν την κόμη, τη γενειάδα και το μουστάκι, τις αναλύουμε ως προς το χρώμα τους και τις συγκρίνουμε με τις τιμές του σαρκώματος. Ειδικά για την κόμη, αυτή χωρίζεται σε τμήματα, έτσι ώστε να μπορεί να εξαχθεί πληροφορία για το αν η κεφαλή της μορφής είναι όλη φαλακρή ή εν μέρει.

Σύμφωνα με τον Διονύσιο εκ Φουρνά, η νεανική μορφή απεικονίζεται χρησιμοποιώντας ελάχιστες ρυτίδες στην περιοχή του μετώπου, αντίθετα με μια ώριμη μορφή που απεικονίζεται με πολλές ρυτίδες. Χρησιμοποιώντας το μετασχηματισμό Hough (Ενότητα), μετράμε το πλήθος των οριζοντίων και κατακόρυφων γραμμών που απεικονίζονται στην περιοχή του μετώπου και έτσι μπορούμε να πάρουμε την πληροφορία για την ηλικία της μορφής που απεικονίζεται.

Ο τρόπος που ζωγραφίζεται η περιοχή της παρειάς σχετίζεται με την ηλικία της μορφής. Προκειμένου να εξάγουμε πληροφορία σχετικά με την παρειά, χρησιμοποιούμε την παρειά που ανήκει στη μεγαλύτερη πλευρά του προσώπου, αφού αυτή καταλαμβάνει μεγαλύτερη επιφάνεια και περιέχει μεγαλύτερη πληροφορία. Ο αλγόριθμος χωρίζει την παρειά σε δύο νοητά οριζόντια τμήματα. Το πάνω τμήμα της παρειάς ζωγραφίζεται χρησιμοποιώντας σαρκώματα και φωτίσματα. Όταν η μορφή προς απεικόνιση είναι ώριμη, το κάτω τμήμα ζωγραφίζεται αρχίζοντας από σαρκώματα και, χρησιμοποιώντας προπλασμό, μπορεί να φτάσει μέχρι και απαλό σκούρεμα. Αντίθετα, όταν η μορφή είναι νεανική, το κάτω τμήμα ζωγραφίζεται χρησιμοποιώντας σαρκώματα και φωτίσματα. Η τεχνική της ανάλυσης χρωμάτων των σαρκωμάτων (Ενότητα) χρησιμοποιείται ξανά και εξάγουμε συμπεράσματα για την ηλικία της μορφής.

Το σχήμα και το μήκος της γενειάδας καθορίζονται με τη βοήθεια της προαναφερθείσας μετρικής - του ύψους της μύτης και την ασαφή διαμέριση των κατηγοριών του μήκους και του σχήματος. Για παράδειγμα, σύμφωνα με τον Διονύσιο, το πλάτος του γενιού ενός οξυγένη είναι περίπου ίσο με $1/2 \cdot H$, όταν αυτό μετρηθεί στο τελευταίο τμήμα του γενιού, ενώ ενός πλατυγένη είναι περίπου ίσο με $3 \cdot H$. Αντίστοιχα ισχύουν και για το μήκος μίας γενειάδας. Τέλος, η μορφή του γενιού μπορεί να ανιχνευθεί με τον υπολογισμό των χωρισμάτων που σχηματίζονται στο τελείωμα της γενειάδας. Υιοθετώντας τον ανιχνευτή Harris-Laplace [92] εξάγουμε το πλήθος των γωνιών του συγκεκριμένου τμήματος και με τη χρήση ασαφούς διαμέρισης μπορούμε να προβούμε σε συμπεράσματα για τη μορφή του γενιού.

4.4.3.1 Δημιουργία σώματος ισχυρισμών

Η συλλογή και αξιολόγηση των πληροφοριών που εξάγεται από κάθε σημασιολογική περιοχή είναι ικανή, με τη χρήση οντολογιών, να εξάγει γνώση, όπως θα αναλυθεί στο Κεφάλαιο . Η δημιουργία σώματος ισχυρισμών βασίζεται στη συλλογή όλων των

πληροφοριών που εξάγονται από τη μονάδα ανάλυσης Βυζαντινών εικόνων (Σχήμα 4.4).

Στο σώμα ισχυρισμών έχουμε μια μορφή που χωρίζεται σε δύο τμήματα (segments), στο πρόσωπο και στο σώμα και αυτά μπορούν να αναλυθούν περαιτέρω στα τμήματα που τα απαρτίζουν. Ας θεωρήσουμε s το πρόσωπο της μορφής και τα τμήματα που το απαρτίζουν ως εξής: s_1 τα μαλλιά, s_2 το μέτωπο, s_3 την παρειά της μεγάλης πλευράς, s_4 την περιοχή που μπορεί να απεικονίζεται μουστάκι, s_5 την περιοχή που μπορεί να απεικονίζεται γενειάδα.

Κάθε σώμα ισχυρισμών αρχικά αναφέρει ποια τμήματα, σύμφωνα με μονάδα επεξεργασίας εικόνας, έχουν ανιχνευθεί και τα δηλώνει (π.χ. $s : Face$, $s_1 : Hair$). Έπειτα για κάθε τμήμα, καταγράφονται τα ανιχνευμένα χαρακτηριστικά μαζί με τις τιμές τους. Για παράδειγμα, από την ανάλυση της εικόνας του Χριστού, ως προς την κατηγοριοποίηση του μαλλιού και συγκεκριμένα ως προς το χαρακτηριστικό χρώμα, εξάγονται τα εξής:

$$s_1 : BlackHair \geq 0.81, s_1 : GrayHair \geq 0.29 \text{ και } s_1 : WhiteHair \geq 0.19.$$

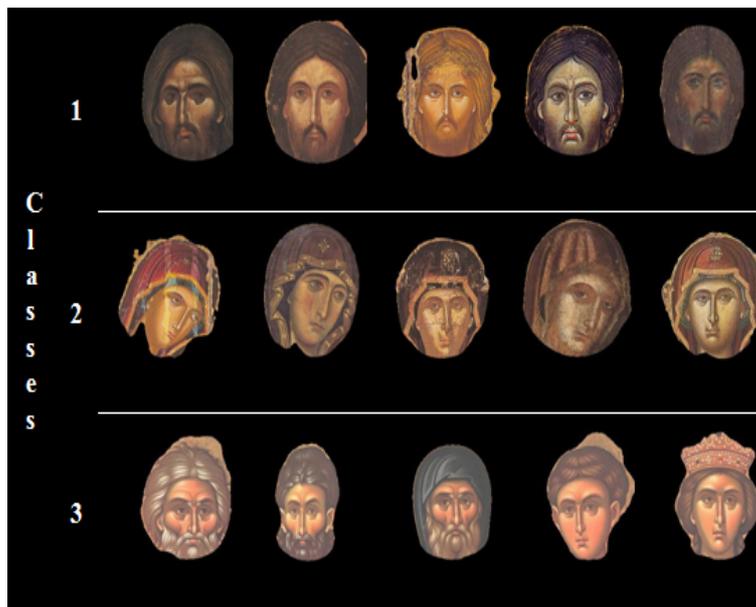
Αντίστοιχα αποτελέσματα εξάγονται από τη μονάδα επεξεργασίας βυζαντινής εικόνας για όλα τα σημασιολογικά τμήματα του προσώπου και η συλλογή αυτών σε ένα κοινό αρχείο είναι το σώμα ισχυρισμών της Βυζαντινής εικόνας. Το σώμα ισχυρισμών δίνεται ως είσοδος στη μονάδα αναπαράστασης γνώσης και συλλογιστικής, η οποία μπορεί να αποφανθεί αυτόματα για το ποια μορφή απεικονίζεται στην εικόνα που μόλις αναλύσαμε, όπως θα δούμε στην Ενότητα 5.4.1.

4.4.3.2 Πειραματικά αποτελέσματα

Το σύνολο των δεδομένων για τα πειράματά μας αποτελείται από 2000 βυζαντινές εικόνες που ανήκουν σε τρεις κατηγορίες: Ιησούς, Παναγία και άλλοι Άγιοι. Στο Σχήμα 4.12 απεικονίζονται πέντε κεφαλές από κάθε κατηγορία, οι οποίες έχουν εξαχθεί από Βυζαντινές εικόνες με τη μέθοδο που αναφέρθηκε στην Ενότητα 4.4.1. Αρχικά, το σύνολο δεδομένων διαιρείται με τυχαίο τρόπο σε δύο μέρη, κρατώντας τις αναλογίες κάθε κατηγορίας ίσες με αυτές που παρουσιάζονται σε ολόκληρο το σύνολο δεδομένων.

Το πρώτο μέρος αποτελείται από 800 Βυζαντινές εικόνες και θεωρείται ως το σύνολο κατάρτισης D , ενώ το υπόλοιπο (1200 Βυζαντινές εικόνες) συνθέτει το σύνολο ελέγχου UD . Το σύνολο δεδομένων εκπαίδευσης D χωρίζεται περαιτέρω σε δύο ίσα υποσύνολα, PD με τα θετικά και ND με τα αρνητικά δείγματα. Το πείραμα μας εστιάζει στη σωστή κατηγοριοποίηση εικόνων του Χριστού και της Παναγίας σε αντίθεση με όλους τους άλλους Αγίους. Για να αξιολογήσουμε την ακρίβεια της ταξινόμησης υιοθετήσαμε τη μετρική F-score, ένα συνδυασμό ακρίβειας και ανάκληση.

Οι τέσσερις αλγόριθμοι ταξινόμησης, το NB, το FNB, το SVM και το πολυστρωματικό Perceptron με μάθηση οπίσθιας διάδοσης (NNBP) εκπαιδεύονται, ο κάθε ένας χωριστά, χρησιμοποιώντας τα υποσύνολα PD και ND . Το χαρακτηριστικό διάνυσμα, που περιέχει τα χαρακτηριστικά από τους 4 περιγραφείς MPEG-7 και τις ροπές H_u , χρησιμοποιείται για να εκπαιδεύσει τα πέντε νευρωνικά δίκτυα, κάθε ένα από τα οποία αποτελείται από δύο κρυφά επίπεδα, που ποικίλλουν σε αριθμό νευρώνων (3-7 για το πρώτο επίπεδο και 1-3 για το δεύτερο επίπεδο). Η διάσταση των χαρακτηριστικών διανυσμάτων έχει μειωθεί για τους ταξινομητές ως εξής : 10 για το NB, 10 για το FNB, 9 για το SVM και 7 για το πολυστρωματικό Perceptron με μάθηση οπίσθιας διάδοσης (NNBP), σαν αποτέλεσμα της ανάλυσης κύριων



Σχήμα 4.12: Κατηγορίες κεφαλής (1) Χριστός, (2) Παναγία και (3) Άλλοι Άγιοι

συνιστωσών. Η όλη διαδικασία επαναλήφθηκε 50 φορές με σκοπό τη μείωση των διακυμάνσεων.

Για την αξιολόγηση της ορθότητας της ταξινόμησης χρησιμοποιήθηκαν τα εξής μέτρα: ακρίβεια, ορθότητα, ανάκληση και το F-score.

Ως ακρίβεια (accuracy) ορίζουμε την αναλογία όλων των προβλέψεων που ήταν σωστές με το πλήθος των προβλέψεων: $accuracy = \frac{\text{πλήθος σωστών προβλέψεων}}{\text{πλήθος προβλέψεων}}$.

Ως ορθότητα (precision - degree of soundness) κάθε κλάσης ορίζουμε τη δεσμευμένη πιθανότητα να ταυτίζεται η κλάση που προβλέπει ένας ταξινομητής για ένα στιγμιότυπο με την πραγματική του κλάση: $precision_c = \frac{\text{πλήθος σωστών προβλέψεων κλάσης } c}{\text{πλήθος προβλέψεων κλάσης } c}$ και για όλες τις κλάσεις: $precision = \sum_i precision_{c_i} \cdot P(c_i)$, όπου $P(c_i)$ η πιθανότητα να ταξινομηθεί μια εικόνα στην κλάση i .

Ως ανάκληση (recall - degree of completeness) για κάθε κλάση ορίζουμε τη δεσμευμένη πιθανότητα αν ένα στιγμιότυπο ανήκει σε μια κλάση, έστω c , αυτή να αναγνωριστεί σωστά από τον ταξινομητή: $recall_c = \frac{\text{πλήθος σωστών προβλέψεων κλάσης } c}{\text{πλήθος δεδομένων κλάσης } c}$ και για όλες τις κλάσεις: $recall = \sum_i recall_{c_i} \cdot P(c_i)$ όπου $P(c_i)$ η πιθανότητα να ανήκει η εικόνα στην κλάση i .

Στην πράξη οι δύο παραπάνω μετρικές δεν μπορούν να εκτιμηθούν χωριστά, καθώς παρέχουν μια αλληλοσυμπληρούμενη εικόνα της αποτελεσματικότητας ενός ταξινομητή. Ένα μέτρο που τα συνδυάζει είναι η συνάρτηση F (F-score), που ορίζεται ως εξής:

$$F = \frac{2 \cdot precision \cdot recall}{precision + recall}$$

Στον Πίνακα 4.7 παρουσιάζεται η μέση τιμή και η διακύμανση του F-score για κάθε αλγόριθμο. Ο ταξινομητής NNBP για ένα συγκεκριμένο υποσύνολο UD παρέχει καλύτερη ταξινόμηση από τους άλλους ταξινομητές.

Η τελική επιλογή του αποδοτικότερου ταξινομητή μεταξύ των πέντε γίνεται βάσει του F-score και η σωστή αναγνώριση επιτυγχάνεται κατά 90% για όλες τις

Πίνακας 4.7: Μέση τιμή και διακύμανση του συνόλου δεδομένων UD

	NB	FNB	SVM	NNBP
Μέση Τιμή	0.347	0.541	0.608	0.895
Διακύμανση	0.020	0.012	0.086	0.038

περιπτώσεις, κάτι που αποδίδεται στην επιτυχία του αλγορίθμου. Στο 10% των λανθασμένων αποτελεσμάτων περιλαμβάνονται οι εξής κατηγορίες:

- αγιογραφίες με μορφές που μοιάζουν στον Χριστό όπως ο Άγιος Αρτέμιος, και κατατάσσονται στην κατηγορία Χριστός,
- αγένειοι άνδρες, συνήθως νεαρής ηλικίας (π.χ. Ο Άγιος Δημήτριος) που το σύστημα τους κατατάσσει στην κατηγορία Παναγία,
- αγιογραφίες με τη μορφή του Χριστού ή της Παναγίας, οι οποίες είναι χαμηλής ανάλυσης, θολές ή θορυβώδεις και κατατάσσονται στην κατηγορία Άλλοι Άγιοι

Στη μέθοδο που προτάθηκε μπορούμε να επισημάνουμε μερικούς παράγοντες που φαίνονται να διαδραματίζουν σημαντικό ρόλο στην ακρίβεια των αποτελεσμάτων του προτεινόμενου αλγορίθμου:

- Η σχεδόν τέλεια ομοιότητα μερικών μορφών που παρουσιάζεται στις κεφαλές των αγιογραφιών, μας οδήγησε στη χρήση του περιεχομένου εικόνων, την εκμετάλλευση δηλαδή κάθε πληροφορίας που μπορεί να εξαχθεί από την κεφαλή, από τα ενδύματα, τα αντικείμενα που συνοδεύουν τη μορφή. Με αυτόν τον τρόπο θα οδηγηθούμε σε αναγνώριση της κάθε μορφής,
- Ο θόρυβος έχει επιπτώσεις στην ακρίβεια ταξινόμησης,
- Η ακρίβεια του αλγορίθμου εξαγωγής της κεφαλής είναι ένας κρίσιμος παράγοντας προς τη σωστή ταξινόμηση.

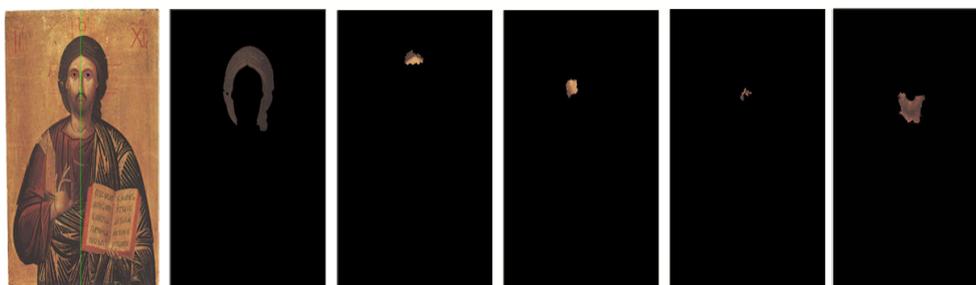
Έπειτα από την πρώτη μέθοδο που χρησιμοποιήσαμε, προχωρήσαμε στην υλοποίηση της ιδέας της ανίχνευσης του προσώπου και των χαρακτηριστικών του και τη χρήση της γνώσης που προσφέρει το εγχειρίδιο του Διονυσίου εκ Φουρνά με σκοπό την αυτόματη κατηγοριοποίηση των εικόνων. Έτσι σχεδιάσαμε και υλοποιήσαμε το σύστημα του Σχήματος 4.3. Ουσιαστικά εφαρμόζουμε τις τεχνικές που αναλύθηκαν προηγουμένως στο υποσύστημα επεξεργασίας εικόνας (Ενότητες 4.4.2.1, 4.4.2.2, 4.4.2.3, 4.4.2.3.1) και τα πειραματικά τους αποτελέσματα παρουσιάζονται σε επόμενο κεφάλαιο. Για την αξιολόγηση του συστήματος, χρησιμοποιήθηκε ένα σύνολο από 2000 Βυζαντινές ψηφιακές εικόνες χρονολογημένες από το 13ο αιώνα, με ένα σύνολο από διαφορετικές μορφές Αγίων. Αυτό το σύνολο προσφέρθηκε από το Ίδρυμα του Όρους Σινά και αποτελείται από πενήντα διαφορετικές μορφές Αγίων (Χριστός, Παναγία, Αγία Αικατερίνη, Άγιος Πέτρος, Άγιος Παύλος, κ.τ.λ.).

Στο Σχήμα 4.13, εμφανίζονται τα αποτελέσματα από τον εντοπισμό της κεφαλής, των ματιών, της μύτης και του άξονα αντικίνησης για τέσσερις χαρακτηριστικές εικόνες. Η απόδοση του αλγορίθμου εντοπισμού προσώπου (Ενότητα 4.4.2.1.2), αφού πρώτα αυτός εκπαιδεύτηκε με βυζαντινές εικόνες, είναι αρκετά ικανοποιητική (της τάξης του 90%). Το μεγαλύτερο πρόβλημα εντοπίζεται σε κατεστραμμένες εικόνες, στις οποίες ένα τμήμα του προσώπου λείπει. Αφού εντοπιστεί σωστά η

Κεφάλαιο 4. Ανάλυση και τεκμηρίωση πολυμεσικού περιεχομένου



Σχήμα 4.13: Εντοπισμός προσώπου, ματιών, μύτης και άξονα αντικίνησης



Σχήμα 4.14: Εντοπισμός σημασιολογικών τμημάτων προσώπου

περιοχή του προσώπου, ο αλγόριθμος ανίχνευσης των ματιών, καθώς επίσης και ο αλγόριθμος εντοπισμού μύτης και άξονα αντικίνησης (Ενότητα 4.4.2.2) λειτουργούν σχεδόν τέλεια, σε όσες εικόνες έχει ανιχνευθεί σωστά το πρόσωπο.

Η ανάλυση των χρωμάτων των σαρκωμάτων (Ενότητα 4.4.3) καθορίζει τις φωτεινότητες των βασικών χρωμάτων που ο αγιογράφος χρησιμοποίησε στην προς ανάλυση Βυζαντινή εικόνα. Η περιοχή του προσώπου, όπως αναφέρθηκε, περιέχει πέντε σημασιολογικά τμήματα: τα μαλλιά, το μέτωπο, την παρειά, την περιοχή του μουσταχίου και την περιοχή του γενειού. Χρησιμοποιώντας το ύψος της μύτης που υπολογίστηκε από τον αλγόριθμο και λαμβάνοντας υπόψη τους κανόνες του εγχειριδίου, μπορούμε να καθορίσουμε, με ακρίβεια, τις θέσεις των εικονοστοιχείων του φόντου και του κάθε σημασιολογικού τμήματος που θέλουμε να εξετάσουμε. Αυτές οι θέσεις αποτελούν την είσοδο στον αλγόριθμο κατάτμησης (Ενότητα 4.4.2.3.1). Η κατάτμηση κάθε σημασιολογικού τμήματος επιτυγχάνει κατά 96% σωστή εξαγωγή των τμημάτων του προσώπου.

Για κάθε εξαγόμενο τμήμα εξετάζουμε περαιτέρω τα χαρακτηριστικά του, όπως αυτά περιγράφηκαν στους Πίνακες 4.1, 4.2, 4.3, 4.4 και 4.5, χρησιμοποιώντας τις τεχνικές που αναφέρθηκαν στην Ενότητα 4.4.2.3. Έπειτα, με τη βοήθεια της ασαφούς διαμέρισης ερμηνεύεται κάθε χαρακτηριστικό, αποδίδοντας σε αυτό ένα βαθμό συμμετοχής για κάθε μία από τις κλάσεις του. Ο μέσος χρόνος απόκρισης του προτεινόμενου συστήματος είναι 15 sec σε ένα κοινό PC (CPU Intel Core 2 Duo 2.13GHz και RAM 2GB). Η συλλογή όλων των ασαφών ερμηνειών

των χαρακτηριστικών των σημασιολογικών τμημάτων του προσώπου αποτελεί την περιγραφή της εικόνας, δηλαδή το σώμα ισχυρισμών (A-Box), όπως αυτό εξήχθη μέσω του συστήματος ανάλυσης εικόνας. Για παράδειγμα, το A-box που δημιουργήθηκε από το σύστημα για τα τμήματα της βυζαντινής εικόνας του Χριστού που απεικονίζονται στο Σχήμα 4.14, παρουσιάζεται στον Πίνακα 4.8.

Πίνακας 4.8: Σώμα Ισχυρισμών όπως αυτό εξήχθη αυτόματα από τη μονάδα επεξεργασίας Βυζαντινής εικόνας

(related j018Seg7Face j018Seg1Hair hasPart)
(instance j018Seg1Hair BlackHair >= 0.90)
(instance j018Seg1Hair GrayHair >= 0.17)
(instance j018Seg1Hair WhiteHair >= 0.10)
(instance j018Seg1Hair HairyHair >= 0.68)
(instance j018Seg1Hair SparseHair >= 0.32)
(instance j018Seg1Hair BelowEarsHair >= 1.00)
(instance j018Seg1Hair ThinHairOnTheLeft >= 1.00)
(instance j018Seg1Hair StraightHair >= 0.68)
(instance j018Seg1Hair CurlyHair >= 0.22)
(instance j018Seg1Hair WildHair >= 0.11)
(related j018Seg7Face j018Seg4Forehead hasPart)
(instance j018Seg4Forehead YoungForehead >= 1.00)
(related j018Seg7Face j018Seg2Cheek hasPart)
(instance j018Seg2Cheek YoungCheek >= 0.66)
(instance j018Seg2Cheek OldCheek >= 0.34)
(related j018Seg7Face j018Seg5Moustache hasPart)
(instance j018Seg5Moustache DarkMoustache >= 0.75)
(instance j018Seg5Moustache FairMoustache >= 0.25)
(instance j018Seg5Moustache StraightMoustache >= 0.78)
(instance j018Seg5Moustache CurlyMoustache >= 0.22)
(related j018Seg7Face j018Seg3Beard hasPart)
(instance j018Seg3Beard ShortBeard >= 1.00)
(instance j018Seg3Beard CircleLikedBeard >= 1.00)
(instance j018Seg3Beard HairyBeard >= 0.06)
(instance j018Seg3Beard SparseBeard >= 0.94)
(instance j018Seg3Beard StraightBeard >= 0.89)
(instance j018Seg3Beard CurlyBeard >= 0.11)
(instance j018Seg3Beard HairyChin >= 0.01)
(instance j018Seg3Beard SparseBeardChin >= 0.89)
(instance j018Seg3Beard DarkBeard >= 0.65)
(instance j018Seg3Beard FairBeard >= 0.35)

Αυτό το ασαφές σώμα ισχυρισμών μαζί με το ασαφές σώμα ορολογίας, όπως θα δούμε στο Κεφάλαιο 5 συνδυάζονται με σκοπό την εξαγωγή γνώσης. Το ασαφές σώμα ισχυρισμών συνοδεύει την εικόνα και αποτελεί τα μεταδεδομένα της, συμβάλλοντας στη γρήγορη ταξινόμησή της καθώς και στην εξόρυξή της από βάση πολυμεσικών δεδομένων.

Σε αυτό το πλαίσιο μπορούμε να μιλήσουμε για ένα ολοκληρωμένο σύστημα εξαγωγής γνώσης από τα πρόσωπα που απεικονίζονται σε Βυζαντινές εικόνες. Το προτεινόμενο σύστημα μπορεί να υποστηρίξει αγιογράφους, βυζαντινούς αναλυτές, ιστορικούς, αρχαιολόγους και ερευνητές με σκοπό να αναλύσουν καλύτερα, να ερμηνεύσουν και να καταλάβουν τα Βυζαντινά έργα τέχνης, τους συσχετισμούς και την κοινή εξέλιξη των χαρακτηριστικών, των χρωμάτων και του ύφους της αγιογραφίας καθώς και πολλών άλλων χαρακτηριστικών γνωρισμάτων. Εδώ πρέπει να αναφέρουμε ότι μέχρι στιγμής οι διαδικασίες για ταξινόμηση των αγιογραφιών βασίζονται στις ικανότητες παρατήρησης και αντίληψης των ιστορικών της τέχνης.

Κατά συνέπεια, η έρευνα με σκοπό την ανάλυση σχεδίων (pattern) και την ταξινόμηση των περιοχών στις Βυζαντινές εικόνες θα συμβάλλει θετικά στην ανάπτυξη έμπειρων συστημάτων για την αναγνώριση των μορφών και των παραλλαγών τους στη Βυζαντινή τέχνη.

Κεφάλαιο 5

Αναπαράσταση Γνώσης και Συστήματα Εξ Αποστάσεως Εκπαίδευσης

5.1 Εισαγωγή

Η αναπαράσταση της γνώσης είναι το σύνολο των συντακτικών και σημασιολογικών παραδοχών οι οποίες καθιστούν δυνατή την περιγραφή ενός κόσμου. Τα συστήματα γνώσης, τα οποία παρουσιάζουν μεγάλη ακμή στις μέρες μας, χρησιμοποιούν αυτήν την αναπαράσταση με σκοπό την εξαγωγή συμπερασμάτων για τον κόσμο που περιγράφουν. Πυρήνας του συστήματος γνώσης είναι η μηχανή συλλογιστικής στην οποία ενσωματώνεται τόσο συγκεκριμένη όσο και αφηρημένη γνώση με στόχο την εξαγωγή νέας γνώσης.

Στόχος της δικής μας εργασίας είναι η παροχή τόσο συγκεκριμένης όσο και αφηρημένης γνώσης από το πεδίο της Βυζαντινής αγιογραφίας σε μία μηχανή συλλογιστικής. Συγκεκριμένα, η συγκεκριμένη γνώση παρουσιάζεται με τη μορφή οντολογίας περιγραφής των μορφών Αγίων, αξιοποιώντας τη γνώση που παρέχουν οι τεχνικές αγιογράφησης των Βυζαντινών εικόνων όπως αυτές περιγράφονται στο εγχειρίδιο του Διονυσίου εκ Φουρνά.

Απώτερος σκοπός του εγχειρήματός μας είναι η δημιουργία νέας γνώσης, με τη βοήθεια της μηχανής συλλογιστικής, συνδυάζοντας την αφηρημένη γνώση, όπως αυτή προέκυψε από τον αλγόριθμο επεξεργασίας Βυζαντινών εικόνων (Ενότητα 4.4.2) με τη συγκεκριμένη γνώση.

Αυτή η νέα γνώση μπορεί να αξιοποιηθεί περαιτέρω και να θεωρηθεί ως μεταδεδομένο της εικόνας το οποίο θα αποτελεί την τεκμηρίωση της εικόνας και θα μπορεί να οδηγεί σε γρήγορη κατηγοριοποίηση και ανάκτηση της μέσα σε ένα ηλεκτρονικό σύστημα όπως ένα σύστημα μαθησιακής τεχνολογίας.

Σε αυτό το κεφάλαιο θα παρουσιάσουμε αρχικά το τι είναι μια οντολογία και πως μπορεί αυτή να χρησιμοποιηθεί τόσο από ένα σύστημα μαθησιακής τεχνολογία όσο και από το σημασιολογικό ιστό. Στη συνέχεια θα αναφερθούμε αναλυτικά στις τυπικές μεθόδους αναπαράστασης γνώσης και συλλογιστικής, και θα δείξουμε πως η έννοια της ασάφειας μπορεί να εφαρμοστεί στις περιγραφικές λογικές και σε συνδυασμό με μηχανισμούς εξαγωγής συμπερασμάτων να εξάγει νέα γνώση με μεγάλη συνέπεια. Τέλος θα παρουσιάσουμε τη μονάδα αναπαράστασης γνώσης και συλλογιστική (Σχήμα 4.5) η οποία αποτελείται από το σώμα ασαφών ρόλων, το σώμα ασαφούς ορολογίας και από τη μηχανή συλλογισμού. Το πεδίο των συστημάτων γνώσης στο χώρο της βυζαντινής αγιογραφίας είναι ακόμα πολύ ανοικτό και επιτρέπει

πλήθος εναλλακτικών λύσεων και προσεγγίσεων. Η εστίαση στους κανόνες του Διονυσίου εκ Φουρνά πιστεύουμε ότι αποτελεί αρκετά καλή βάση για την εξέλιξη και προσαρμογή του παρόντος.

5.2 Οντολογίες και Σημασιολογικός Ιστός

Η λέξη οντολογία ετυμολογικά προέρχεται από τη λέξη *ον* και τη λέξη *λέγω* και είναι ο κλάδος ο οποίος ασχολείται με την ουσία των όντων δηλαδή την ανάλυση των όντων με σκοπό την εύρεση της βαθύτερης έννοιας που τα συνδέει. Με άλλα λόγια η οντολογία είναι ένας σαφής εννοιολογικός προσδιορισμός με σκοπό την κατανόηση ενός πεδίου ενδιαφέροντος. Ο στόχος μιας οντολογίας είναι η χαρτογράφηση ενός πεδίου ενδιαφέροντος έτσι ώστε να καθοριστεί ένα κοινό λεξιλόγιο των βασικών εννοιών του πεδίου, καθώς και οι ιδιότητες και οι περιορισμοί αυτών των εννοιών. Η επίτευξη αυτού του στόχου διασφαλίζει την επικοινωνία μεταξύ συστημάτων που σκοπεύουν να χρησιμοποιήσουν κοινή γλώσσα, δηλαδή κοινή οντολογία.

Η οντολογία λοιπόν είναι μία σαφής περιγραφή ενός πεδίου και περιέχει έννοιες καθώς και τις ιδιότητες και τους περιορισμούς τους. Ο στόχος μιας οντολογίας είναι ο καθορισμός ενός κοινού λεξιλογίου ώστε να είναι εφικτή η κατανόησή του από διαφορετικά συστήματα καθώς και η ενοποίηση ετερογενούς πληροφορίας. Με τη χρήση οντολογιών μπορούμε να αναπαραστήσουμε και να μοντελοποιήσουμε τη γνώση που ενσωματώνει το σύστημα.

Η χρήση των οντολογιών στο διαδίκτυο κρίνεται όλο και πιο απαραίτητη, μιας και οι πληροφορίες που αυτό διαθέτει αυξάνουν ανεξέλεγκτα. Για παράδειγμα, αν κάποιος θελήσει να ανακτήσει μια πληροφορία από το διαδίκτυο θα πρέπει να την αναζητήσει κάνοντας χρήση κάποιας μηχανής αναζήτησης. Το βασικό πρόβλημα που θα αντιμετωπίσει είναι το ότι η σύνδεση που μπορεί να επιτευχθεί μεταξύ της σημασίας της πληροφορίας που αναζητά και της καταχωρημένης στο διαδίκτυο πληροφορίας περιορίζεται στην αξιοπιστία της μηχανής αναζήτησης και στα μεταδεδομένα που ο δημιουργός του εγγράφου έχει δηλώσει. Έτσι πολλές φορές παρατηρείται η εύρεση άσχετων πληροφοριών σε σχέση με τις πληροφορίες που ζητήθηκαν.

Ένα σενάριο που όπως διαφαίνεται ξεπερνά αυτό το εμπόδιο, είναι να μπορούν οι υπολογιστές να έχουν πρόσβαση σε δομημένη πληροφορία (μία παγκόσμια κοινή κατανομημένη βάση δεδομένων) και σε ένα σύνολο από συμπερασματικούς κανόνες. Με αυτό το τρόπο θα είναι δυνατή η ανάπτυξη συστημάτων που αυτόματα θα συνδυάζουν τα πεδία των ερωτημάτων του χρήστη και θα παράγουν συμπεράσματα τα οποία θα οδηγούν το χρήστη στο επιθυμητό για αυτόν αποτέλεσμα, δηλαδή στην ακριβή πληροφόρησή του.

Σε αυτό το πλαίσιο έχουν προταθεί νέα διαδικτυακά πρότυπα με σκοπό να κάνουν κατανοητή τη σημασιολογία της πληροφορίας από τους υπολογιστές. Εκατομμύρια μικρές μηχανές συμπερασματολογίας παρέχουν αυτόματη επίτευξη εργασιών έχοντας πρόσβαση στη πληροφορία. Ο Berners-Lee και τους συνεργάτες του το 2001 παρουσίασαν μια σημαντική εφαρμογή των οντολογιών [22], τον σημασιολογικό ιστό [23].

Ως σημασιολογικός ιστός ορίζεται το δίκτυο στο οποίο μπορούν οι υπολογιστές να έχουν πρόσβαση σε μία συλλογή πληροφοριών και σε ένα σύνολο από συμπερασματικούς κανόνες με σκοπό την αυτόματη παραγωγή λογικής. Ο σημασιολογικός ιστός δεν είναι ένας ξεχωριστός ιστός αλλά μπορεί κάλλιστα να θεωρηθεί ως μια επέκταση του web ιστού με σκοπό την έξυπνη ενσωμάτωση

συστημάτων και τη λειτουργική συνεργασία τους. Ο Tim Berners-Lee, που επινόησε τον Παγκόσμιο Ιστό το 1989, αναφέρει ότι ο σημασιολογικός ιστός είναι μια επέκταση του σημερινού ιστού όπου η πληροφορία έχει καλά καθορισμένο νόημα, καθιστώντας τη συνεργασία μεταξύ ανθρώπων και υπολογιστών πιο αποτελεσματική [22].

Ο σημασιολογικός ιστός έχει αρχίσει και δημιουργείται με την εξαγωγή γνώσης από την υπάρχουσα γνώση. Για την επίτευξη αυτής της εξαγωγής χρησιμοποιούνται μηχανές συμπερασματολογίας οι οποίες βασίζονται σε αλγόριθμους γενικευμένης λογικής (General logic based inference engines). Στόχος των μηχανών συμπερασματολογίας είναι η παραγωγή λογικής χρησιμοποιώντας τη σημασιολογική πληροφορία που συσχετίζεται με μεταδεδομένα που συνοδεύουν την πληροφορία που διακινείται μέσω διαδικτύου.

Η επέκταση του παγκόσμιου ιστού σε σημασιολογικό ιστό έχει ως σκοπό την αυτοματοποίηση των λειτουργιών και των εφαρμογών του διαδικτύου, όπως είναι οι μηχανές αναζήτησης και οι πράκτορες (agents). Η αυτοματοποίηση αυτή μπορεί να επιτευχθεί μόνον εφόσον η γνώση και η πληροφορία που υπάρχει αποθηκευμένη και δημοσιευμένη αυτή τη στιγμή στο σημερινό Παγκόσμιο Ιστό αποκτήσει τυπικό νόημα (formal meaning) και σημασιολογία (semantics) και δομηθεί με ένα τέτοιο τρόπο ώστε να γίνεται κατανοητή από τις μηχανές που την επεξεργάζονται (machine understandable). Από τη στιγμή που η πληροφορία θα είναι δομημένη με έναν σημασιολογικά πλούσιο τρόπο θα ενισχύεται ο διαμοιρασμός (sharing) και η επαναχρησιμοποίησή της (reusability) επιτυγχάνοντας τη διαλειτουργικότητα (interoperability) και συνδεσιμότητα (interconnectivity) ετερογενών (heterogeneous) συστημάτων και εφαρμογών.

5.3 Τυπικές μέθοδοι αναπαράστασης γνώσης και συλλογιστική

Ο άνθρωπος στην προσπάθειά του να επικοινωνήσει με τους άλλους επινόησε την γλώσσα, δίνοντας κατά αυτό τον τρόπο νόημα στις εκφράσεις του. Στην εργασία [3] αναφέρεται ότι "Μελετώντας τους μηχανισμούς και τις γραμματικές δομές της γλώσσας είναι στην πραγματικότητα σαν να μελετάμε το διανοητικό αγώνα του ανθρώπου να εκφράσει μέσα από ένα σύστημα (τη γραμματική που έχει επινοήσει, συνεχώς καλλιεργεί και μεταβάλλει) το πολύ σύνθετο πλέγμα λεπτών νοητικών σχέσεων που συνιστά τον ανθρώπινο λόγο ως έκφραση του βασικού προνομίου του ανθρώπου, του πνεύματος". Θέλοντας ο άνθρωπος να εμψυχήσει αυτό το βασικό του προνόμιο σε μία μηχανή χρειάζεται να την εφοδιάσει με τον απαραίτητο μηχανισμό που θα της προσδώσει αυτήν την ικανότητα. Δηλαδή, ο άνθρωπος πρέπει να επινοήσει τρόπους και μηχανισμούς αναπαράστασης της γνώσης που διαθέτει, έτσι ώστε να μπορέσει να δημιουργήσει μία ευφυή μηχανή δηλαδή μια μηχανή ικανή να προσδώσει συλλογιστική.

Σε αυτή την κατεύθυνση κινείται το ερευνητικό πεδίο της τεχνητής νοημοσύνης επιχειρώντας όχι μόνο να κατανοήσει αλλά και να κατασκευάσει νοήμονες οντότητες. Η αναπαράσταση της γνώσης είναι μια περιοχή της τεχνητής νοημοσύνης που ασχολείται με το πώς τυπικά "σκεφτόμαστε" προσπαθώντας να ορίσει ένα σύστημα συμβόλων τέτοιο ώστε να είναι ικανό να αντιπροσωπεύσει "μια περιοχή του λόγου" (ένα επίχειρημα). Η αναπαράσταση γνώσης ουσιαστικά είναι ένα σύνολο συντακτικών και σημασιολογικών παραδοχών, οι οποίες καθιστούν δυνατή τη

περιγραφή ενός κόσμου. Ο όρος σύστημα γνώσης ή σύστημα βασισμένο στη γνώση (knowledge-based system) αποδίδεται σε ένα ευφυές σύστημα λογισμικού στο οποίο η γνώση αναπαρίσταται χρησιμοποιώντας τυπικές μεθόδους αναπαράστασης γνώσης και συλλογιστικής.

Προκειμένου η γνώση και η πληροφορία να περιγραφεί με έναν τυπικό (formal) τρόπο ο οποίος θα δηλώνει τη σημασία της, χρησιμοποιήθηκαν διάφορες γλώσσες αναπαράστασης γνώσης (knowledge representational languages), όπως είναι η Προτασιακή Λογική (Propositional Logic) [91] ή η Κατηγορηματική Λογική Πρώτης Τάξης (First-Order Predicate Logic) [91], οι οποίες έχουν αναπτυχθεί από μαθηματικούς, και άλλες όπως τα Σημασιολογικά Δίκτυα (Semantic Networks)[108] και οι Περιγραφικές Λογικές (ΠΛ) (Description Logics ή DLs) [16], οι οποίες έχουν αναπτυχθεί από επιστήμονες του κλάδου της πληροφορικής. Οι γλώσσες αυτές αποτελούνται από ένα αλφάβητο (alphabet), ένα συντακτικό (syntax) και μια σημασιολογία (semantic). Επιπλέον είναι εφοδιασμένες με ένα μηχανισμό εξαγωγής συμπερασμάτων (reasoning algorithm), ο οποίος χρησιμοποιώντας την υπάρχουσα γνώση να δύναται να παράγει νέα γνώση.

Η συλλογιστική (reasoning) ουσιαστικά είναι η μέθοδος που συνδυάζει τμήματα υπάρχουσας γνώσης με τέτοιο τρόπο ώστε να εξάγονται αποτελέσματα τα οποία θα πλησιάζουν σε λογική αυτά της ανθρώπινης συλλογιστικής και σκέψης. Η εξαγωγή συμπερασμάτων εξαρτάται όχι μόνο από την συλλογιστική αλλά και από τον τρόπο που έχει δομηθεί και κωδικοποιηθεί η γνώση χρησιμοποιώντας έναν τυπικό τρόπο αναπαράστασης.

Αρχικά, από την εποχή του Αριστοτέλη, έγινε προσπάθεια να καταγραφεί η ανθρώπινη γνώση με έναν τυπικό τρόπο αποφεύγοντας τη χρήση της φυσικής γλώσσας (η οποία δεν μπορεί να αναπαρασταθεί εύκολα με μαθηματικό τρόπο) και κατασκευάζοντας γλώσσες αναπαράστασης γνώσης. Οι γλώσσες αυτές περιέχουν ένα αλφάβητο, ένα συντακτικό και μια σημασιολογία ακριβώς όπως και η φυσική γλώσσα. Επιπροσθέτως, οι γλώσσες αναπαράστασης γνώσης περιλαμβάνουν ένα μηχανισμό εξαγωγής συμπερασμάτων (reasoning algorithm). Ο μηχανισμός εξαγωγής συμπερασμάτων χρησιμοποιώντας μια γλώσσα αναπαράστασης γνώσης περιγράφει με κανόνες την αρχική γνώση του πεδίου και μπορεί να εξάγει νέα γνώση και νέα συμπεράσματα από τα αρχικά γεγονότα. Όπως γίνεται αντιληπτό ο μηχανισμός αυτός είναι ιδιαίτερα σημαντικός στους επιστήμονες του κλάδου των υπολογιστών, καθώς παρέχει ένα νέο τρόπο για την εκμάθηση νέας γνώσης από την ήδη υπάρχουσα.

Η περιγραφική λογική (description logics) έχει αποδειχθεί κατάλληλη σε πολλές εφαρμογές αφού έχει πολύ μεγάλη εκφραστική δύναμη (expressive power) και διατηρεί αρκετά χαμηλά την υπολογιστική πολυπλοκότητα. Εν τούτοις, η εγγενής αβεβαιότητα που υπάρχει στις πραγματικές εφαρμογές επιβάλλει την επέκταση των σύγχρονων φορμαλισμών λογικής ώστε να αναπαριστούν ασαφείς έννοιες (fuzzy concepts) και να προσαρμόζουν τις υπάρχουσες μηχανές συμπερασματολογίας (reasoners) ώστε να διαχειρίζονται τέτοιου είδους γνώση.

5.3.1 Γλώσσες αναπαράστασης γνώσης

Στην προσπάθεια καθορισμού τυπικής σημασιολογίας στη γλώσσα των σημασιολογικών δικτύων όπως προαναφέρθηκε προέκυψε η γλώσσα των Περιγραφικών Λογικών (ΠΛ) [15]. Σκοπός των ΠΛ είναι η δημιουργία μιας

δομημένης, τυπικής και κατανοητής περιγραφή της γνώσης ενός πεδίου εφαρμογής η οποία να παρέχει πλούσια εκφραστικότητα, να είναι αποφάνσιμη καθώς και η σημασιολογία να είναι μοντελοθεωρητική. Σε αυτό το πλαίσιο έχουν προταθεί διάφορες ΠΛ και γλώσσες αναπαράστασης όπως θα δούμε στη συνέχεια.

5.3.1.1 Περιγραφικές Λογικές και SHIN

Οι Περιγραφικές Λογικές (ΠΛ) (Description Logics - DLs) [15] αποτελούν μια οικογένεια γλωσσών αναπαράστασης γνώσης βασισμένων στη λογική που έχουν σχεδιαστεί για την καταγραφή της γνώσης αλλά και τη διενέργεια εργασιών συλλογιστικής πάνω σε αυτή, με ένα δομημένο και κατανοητό τρόπο.

Οι ΠΛ αποτελούνται από ένα αλφάβητο (alphabet), ένα συντακτικό (syntax) και μια σημασιολογία (semantics). Πιο τυπικά αποτελείται από ένα αλφάβητο (alphabet) διακεκριμένων ατομικών εννοιών (atomic concepts) (C), ατομικών ρόλων (atomic roles) (R) και ατόμων (individuals) (I). Ταυτόχρονα όμως προσφέρουν και ένα σύνολο από κατασκευαστές οι οποίοι επενεργούν πάνω στις έννοιες και στους ρόλους και συντελούν στην κατασκευή περισσότερο πολύπλοκων εννοιών και εκφράσεων που ονομάζονται περιγραφές εννοιών ή σύνθετες έννοιες.

Επιπλέον οι ΠΛ μας δίνουν τη δυνατότητα να περιγράψουμε και σχέσεις ανάμεσα στις σύνθετες έννοιες με τη μορφή αξιωμάτων που καλούνται αξιώματα ορολογίας (terminological axioms). Πιο συγκεκριμένα, αν C και D είναι έννοιες ΠΛ τότε τα αξιώματα ορολογίας έχουν τη μορφή $C \sqsubseteq D$ (αξίωμα υπαγωγής - subsumption/inclusion axioms) και $C \equiv D$ (αξίωμα ισοδυναμίας - equivalence axioms) (οι σχέσεις υπαγωγής είναι σημαντικές μιας και δημιουργούν ιεραρχίες μεταξύ των εννοιών). Ένα σύνολο από αξιώματα υπαγωγής και ισοδυναμίας αποτελούν ένα σώμα ορολογίας (Terminological Box - TBox).

Για να αποκτήσει μια έννοια συγκεκριμένο νόημα και σημασία θα πρέπει να της αποδώσουμε μια συγκεκριμένη ερμηνεία. Ως ερμηνεία (interpretation) \mathcal{I} ορίζεται από ένα ζεύγος $(\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ όπου $\Delta^{\mathcal{I}}$ είναι ένα μη κενό σύνολο που καλείται χώρος ερμηνείας (domain of interpretation) και περιέχει ένα σύνολο αντικειμένων (objects) και $\cdot^{\mathcal{I}}$ είναι μια συνάρτηση ερμηνείας (interpretation function) η οποία απεικονίζει κάθε άτομο $a \in \mathcal{I}$ σε ένα αντικείμενο $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$, κάθε ατομική έννοια $CN \in \mathcal{C}$ σε ένα υποσύνολο $CN^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ και κάθε ατομικό ρόλο $RN \in \mathcal{R}$ σε μια δυαδική σχέση $RN^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$.

Η συνάρτηση ερμηνείας μπορεί να επεκταθεί περαιτέρω και να δώσει ερμηνεία και σε αξιώματα. Έτσι αν μια ερμηνεία \mathcal{I} ικανοποιεί ένα σώμα ισχυρισμών TBox, δηλαδή ικανοποιεί όλα τα αξιώματα υπαγωγής και ισοδυναμίας που υπάρχουν στο TBox, τότε λέμε ότι είναι μοντέλο (model) του TBox. Το TBox ουσιαστικά αποτελεί ένα σημαντικό τμήμα της γνώσης μας, το οποίο θέτει περιορισμούς σε πιθανά μοντέλα.

Οι ΠΛ μας δίνουν επιπλέον τη δυνατότητα να κάνουμε υποθέσεις όσο αφορά τα άτομα του κόσμου τον οποίο μοντελοποιούμε. Μας δίνει δηλαδή τη δυνατότητα καθορισμού σχέσεων στιγμιότυπου (instance relations) ανάμεσα σε ένα άτομο ή ζευγάρια ατόμων και μια έννοια ή ρόλο τα οποία ονομάζονται ισχυρισμοί (assertions). Οι ισχυρισμοί εννοιών έχουν τη σύνταξη $a : C$ ενώ οι ισχυρισμοί ρόλων αναπαρίσταται ως $(a, b) : R$. Το σύνολο των ισχυρισμών αποτελεί το σώμα ισχυρισμών (Assertion Box - ABox).

Μια ερμηνεία \mathcal{I} ικανοποιεί έναν ισχυρισμό $a : C((a, b) : R)$ αν $a^{\mathcal{I}} \in C^{\mathcal{I}}((a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}})$ ενώ ικανοποιεί ένα σώμα ισχυρισμών ABox αν ικανοποιεί όλους τους ισχυρισμούς που υπάρχουν στο ABox, τότε λέμε ότι είναι μοντέλο (model) του ABox.

Μια ερμηνεία \mathcal{I} ικανοποιεί ένα σώμα ισχυρισμών ABox με βάση το (μβτ) σώμα ορολογίας TBox αν η \mathcal{I} είναι μοντέλο του ABox αλλά ταυτόχρονα είναι και μοντέλο του TBox. Έτσι μια ΠΛ βάση γνώσης (knowledge base) ορίζεται ως ένα ζευγάρι ενός TBox και ενός ABox.

Στη βιβλιογραφία έχουν προταθεί διάφορες γλώσσες ΠΛ για αναπαράσταση γνώσης οι οποίες μπορούν να διαχωριστούν ανάλογα με τους κατασκευαστές που αυτές διαθέτουν. Για παράδειγμα η \mathcal{ALC} είναι η ελάχιστη εκφραστική ΠΛ η οποία περιλαμβάνει ένα αλφάβητο πρωτογενών εννοιών και ρόλων, και ένα σύνολο κατασκευαστών (άρνηση \neg μόνο για ατομικές έννοιες, τομή \sqcap , ένωση \sqcup , υπαρξιακού περιορισμού \exists και περιορισμού τιμής \forall) καθώς και από δύο περιγραφές εννοιών (καθολική έννοια \top , κενή έννοια \perp).

Προσθέτοντας επιπλέον κατασκευαστές εννοιών στην ΠΛ \mathcal{ALC} [116] μπορούμε να δημιουργήσουμε περισσότερο εκφραστικές γλώσσες οι οποίες θα μας δίνουν τη δυνατότητα να περιγράψουμε πιο πολύπλοκες έννοιες. Για παράδειγμα η \mathcal{SHIN} επεκτείνει την ΠΛ \mathcal{ALC} χρησιμοποιώντας τρία νέα αξιώματα ρόλων καθώς και ένα νέο κατασκευαστή. Τα αξιώματα ρόλων που περιλαμβάνονται στην \mathcal{SHIN} είναι οι μεταβατικοί ρόλοι $Trans()(\mathcal{S})$, η ιεραρχία ρόλων $\sqsubseteq (\mathcal{H})$ και η αντίστροφοι ρόλοι $\bar{()}$. Ο νέος κατασκευαστής είναι ο κατασκευαστής περιορισμού πληθικότητας (number restriction) $\geq \eta \leq ()$.

Ο κατασκευαστής πληθικότητας αποτελείται από δύο κατασκευαστές: τον κατασκευαστή-έννοια το-πολύ (at most) $\leq nR$ και τον κατασκευαστή-έννοια $\geq nR$ όπου n φυσικός αριθμός και R ρόλος. Με αυτούς τους κατασκευαστές μπορούμε να περιγράψουμε έννοιες π.χ. μορφές Αγίων που έχουν ακριβώς ένα πλοχμό:

Μορφή $\top \geq 1$ έχειΠλοχμό $\top \leq 1$ έχειΠλοχμό.

5.3.1.2 Ασαφείς Περιγραφικές Λογικές και f-SHIN

Σε αυτήν την ενότητα θα γίνει σύντομη περιγραφή της γραφής της DL f-SHIN η οποία είναι μια ασαφής επέκταση της DL SHIN όπως αυτή αναφέρεται στο [124]. Παρόμοια με τις σαφείς γλώσσες περιγραφικής λογικής, μια ασαφής γλώσσα περιγραφικής λογικής αποτελείται από ένα αλφάβητο των διάκριτων ονομάτων εννοιών (\mathbf{C}), των ονομάτων ρόλων (\mathbf{R}) και των ατομικών ονομάτων (\mathbf{I}), μαζί με ένα σύνολο κατασκευαστών για να κατασκευάσει περιγραφές, έννοιες και ρόλους.

Η συνάρτηση ερμηνείας, που περιγράψαμε στην Ενότητα 5.3.1.1, μπορεί να επεκταθεί περαιτέρω και να δώσει ερμηνεία και σε περιγραφές εννοιών, για παράδειγμα στο $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$ η έννοια \top ερμηνεύεται ως το σύνολο που περιέχει όλα τα αντικείμενα του χώρου ενώ στο $\perp^{\mathcal{I}} = \emptyset$ η έννοια \perp δηλώνει το κενό σύνολο.

Το σύνολο των f-SHIN εννοιών είναι το μικρότερο σύνολο στο οποίο ισχύουν τα εξής [122]:

1. Αν $C \in \mathbf{C}$, τότε C είναι μια f-SHIN-έννοια,
2. Οι ειδικές έννοιες \perp και \top είναι f-SHIN-έννοιες,
3. Αν C και D είναι f-SHIN έννοιες, το R είναι ένας f-SHIN ρόλος, το S είναι ένας απλός ρόλος και $n \in \mathbb{N}$, τότε $(\neg C)$, $(C \sqcup D)$, $(C \sqcap D)$, $(\forall R.C)$, $(\exists R.C)$, $(\geq nS)$ και $(\leq nS)$ είναι επίσης f-SHIN έννοιες.

Οι έννοιες \perp και \top ονομάζονται κενή (bottom) και καθολική (top) έννοια, αντίστοιχα. Οι έννοιες $(\forall R.C)$, $(\exists R.C)$ ονομάζονται περιορισμός τιμής (value restriction) και υπαρξιακός περιορισμός (exist restriction) αντίστοιχα, ενώ οι έννοιες

Κατασκευαστής	Συντακτικό	Σημασιολογία
Καθολική Έννοια	\top	$\top^{\mathcal{I}}(a) = 1$
Κενή Έννοια	\perp	$\perp^{\mathcal{I}}(a) = 0$
Γενική Άρνηση	$\neg C$	$(\neg C)^{\mathcal{I}}(a) = c(C^{\mathcal{I}}(a))$
Σύνδεση	$C \& D$	$(C \cap D)^{\mathcal{I}}(a) = t(C^{\mathcal{I}}(a), D^{\mathcal{I}}(a))$
Διάζευξη	$C \sqcup D$	$(C \cup D)^{\mathcal{I}}(a) = u(C^{\mathcal{I}}(a), D^{\mathcal{I}}(a))$
Υπαρξιακός Περιορισμός	$\exists R.C$	$(\exists R.C)^{\mathcal{I}}(a) = \sup_{b \in \Delta^{\mathcal{I}}} \{t(R^{\mathcal{I}}(a, b), C^{\mathcal{I}}(b))\}$
Περιορισμός Τιμής	$\forall R.C$	$(\forall R.C)^{\mathcal{I}}(a) = \inf_{b \in \Delta^{\mathcal{I}}} \{\mathcal{J}(R^{\mathcal{I}}(a, b), C^{\mathcal{I}}(b))\}$
Περιορισμός το λιγότερο	$\leq pR$	$\inf_{b_1, \dots, b_{p+1} \in \Delta^{\mathcal{I}}} \mathcal{J}(t_{i=1}^{p+1} R^{\mathcal{I}}(a, b_i), u_{i < j} \{b_i = b_j\})$
Περιορισμός το πολύ	$\geq pR$	$\sup_{b_1, \dots, b_p \in \Delta^{\mathcal{I}}} t(t_{i=1}^p R^{\mathcal{I}}(a, b_i), t_{i < j} \{b_i \neq b_j\})$
Αντίστροφοι Ρόλοι	R^-	$(R^-)^{\mathcal{I}}(b, a) = R^{\mathcal{I}}(a, b)$
Ισοδυναμία	$C \equiv D$	$\forall a \in \Delta^{\mathcal{I}}. C^{\mathcal{I}}(a) = D^{\mathcal{I}}(a)$
Υποέννοιες	$C \sqsubseteq D$	$\forall a \in \Delta^{\mathcal{I}}. C^{\mathcal{I}}(a) \leq D^{\mathcal{I}}(a)$
Μεταβατικοί ρόλοι	$\text{Trans}(R)$	$\forall a, b \in \Delta^{\mathcal{I}}. R^{\mathcal{I}}(a, b) \geq \sup_{c \in \Delta^{\mathcal{I}}} \{t(R^{\mathcal{I}}(a, c), R^{\mathcal{I}}(c, b))\}$
ΥποΡόλοι	$R \sqsubseteq S$	$\forall a, b \in \Delta^{\mathcal{I}}. R^{\mathcal{I}}(a, b) \leq S^{\mathcal{I}}(a, b)$
Έννοιες Ισχυρισμών	$\langle a : C \triangleright n \rangle$	$C^{\mathcal{I}}(a^{\mathcal{I}}) \triangleright n$
Ρόλοι Ισχυρισμών	$\langle \langle a, b \rangle : R \triangleright n \rangle$	$R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \triangleright n$

Πίνακας 5.1: Σημασιολογία Εννοιών και Ρόλων

τις μορφής ($\geq nS$) και ($\leq nS$) ονομάζονται ως περιορισμοί πληθικότητας (number restrictions) το λιγότερο (at-most) από τους οποίους ο ($\geq nS$) αναφέρεται σαν περιορισμός το λιγότερο (at-least) ενώ ο ($\leq nS$) σαν περιορισμός το πολύ (at-most).

Σε αντίθεση με τις σαφείς ΠΛ, η σημασιολογία των ασαφών ΠΛ ορίζεται με τη βοήθεια των ασαφών ερμηνειών (fuzzy interpretations) [126]. Μία ασαφής ερμηνεία αποτελείται από ένα ζεύγος $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ όπου ο χώρος ερμηνείας $\Delta^{\mathcal{I}}$ είναι ένα μη κενό σύνολο αντικειμένων και $\cdot^{\mathcal{I}}$ είναι μια ασαφής συνάρτηση ερμηνείας, η οποία αντιστοιχεί ένα άτομο a σε ένα στοιχείο $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$ και μία ατομική έννοια A (όνομα ρόλου R) σε μία συνάρτηση συμμετοχής $A^{\mathcal{I}} : \Delta^{\mathcal{I}} \rightarrow [0, 1]$ και έναν ατομικό ρόλο $R \in \mathbf{R}$ σε μία συνάρτηση συμμετοχής ($R^{\mathcal{I}} : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0, 1]$)

Χρησιμοποιώντας τις θεωρητικές λειτουργίες των ασαφών συνόλων, η συνάρτηση ασαφών ερμηνειών μπορεί να επεκταθεί και να προσδώσει σημασιολογία σε πολύπλοκες έννοιες, ρόλους και αξιώματα. Οι τυπικοί ασαφής λειτουργοί του $1-x$ (c) για ασαφή άρνηση και \max (u), \min (t) για σαφή ένωση και τομή χρησιμοποιούνται στην περίπτωση μας. Το πλήρες σημασιολογικό σύνολο απεικονίζεται στον Πίνακα 5.1:

Μία γνώση βασισμένη σε f-SHLN είναι μία πλειάδα (τριπλέτα -triple) $\langle \mathcal{T}, \mathcal{R}, \mathcal{A} \rangle$, όπου \mathcal{T} είναι ένα ασαφές σώμα ορολογίας (terminological box - $TBox$), \mathcal{R} είναι ένα ασαφές σώμα ρόλων (role box - $RBox$ και \mathcal{A} είναι ένα ασαφές σώμα ισχυρισμών (assertional box - $ABox$).

Το $TBox$ είναι ένα πεπερασμένο σύνολο από αξιώματα ασαφών εννοιών τα οποία όσα είναι της μορφής $C \sqsubseteq D$ ονομάζονται αξιώματα υπαγωγής ασαφών εννοιών (fuzzy concept inclusion axioms) ενώ όσα είναι της μορφής $C \equiv D$ καλούνται αξιώματα ισοδυναμίας ασαφών εννοιών (fuzzy concept equivalence axioms), όπου C, D είναι έννοιες, τέτοιες ώστε η C να είναι υπαγωγή και η C είναι ισοδύναμη με την D , αντίστοιχα. Μια ασαφής ερμηνεία I ικανοποιεί το αξίωμα $C \sqsubseteq D$ αν

$\forall \alpha \in \Delta^I, C^I(\alpha) \geq D^I(\alpha)$ ενώ ικανοποιεί το $C \equiv D$ αν $C^I(\alpha) = D^I(\alpha)$. Μια ασαφής ερμηνεία I ικανοποιεί ένα f-SHLN σώμα ισχυρισμών T αν ικανοποιεί κάθε αξίωμα στο T . Τότε λέμε ότι η I είναι μοντέλο (model) του T .

Το $RBox$ είναι ένα πεπερασμένο σύνολο από αξιώματα ασαφών ρόλων τα οποία όσα είναι της μορφής $Trans(R)$ ονομάζονται αξιώματα μεταβατικών ασαφών ρόλων (fuzzy transitive role axioms), ενώ αξιώματα ασαφών ρόλων της μορφής $R \sqsubseteq S$ ονομάζονται αξιώματα υπαγωγής ασαφών ρόλων (fuzzy role inclusion axioms), όπου R είναι μια μετάβαση και R είναι υπορόλος του S αντίστοιχα. Μια ασαφής ερμηνεία \mathcal{I} ικανοποιεί ένα αξίωμα $Trans(R)$ αν $\forall a, c \in \Delta^{\mathcal{I}} x \Delta^{\mathcal{I}}, R^{\mathcal{I}}(a, c) \leq \sup_{b \in \Delta^{\mathcal{I}}} \{t(R^{\mathcal{I}}(a, b), R^{\mathcal{I}}(b, c))\}$ ενώ ικανοποιεί το $R \sqsubseteq S$ αν $\forall (\alpha, b) \in \Delta^{\mathcal{I}} x \Delta^{\mathcal{I}}, R^{\mathcal{I}}(\alpha, b) \leq S^{\mathcal{I}}(\alpha, b)$. Η I ικανοποιεί ένα f-SHLN $RBox$ αν ικανοποιεί όλα τα αξιώματα στο R . Σε αυτή την περίπτωση η \mathcal{I} λέγεται μοντέλο του R . Ένα σύνολο από αξιώματα υπαγωγής ρόλων καθορίζει μια ιεραρχία ρόλων (role hierarchy) R_h . Επίσης ισχύει ότι αν $R \sqsubseteq S$, τότε και $Inv(R) \sqsubseteq Inv(S)$, όπως και στην κλασική περίπτωση.

Ένα ασαφές σώμα ισχυρισμών (assertional box) $ABox$ είναι ένα πεπερασμένο σύνολο από ασαφή αξιώματα ατόμων τα οποία αποκαλούνται ασαφείς ισχυρισμοί. Ένας ασαφής ισχυρισμός (fuzzy assertion) είναι μια δήλωση της μορφής $a : C \bowtie n$, $(a, b) : R$, όπου $\bowtie \in (\geq, >, \leq, <)$, $a = b$ ή $a \neq b$, $\forall a, b \in \mathbf{I}$. Επιπλέον θα γράφεται $a : C = n$ αντί να γράφονται οι δυο ασαφείς ισχυρισμοί της μορφής $a : C \leq n$ και $a : C \geq n$.

5.3.2 Υπηρεσίες και μηχανισμοί εξαγωγής συμπερασμάτων

Στην προηγούμενη ενότητα παρουσιάστηκε η δημιουργία βάσεων γνώσεων με τον καθορισμό του σώματος ορολογίας και του σώματος ισχυρισμών χρησιμοποιώντας ΠΛ και ασαφή ΠΛ. Για να μπορέσουμε να εξάγουμε γνώση ή νέα γνώση από τη βάση γνώσης θα πρέπει να εφοδιάσουμε το σύστημά μας με υπηρεσίες εξαγωγή συμπερασμάτων τις οποίες παρέχουν οι ΠΛ. Οι ΠΛ παρέχουν υπηρεσίες τόσο πάνω στα σώματα ορολογίας όσο και στα σώματα ισχυρισμών.

Έστω ένα $Tbox$. Θα λέμε ότι μία έννοια C :

- ικανοποιεί (satisfy) μβτ το αν \exists μοντέλο του : $C^{\mathcal{I}} = \emptyset$
- υπάγεται (subsumes) στην έννοια D μβτ το αν $C^{\mathcal{I}} \subseteq D^{\mathcal{I}} \forall$ μοντέλο \mathcal{I} του ($\models C \sqsubseteq D$)
- ισοδυναμεί (be equivalent) με την έννοια D μβτ το αν $C^{\mathcal{I}} = D^{\mathcal{I}} \forall$ μοντέλο \mathcal{I} του ($\models C \equiv D$)
- είναι ξένη (disjoint with) με την έννοια D μβτ το αν $C^{\mathcal{I}} \cap D^{\mathcal{I}} = \emptyset \forall$ μοντέλο \mathcal{I} του

Όλες οι υπηρεσίες εξαγωγή συμπερασμάτων μπορούν να αναχθούν σε ακριβώς μία και συγκεκριμένα να αναχθούν στο πρόβλημα μη ικανοποιησιμότητας [44]. Πιο συγκεκριμένα, αν η ΠΛ περιέχει άρνηση τότε έχουμε τις εξής ισοδυναμίες :

Θα λέμε ότι η έννοια C :

- υπάγεται (subsumes) στην έννοια D αν η έννοια $C \sqcap \neg D$ είναι μη-ικανοποιήσιμη

- *ισοδυναμεί* (be equivalent) με την έννοια D αν η έννοια $C \sqcap \neg D$ και η έννοια $\neg C \sqcap D$ είναι μη-ικανοποιήσιμες
- είναι *ξένη* (disjoint with) με την έννοια D αν η έννοια $C \sqcap D$ είναι μη-ικανοποιήσιμη

Αν είναι ένα $ABox$ και ένα $TBox$ τότε :

- Το είναι *συνεπές* (consistent) μβτ αν \exists μοντέλο του το οποίο είναι και μοντέλο του
- Το *συνεπάγεται* (entails) έναν ισχυρισμό ϕ μβτ , αν κάθε μοντέλο του και του ικανοποιεί τον ισχυρισμό ($\models \phi$)

Τα $ABoxes$ [96] ουσιαστικά αποτελούν μια αφαίρεση ενός μόνο κομματιού του κόσμου μας και ποτέ δεν αντιπροσωπεύουν ολόκληρο τον κόσμο. Άρα αν κάποιος ισχυρισμός δε βρίσκεται μέσα στο $ABox$ αυτό δεν σημαίνει ότι ο ισχυρισμός δεν ισχύει αλλά ότι δε ξέρουμε αν ισχύει και ότι μπορούμε να δημιουργήσουμε ερμηνείες στις οποίες ο ισχυρισμός αυτός να ικανοποιείται και άλλες στις οποίες μπορεί να μην ικανοποιείται. Αυτή η υπόθεση ονομάζεται υπόθεση ανοικτού κόσμου.

Έχοντας ορίσει τη σύνταξη, τη σημασιολογία και τις υπηρεσίες εξαγωγής συμπερασμάτων πρέπει να ορίσουμε αυτοματοποιημένους τρόπους (αλγόριθμους) οι οποίοι θα λύνουν το πρόβλημα της συνέπειας και άρα μεταβατικά οποιοδήποτε πρόβλημα μιας ΠΛ γλώσσας. Στόχος αυτών των αλγόριθμων είναι να "κτίσουν" μια ερμηνεία που να ικανοποιεί τους ισχυρισμούς απλοποιώντας πολύπλοκες έννοιες (με τη βοήθεια των κατασκευαστών) έως ότου φτάσουν σε απλούστερες των οποίων η ικανοποιησιμότητα ή μη είναι προφανής. Για την απλοποίηση αυτών των εννοιών δημιουργούνται κανόνες που βασίζονται στη σημασιολογία των κατασκευαστών της γλώσσας. Αυτοί οι αλγόριθμοι ονομάζονται αλγόριθμοι πινάκων (tableaux algorithms). Για κάθε ΠΛ έχει προταθεί μια tableaux διαδικασία που μπορεί να αποφανθεί για το πρόβλημα της συνέπειας σε ένα σώμα ισχυρισμών.

Στις ασαφείς ΠΛ, το πρόβλημα της συνέπειας ελέγχεται συνήθως με αλγόριθμους ασαφούς tableaux για το $ABox$ [124], το οποίο αποτελεί μια αφαίρεση (abstraction) ενός μοντέλου του A . Με την έννοια αφαίρεση εννοούμε ότι ενώ στη γενική περίπτωση σε ένα μοντέλο τα αντικείμενα μπορεί να είναι τυχαία συνδεδεμένα μεταξύ τους, σε ένα tableau οι κόμβοι συνδέονται με έναν τέτοιο τρόπο ώστε να δημιουργείται μια δεντρική (tree-like) ή δασική (forest-like) δομή.

Το ασαφές tableau είναι μια επέκταση της δομής tableau που παρουσιάστηκε στην εργασία [63] λαμβάνοντας υπόψη τη διαχείριση των βαθμών συμμετοχής καθώς και την εισαγωγή ονοματικών εννοιών [124]. Για τη συνέχεια, και χωρίς βλάβη της γενικότητας θεωρούμε ότι όλες οι έννοιες C που εμφανίζονται σε έναν ασαφή ισχυρισμό $(a : C) \cdot n \in A$ βρίσκονται αφενός στην κανονική μορφή θετικής ανισότητας και αφετέρου οι έννοιες C σε κανονική μορφή άρνησης. Κάτι τέτοιο είναι δυνατό γιατί οι τελεστές της ασαφούς λογικής $f\text{-SHLN}$ [123] ικανοποιούν όλες τις ιδιότητες που χρειάζονται για να ισχύει η ιδιότητα της κανονικής μορφής άρνησης.

Η ορθότητα (soundness) των αλγόριθμων tableaux ελέγχεται με το αν το σώμα ισχυρισμών είναι συνεπές (οπότε μπορούμε να κατασκευάσουμε ένα μοντέλο για αυτό) ενώ η πληρότητα (completeness) ελέγχεται με το αν ο αλγόριθμος τερματίζει και τουλάχιστον ένα από τα σώματα ισχυρισμών που δημιουργήθηκαν δεν περιέχει συγχρούσεις, δοθέντος ενός αρχικού σώματος ισχυρισμών.

Ο αλγόριθμος tableau για την *f-SHIN* υποστηρίζεται πλήρως από τη μηχανή ασαφούς συμπερασματολογίας FiRE¹ και η οποία χρησιμοποιήθηκε στο πλαίσιο της παρούσας εργασίας. Για περισσότερες πληροφορίες σχετικά με την *f-SHIN* τη μηχανή ασαφούς συμπερασματολογίας FiRE μπορείτε να ανατρέξετε στην εργασία [125].

5.4 Συστήματα γνώσης και συστήματα μαθησιακής τεχνολογίας

Ένα σύστημα γνώσης ουσιαστικά είναι ένα υπολογιστικό συστήματα που επιδιώκει να λειτουργεί όπως ο άνθρωπος, εστιάζοντας κάθε φορά σε έναν τομέα ειδικευσης. Η ανάπτυξη ενός συστήματος γνώσης προϋποθέτει τη συνεργασία ενός ειδικού στον τομέα πάνω στον οποίο πρόκειται να δημιουργηθεί το σύστημα γνώσης, με ένα μηχανικό γνώσης. Ο ειδικός έχει τις απαραίτητες γνώσης του τομέα ενώ ο μηχανικός γνώσης είναι ο ειδικευμένος επιστήμονας σε συστήματα γνώσης. Ο μηχανικός γνώσης συνεργάζεται με τον ειδικό του τομέα ενδιαφέροντος με σκοπό αρχικά την κατανόηση και μεταφορά της γνώσης και έπειτα το σχεδιασμό του συστήματος και την ορθή αναπαράσταση της γνώσης σε αυτό.

Το ολοκληρωμένο σύστημα γνώσης μπορεί να χρησιμοποιηθεί από τον τελικό χρήστη με σκοπό τη βοήθεια στη λήψη απόφασης πάνω σε ένα ζήτημα του τομέα, χωρίς την παρουσίαση του ειδικού. Με αυτό τον τρόπο ο τελικός χρήστης, συνήθως μη ειδικός στον τομέα ενδιαφέροντος στο οποίο έχει αναπτυχθεί το σύστημα γνώσης, εκπαιδεύεται πάνω στο πεδίο.

Σύμφωνα με την προτεινόμενη επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE (Ενότητα 2.5) οι ειδικοί είναι οι χρήστες που εφοδιάζουν το σύστημα με γνώση την οποία άμεσα ή έμμεσα θα παραλάβει ο χρήστης. Το σύστημα γνώσης ουσιαστικά μπορεί να ενσωματωθεί μέσα στις λειτουργίες που παρέχει η διαδικασία Δάσκαλος. Με αυτό τον τρόπο η διαδικασία Δάσκαλος μπορεί να δημιουργήσει τα κατάλληλα ερωτήματα, και με τη βοήθεια της μηχανής συμπερασματολογίας να ανακτήσει, από την αποθήκη Υλικό Εκμάθησης, το κατάλληλο υλικό για το χρήστη.

Στόχος ενός συστήματος γνώσης είναι η επίλυση προβλημάτων (απάντηση σε ερωτήματα) χρησιμοποιώντας την υπάρχουσα γνώση ή δημιουργώντας νέα γνώση. Αρχικά, προτάθηκαν συστήματα των οποίων η γνώση προερχόταν κυρίως από την εμπειρία ειδικών πάνω σε κάποιον τομέα και χρησιμοποιούσαν κλασικές συλλογιστικές για εξαγωγή νέας γνώσης [16].

Ένα από τα πλεονεκτήματα των συστημάτων γνώσης είναι ότι μπορούν να ενσωματωθούν σε ένα σύστημα μαθησιακής τεχνολογίας και προσφέρουν στο χρήστη του εξειδικευμένη γνώση, χωρίς την ανάγκη ύπαρξης ανθρώπου-δασκάλου και διατηρώντας σε χαμηλό επίπεδο το κόστος της διαδικασίας μάθησης. Επιπλέον, ένα σύστημα γνώση μπορεί να βασίζεται στην αξιοποίηση της γνώσης πολλών ειδικών που μπορεί να έχουν συνθέσει τη γνώση του συστήματος. Τα συστήματα γνώσης έχουν τη δυνατότητα αντιμετώπισης πολύπλοκων προβλημάτων, όπως η αναζήτηση πολύπλοκων σχέσεων στις οντότητες που περιλαμβάνονται.

Τα βασικά συστατικά μέρη ενός συστήματος γνώσης, τα οποία ενσωματώνονται μέσα στη διαδικασία Δάσκαλος του προτεινόμενου συστήματος μαθησιακής

¹<http://www.image.ece.ntua.gr/~nsimou/FiRE/>

τεχνολογίας (Ενότητα 2.5) είναι τα εξής:

- Η *βάση γνώσης* η οποία αναπαριστά σε κωδικοποιημένη μορφή τη γνώση του πεδίου ενδιαφέροντος, όπως αυτή αποδόθηκε από τους ειδικούς του τομέα. Η γνώση αυτή μπορεί να αποδοθεί σε αφηρημένη μορφή σε αιτιοκρατικές καταστάσεις, αποτελούμενη από ένα σύνολο αντικειμένων και ένα σύνολο κανόνων (που καθορίζει τις συσχετίσεις μεταξύ αντικειμένων) ή σε συγκεκριμένη μορφή σε συγκεκριμένες καταστάσεις, αποτελούμενη από γεγονότα και ενδείξεις.
- Το *υποσύστημα αναζήτησης γνώσης* που ελέγχει τη ροή νέας γνώσης από τους ειδικούς προς τη βάση γνώσης. Σκοπός του είναι να ελέγχει αν η γνώση που λαμβάνεται είναι πραγματικά νέα και χρήσιμη για το σύστημα.
- Το *υποσύστημα ελέγχου συνέπειας* το οποίο εντοπίζει ασυνεπή γνώση και εμποδίζει την εισαγωγή νέας γνώσης που έρχεται σε σύγκρουση με την ήδη υπάρχουσα ή ενημερώνει σχετικά τον άνθρωπο-ειδικό που διαχειρίζεται το εκάστοτε σύστημα γνώσης.
- Η *μηχανή συλλογιστικής (inference engine)* η οποία αποτελεί τον πυρήνα των συστημάτων γνώσης. Ο βασικός σκοπός της είναι η εφαρμογή της αφηρημένης γνώσης στη συγκεκριμένη γνώση. Τα συμπεράσματα που λαμβάνονται από τη μηχανή συλλογιστικής μπορεί να βασίζονται είτε σε αιτιοκρατική γνώση, είτε σε πιθανοτική γνώση είτε σε ασαφή γνώση. Όπως είναι αναμενόμενο, ο χειρισμός πιθανοτικής /ασαφούς εξαγωγής συμπερασμάτων είναι αρκετά δυσκολότερος από την αιτιοκρατική περίπτωση. Όμως στην πλειονότητα των εφαρμογών καθίσταται αναγκαίος, γιατί υπεισέρχεται αβεβαιότητα, τόσο σε επίπεδο αφηρημένης γνώσης όσο και σε συγκεκριμένης.
- Το *υποσύστημα αναζήτησης γνώσης* που έχει ως στόχο τον εντοπισμό και την αναζήτηση ελλειπόντων γνώσης σε περιπτώσεις στις οποίες δεν επαρκεί η υπάρχουσα γνώση για την εξαγωγή συμπερασμάτων. Η αναζήτηση μπορεί να γίνει είτε από τον ειδικό, είτε αυτόματα π.χ. σε βάσεις δεδομένων ή στο internet.
- Το *υποσύστημα επεξήγησης* το οποίο παρέχει στον ειδικό τον τρόπο με τον οποίο το σύστημα γνώσης κατέληξε στα συμπεράσματά του. Το σύστημα γνώσης δηλαδή δικαιολογεί την απόφασή του μέσω αυτού του υποσυστήματος. Η ικανοποιητική λειτουργία του υποσυστήματος αυτού είναι ουσιώδης στην εξασφάλιση της εγκυρότητας των αποτελεσμάτων ενός συστήματος γνώσης.
- Το *υποσύστημα μάθησης* που παρέχει μηχανισμούς συνεχούς προσαρμογής της γνώσης από δεδομένα εκπαίδευσης. Υπάρχουν δύο είδη μάθησης στα οποία η αναπαράσταση της γνώσης γίνεται συνήθως χρησιμοποιώντας ως δομική μονάδα τους κανόνες: α) *Δομική μάθηση* που αναφέρεται στην εκμάθηση νέων κανόνων, στην αναπροσαρμογή των υπαρχόντων και στη διαγραφή κανόνων που έρχονται σε αντίθεση με τα δεδομένα εκπαίδευσης και β) *Παραμετρική μάθηση* στην οποία χρησιμοποιείται η μέθοδος μέγιστης πιθανοφάνειας με σκοπό τη βέλτιστη προσαρμογή των παραμέτρων στα δεδομένα εκπαίδευσης.

5.4.1 Αναπαράσταση πληροφορίας από ανάλυση πολυμεσικών δεδομένων με χρήση ασαφούς γνώσης

Η μονάδα Αναπαράστασης Γνώσης και Συλλογιστική του προτεινόμενου συστήματος αυτόματης κατηγοριοποίησης Βυζαντινών εικόνων με βάση τη γνώση (Σχήμα 4.3) έχει σκοπό την κατηγοριοποίηση της εικόνας.

Η μονάδα αναπαράστασης γνώσης και συλλογιστικής αποτελείται από το σώμα ασαφών ρόλων, το σώμα ασαφούς ορολογίας και από τη μηχανή συλλογισμού (Σχήμα 4.5). Αναλυτικότερα, το σώμα ασαφών ρόλων περιέχει όλες τις ιδιότητες που έχουν αποδοθεί στα σημασιολογικά τμήματα από τη μονάδα ανάλυσης Βυζαντινής εικόνας, προσδιορισμένες με χρήση ασαφών ισχυρισμών. Το σώμα ασαφούς ορολογίας ουσιαστικά περιλαμβάνει τη γνώση του πεδίου, δομημένη με τη χρήση της περιγραφικής λογικής. Η μηχανή συλλογιστικής που χρησιμοποιείται από το σύστημά μας είναι η μηχανή FiRE (www.image.ece.ntua.gr/~nsimou/FiRE) η οποία υποστηρίζει πλήρως την f-SHIN λογική (Ενότητα 5.3). Η μηχανή FiRE αξιολογεί το μέγιστο πιθανό βαθμό συμμετοχής ενός ατόμου- σημασιολογικού τμήματος σε μια έννοια.

Η μηχανή συλλογιστικής, όπως προαναφέρθηκε, αποτελεί τον πυρήνα του συστήματος γνώσης στον οποίο ενσωματώνεται η αφηρημένη και η συγκεκριμένη γνώση. Η αφηρημένη γνώση αντιπροσωπεύεται από το σώμα ισχυρισμών (Ενότητα 4.4.3.1) το οποίο βασίζεται στη συλλογή όλων των πληροφοριών που εξάγονται από μία ψηφιοποιημένη βυζαντινή εικόνα με τη βοήθεια της μονάδας ανάλυσης Βυζαντινών εικόνων (Ενότητα 4.4.2) και περιέχει όλες τις ιδιότητες που έχουν αποδοθεί στα σημασιολογικά τμήματα μιας εικόνας, προσδιορισμένα με χρήση ασαφών ισχυρισμών. Η συγκεκριμένη γνώση αντιπροσωπεύεται από το σώμα ορολογίας (TBox) και παρουσιάζεται με τη μορφή οντολογίας περιγραφής των Βυζαντινών μορφών.

Με βάση την τεχνογνωσία που αποκτήθηκε από τη μελέτη και εφαρμογή του εγχειριδίου του Διονυσίου εκ Φουρνά και με τη βοήθεια των ειδικών σε θέματα Βυζαντινής αγιογραφίας συγκεκριμενοποιήθηκε η κοινή οντολογία θεματικών κατηγοριών για βυζαντινές ιστάμενες μορφές. Συγκεκριμένα, η οντολογία που δημιουργήθηκε αφορά τη κατηγορία πρόσωπο της μορφής, μιας και αυτή η περιοχή είναι σημασιολογικά πλούσια, σύμφωνα με τον Διονύσιο εκ Φουρνά και μπορούν να εξαχθούν πολλές πληροφορίες από εκεί. Η οντολογία αυτή χρησιμοποιείται σαν μέσο διαχείρισης γνώσης και εξαγωγής συμπερασμάτων από τη μηχανή συλλογιστικής. Η γνώση αποτελείται από ένα ασαφές σώμα ορολογίας (TBox) και ένα ασαφές σώμα ισχυρισμών (ABox). Η ορολογία ορίζει το λεξιλόγιο το οποίο θα χρησιμοποιηθεί από την εφαρμογή ενώ από την άλλη το σώμα ισχυρισμών περιλαμβάνει ισχυρισμούς για τα άτομα κάνοντας χρήση του λεξιλογίου, το οποίο αποτελείται από έννοιες, ρόλους και άτομα.

Για τη δημιουργία της γνώσης έγινε η υπόθεση ότι μία αγιογραφία αποτελείται από τμήματα (Segments) τα οποία συνδέονται μεταξύ τους ορίζοντας το ρόλο περιέχει για την σύνδεση τους. Τα βασικά τμήματα τα οποία εντοπίζονται αρχικά είναι το πρόσωπο, τα μάτια και η μύτη βάση των οποίων εντοπίζονται στη συνέχεια τα τμήματα που αποτελούν το μέτωπο, το μαλλί, το μουστάκι την παρειά και το γενειάδα. Έτσι λοιπόν θεωρούμε ότι μια εικόνα περιέχει όλα τα παραπάνω τμήματα καθώς και ότι το τμήμα του προσώπου περιέχει όλα τα τμήματα τα οποία είναι χαρακτηριστικά του προσώπου (δηλαδή το μέτωπο, το μαλλί, το μουστάκι κτλ).

Ως άτομα της οντολογίας που δημιουργήσαμε χαρακτηρίζαμε τα τμήματα τα οποία

εντοπίζουμε κατά την ανάλυση των Βυζαντινών εικόνων: Πρόσωπο, Μάτια, Μύτη, Μαλλία, Μέτωπο, Παρειά, Μουστάκι, Γενειάδα.

Οι έννοιες που απαρτίζουν την οντολογία μας είναι αυτά που αναφέρθηκαν στους Πίνακες της Ενότητας 4.3. Αναλυτικότερα, οι έννοιες που ορίσαμε και οι οποίες είναι ικανές να χαρακτηρίσουν τα άτομα είναι Δασυμάλλης, Ολιγομάλλης, ΠάνωΑποΑυτιά, ΚάτωΑποΑυτιά, ΠλοχμόςΑριστερά, ΠλοχμόςΔεξιά, ΠλοχμόςΚαιΣτους2Ώμους, Μαυρομάλλης, Γκρίζομάλλης, Πολιασμένος, Ισιοκέφαλος, Αγριοκέφαλος, Σγουροκέφαλος. Νέου, Γέρου, Σγουρό, Ίσιο, Δασυγένης Ολιγογένης, Κοντογένης, ΜακρυγένηςΟυΠολλά, Μακρυγένης, ΜακρυγένηςΩςΤηΜέση, ΜακρυγένηςΩςΤηΖώνη, ΜακρυγένηςΩςΤαΓόνατα, Οξυγένης, ΣτρογγυλογένηςΟλίγον, Στρογγυλογένης, Πλατυγένης, Φουντογένης, Βαδαρογένης, ΤρίχαςΗγριωμένας, έχειΚυματισμό, Κατζαρογένης, ΜεΤρίχας, Γυμνοπώγων, ΑνοιχτόχρωμοΓένυ, ΣκουρόχρωμοΓένυ.

Με παρόμοιο τρόπο ορίστηκαν οι σύνθετες έννοιες που χρησιμοποιούνται για τον γενικό χαρακτηρισμό Βυζαντινών μορφών: ΈχειΠαρειά, ΈχειΠαρειάΝέου, ΈχειΠαρειάΓέρου, ΈχειΜέτωπο, ΈχειΜέτωποΝέου, ΈχειΜέτωποΓέρου, ΈχειΚοντήΓενειάδα, ΈχειΣκούραΓενειάδα (Παράρτημα 8).

Αυτές οι έννοιες χρησιμοποιήθηκαν για την αναπαράσταση ακόμη πιο σύνθετων εννοιών όπως: Άντρας, Γυναίκα, ΝέοςΆντρας, ΏριμοςΆντρας, ΗλικιωμένοςΆντρας, ΝέαΓυναίκα, ΏριμηΓυναίκα... και έννοιες με τις οποίες καθορίζονται κάποιες άγιες μορφές όπως: Χριστός, Παναγία, ΙωάννηςΒαπτιστής, ΙωάννηςΘεολόγος, Λουκάς, Ανδρέας, Σίμων, Αθανάσιος, Χρυστόστομος, Νικόλαος, Βασίλειος, Αντώνιος, Ευθύμιος, ΙωάννηςΤηςΚλίμακος, Μηνάς, Θεόδωρος, Κωνσταντίνος, Θωμάς, Στέφανος, Μάμας, Δημήτριος, Γεώργιος, Προκόπιος, Αικατερίνα, Ελένη, Πελαγία οι οποίες σε συνδυασμό με τους κανόνες όπως αυτοί κατεγράφησαν από τον Διονύσιο εκ Φουρνά χρησιμοποιούνται για την κατηγοριοποίηση των Βυζαντινών εικόνων. Η κατηγοριοποίηση γίνεται με τον ορισμό των κατάλληλων αξιωμάτων (Παράρτημα 8).

Για παράδειγμα η μορφή μιας ώριμης γυναίκας στη Βυζαντινή τέχνη ορίζεται σύμφωνα με τον Διονύσιο εκ Φουρνά ως μια μορφή χωρίς γένυ η οποία έχει παρειές γέρου καθώς και παραπάνω από δυο ρυτίδες στο μέτωπο. Έτσι λοιπόν μια ώριμη γυναίκα σε μια εικόνα απεικονίζεται εάν η εικόνα περιέχει ένα τμήμα το οποίο αποτελεί σύζευξη των εννοιών:

Πρόσωπο και $(\exists$ περιέχει. (Παρειά Π (\forall έχειΠαρειάΓέρου))) και \exists περιέχει. (Μέτωπο Π (\geq έχειΡυτίδα2)) και $(\exists$ περιέχει. (\neg Γενειάδα))

δηλαδή περιέχει Πρόσωπο και περιέχειτμήμα που είναι Παρειά και όλες οι παρειές που έχει είναι γέρου και περιέχει τμήμα που είναι Μέτωπο και έχει πάνω από 2 ρυτίδες και περιέχει τμήμα που δεν είναι Γενειάδα.

Το αντίστοιχο αξίωμα που ορίζει μία Ώριμη Γυναίκα είναι: $\text{ριμηΓυναίκα} \equiv (\exists$ περιέχει. Πρόσωπο Π (\forall περιέχει. Π (\exists περιέχει. (Παρειά Π (\forall έχειΠαρειάΓέρου)))) Π (\exists περιέχει. (Μέτωπο Π (\geq έχειΡυτίδα2))) Π (\exists περιέχει. (\neg Γενειάδα)))

Το σώμα ορολογίας TBox που ορίστηκε για τις Βυζαντινές εικόνες παρατίθεται στο παράρτημα Α της παρούσης διατριβής.

Το TBox στη γνώση μας περιγράφεται με τη χρήση του συντακτικού της μηχανής ασαφούς συμπερασματολογίας FiRE, το οποίο στηρίζεται στη σύνταξη που έχει οριστεί για την αναπαράσταση γνώσης σε συστήματα που βασίζονται σε

ΠΛ (Description-Logic Knowledge Representation System Specification (KRSS) <http://www.dl.kr.org/krssspec.ps>) με τις κατάλληλες επεκτάσεις για την ασάφεια. Το συντακτικό κάνει χρήση του λεξιλογίου το οποίο αποτελείται από έννοιες, ρόλους και άτομα. Για παράδειγμα: (signature :atomic-concepts (Πρόσωπο Μαλλιά Γενειάδα ...):roles ((περιέχει)):individuals (filename segment1 segment2))

Η λέξη κλειδί *signature* χρησιμοποιείται για να δηλωθεί ένα σώμα ορολογίας, με τις λέξεις κλειδιά *atomic-concepts*, *roles* και *individuals* να χρησιμοποιούνται για τη δήλωση των εννοιών, των ρόλων, και των ατόμων αντίστοιχα. Οι έννοιες και τα άτομα δηλώνονται έπειτα από την αντίστοιχη λέξη κλειδί, μέσα σε παρενθέσεις, γράφοντας τα ονόματα τους, χωρισμένα με κενά. Από την άλλη οι ρόλοι, λόγω των επιπρόσθετων ιδιοτήτων τους, δηλώνονται σε ξεχωριστές παρενθέσεις ανά ρόλο μαζί με τις ιδιότητες τους.

Αν θέλουμε να δηλώσουμε πως μια έννοια είναι μεταβατική χρησιμοποιούμε τη λέξη κλειδί *transitive* ακολουθούμενη από το γράμμα t (true) ή από το γράμμα f (false). Αν δεν χρησιμοποιηθεί η λέξη κλειδί *transitive*, το FiRE θεωρεί πως ο ρόλος δεν είναι μεταβατικός. Για τη δήλωση του αντίστροφου ρόλου και τη δήλωση του ρόλου ως υπορόλου ενός πατέρα ρόλου χρησιμοποιούνται οι λέξεις κλειδιά *inverse* και *parent* ακολουθούμενες από τους αντιστοίχους ρόλους.

Τα αξιώματα που διαθέτει η γνώση μας είναι το αξίωμα της υπαγωγής και το αξίωμα της ισότητας. Η σύνταξη τους γίνεται με τη χρήση των λέξεων κλειδιά *implies* και *equivalent*, την έννοια που ορίζεται, η οποία πρέπει να είναι μια ατομική έννοια και την έννοια που δίνει την ιδιότητα της έννοιας που ορίζεται, η οποία μπορεί να είναι μια ατομική έννοια ή μια έννοια που προήλθε από ένα αποτέλεσμα των κατασκευαστών της f-SHIN.

Για παράδειγμα, αν θελήσουμε να ορίσουμε στο FiRE το αξίωμα της 'ΩριμηΓυναίκα που δόθηκε παραπάνω θα γράφαμε:

(equivalent ΩριμηΓυναίκα (some περιέχει (and Πρόσωπο (some περιέχει (and Παρειά (all περιέχει (and ΠαρειέςΓέρου))) (some περιέχει (and Μέτωπο (at-least 2έχειΡυτίδα))) (some περιέχει (not Γενειάδα))))))

Όσον αναφορά στους κατασκευαστές της f-SHIN [124], αυτοί συντάσσονται ως εξής: *top* (Καθολική έννοια), *bottom* (Κενή έννοια), and $C_1 \dots C_n$ (Σύζευξη $C_1 \dots C_n$), or $C_1 \dots C_n$ (Διάζευξη $C_1 \dots C_n$), not C ('Αρνηση), some R C (Υπαρξιακός περιορισμός), all R C (Περιορισμός τιμής), at-least p R (Περιορισμός το λιγότερο), at-most p R (Περιορισμός το πολύ).

Η δήλωση του σώματος ισχυρισμών (ABox) γίνεται με τη χρήση των λέξεων κλειδιά *instance* και *related*. Πιο συγκεκριμένα για να δηλώσουμε πως μια έννοια συμμετέχει σε ένα άτομο κατά ένα βαθμό χρησιμοποιούμε την λέξη κλειδί *instance*, ακολουθούμενη από το άτομο στο οποίο η έννοια συμμετέχει, την έννοια, το σύμβολο ανισότητας που υποδηλώνει εάν ο ισχυρισμός είναι θετικός ή αρνητικός και το βαθμό συμμετοχής. Το άτομο σε αυτή την υλοποίηση πρέπει να έχει δηλωθεί μαζί με το σώμα ορολογίας. Η συσχέτιση δυο ατόμων με ένα ρόλο γίνεται με τη χρήση της λέξης κλειδί *related* ακολουθούμενη από τα άτομα που συνδέονται και το ρόλο που τα συνδέει, για το βαθμό συμμετοχής.

Στο παράδειγμα ισχυρισμών που ακολουθεί, τα άτομα *segment1* και *segment3* συμμετέχουν στις έννοιες Πρόσωπο και Μέτωπο αντίστοιχα με βαθμό συμμετοχής 1

ενώ το `segment2` είναι με βαθμό 0.7 δηλαδή το `segment2` είναι Μαλλιά με βεβαιότητα 0.7. Τέλος φαίνονται και οι σχέσεις του ατόμου `filename` με τα `segment1`, `segment2` και `segment3`.

```
( instance segment1 Πρόσωπο)
( instance segment2 Μαλλιά ≥ 0.7)
( instance segment3 Μέτωπο)
( related filename segment1 περιέχει)
( related filename segment2 περιέχει)
( related filename segment3 περιέχει)
```

Η μηχανή ασαφούς συμπερασματολογίας FiRE υποστηρίζει την ασαφή περιγραφική λογική *f-SHLN* και συνεπώς τις υπηρεσίες εξαγωγής συμπερασμάτων της. Η μηχανή ασαφούς συμπερασματολογίας μπορεί να απαντήσει σε ερώτηση για το εάν ένα άτομο συμμετέχει κατά ένα βαθμό σε μια έννοια σύνθετη ή απλή. Οι ερωτήσεις αυτές συντάσσονται ως εξής:

```
( instance individualName ConceptName)
```

Για παράδειγμα:

```
( instance segment1 Πρόσωπο ≥ 0.9 )( instance segment2 Μαλλί ≥ 0.8 )
```

Τα παραπάνω παραδείγματα ρωτούν αν το άτομο `segment1` συμμετέχει στην έννοια `Πρόσωπο` \hat{u} κατά βαθμό 0.9 ενώ το άτομο `segment2` συμμετέχει στην έννοια `Μαλλί` \hat{u} κατά βαθμό 0.8 αντίστοιχα. Επιπλέον η μηχανή ασαφούς συμπερασματολογίας μπορεί να ελέγξει εάν μια έννοια είναι υποέννοια μιας άλλης έννοιας.

Για παράδειγμα: (ΚεφαλήΜεΚώμη) (and Πρόσωπο (some περιέχει Μαλλί))

Αντίστοιχα με τις ερωτήσεις ικανοποιησιμότητας μια έννοιας υπάρχουν και οι ερωτήσεις μέγιστου ελάχιστου φράγματος οι οποίες υπολογίζουν το βαθμό συμμετοχής ενός ατόμου σε μια έννοια. Η σύνταξη του είναι παρόμοια με την σύνταξη των ερωτήσεων ικανοποιησιμότητας μιας έννοιας μόνο που δεν χρησιμοποιείται βαθμός συμμετοχής.

Τέλος η σύνταξη των συνδυαστικών ερωτήσεων κατώτατων ορίων (*conjunctive threshold queries*) αποτελείται από δύο μέρη, το πρώτο που ορίζει τα άτομα που θα υπολογιστούν ενώ το δεύτερο δηλώνει ότι τις συνθήκες που πρέπει να ισχύουν για τα άτομα. Η παρακάτω ερώτηση ζητά ένα άτομο A , το οποίο πρέπει να συμμετέχει στην έννοια "Πρόσωπο" και να περιέχει 2 segments x και y τα οποία συμμετέχουν στις έννοιες "Μαλλιά" και "Γενειάδα" με τους αντίστοιχους βαθμούς:

```
 $A < \text{Πρόσωπο} (A) \text{ περιέχει} (A, x) \text{ Κώμη} (x) > 0.5 \text{ περιέχει} (x) > 0.5 \text{ Γενειάδα} (y) \geq 0.5$ 
```

5.5 Πειραματικά αποτελέσματα

Στηριζόμενοι στις έννοιες που δύναται να εξάγει ο αλγόριθμος επεξεργασίας εικόνας (Ενότητα 4.4.2.1) δημιουργείται όπως αναφέρθηκε το σώμα ισχυρισμών (Ενότητα 4.4.3.1) το οποίο και χρησιμοποιείται από τη μονάδα αναπαράστασης γνώσης και συλλογιστική (Σχήμα 4.5). Για τη δημιουργία της γνώσης (TBox) είχε γίνει η υπόθεση ότι ένα Πρόσωπο θα σχετίζεται με τμήματα με τον ρόλο "περιέχει". Έτσι λοιπόν με τη χρήση των ασαφών κανόνων-ερωτήσεων (Ενότητα 5.4.1) μπορούμε να κάνουμε ερωτήσεις για Πρόσωπα που περιέχουν διαφορετικά τμήματα και με

διαφορετικούς βαθμούς συμμετοχής. Ενδεικτικά παραδείγματα δίνονται στον Πίνακα 5.2

Κατηγορία	Ερώτημα
Χριστός	$A \equiv \text{Πρόσωπο}(A) \wedge \text{περιέχει}(A,x) \wedge \text{Κώμη}(x) > 0.9 \wedge \text{περιέχει}(A,y) \wedge \text{Γένυ}(y) \geq 0.6$
Παναγία	$A \equiv \text{Πρόσωπο}(A) \wedge \text{περιέχει}(A,x) \wedge \text{Κώμη}(x) > 0.1 \wedge \text{περιέχει}(A,y) \wedge \text{Γένυ}(y) \geq 0$
Αγία Αικατερίνη	$A \equiv \text{Πρόσωπο}(A) \wedge \text{περιέχει}(A,x) \wedge \text{Κώμη}(x) > 0.7 \wedge \text{περιέχει}(A,y) \wedge \text{Γένυ}(y) \geq 0$
Άγιος Λουκάς	$A \equiv \text{Πρόσωπο}(A) \wedge \text{περιέχει}(A,x) \wedge \text{Κώμη}(x) > 0.7 \wedge \text{περιέχει}(A,y) \wedge \text{Γένυ}(y) \geq 0.5$

Πίνακας 5.2: Ασαφή ερωτήματα για τη μηχανή συμπερασματολογίας

Η απόδοση του συστήματος αξιολογείται με το μέτρο της ορθότητας και ανάκλησης. Στον Πίνακα 5.3 παρουσιάζονται τα αποτελέσματα για 20 από τις 50 κατηγορίες εικόνων που διαθέτει το σύνολο δεδομένων μας (Ενότητα 4.4.3.2). Είναι φανερό ότι όσο περισσότερα χαρακτηριστικά του προσώπου περιέχονται στο ABox τόσο πιο βελτιωμένο και ακριβές είναι το συμπέρασμα της μονάδας αναπαράστασης γνώσης και συλλογιστική. Αντίθετα, η απόδοση της μεθόδου μας μειώνεται όταν οι μορφές των Αγίων έχουν όμοια χαρακτηριστικά π.χ. οι περισσότερες Αγίες φορούν μαντήλι. Σε αυτή την περίπτωση, θα μπορούσαμε να χρησιμοποιήσουμε και άλλα στοιχεία που είναι χαρακτηριστικά για την κάθε μορφή όπως αντικείμενα που κρατά (π.χ. σταυρούς, ευαγγέλια, πάπυροι), ειδικά ενδύματα (αρχιερέα, σαρτιωτικού Αγίου κ.τ.λ.), αναγνωριστικά όπως τα τρία αστέρια που διακοσμούν το μαντήλι της Παναγίας. Η ανάλυση των χαρακτηριστικών του ενδύματος καθώς και των παρελκόμενων στοιχείων που διαθέτουν οι μορφές ξεφεύγει από το πλαίσιο αυτής της διατριβής.

Το πεδίο των συστημάτων γνώσης στο χώρο της βυζαντινής αγιογραφίας είναι καινούργιο και επιτρέπει πλήθος εναλλακτικών λύσεων και προσεγγίσεων. Η εστίαση στους κανόνες του Διονυσίου εκ Φουρνά πιστεύουμε ότι αποτελεί αρκετά καλή βάση για την εξέλιξη και προσαρμογή του παρόντος συστήματος σε πιο εξελιγμένες λειτουργίες (αναγνώριση τεχνοτροπίας, χρονολογική κατάταξη κλπ).

α.α.	Κατηγορία	Σύνολο δεδομένων /κατηγορία	FiRE Ορθά Εκτιμώμενα	FiRE Συνολικά Εκτιμώμενα	Ορθότητα	Ανάκληση
1	Χριστός	70	58	67	0,87	0,83
2	Παναγία	60	43	54	0,80	0,72
3	Αγ.Πέτρος	50	38	44	0,86	0,76
4	Αγ.Παύλος	50	39	44	0,89	0,78
5	Αγ.Αικατερίνη	40	30	35	0,86	0,75
6	Αγ.Ιωάννης	40	29	37	0,78	0,73
7	Αγ.Λουκάς	40	28	33	0,85	0,70
8	Αγ.Ανδρέας	40	29	33	0,88	0,73
9	Αγ.Στέφανος	30	24	27	0,89	0,80
10	Αγ.Κων/νος	40	30	36	0,83	0,75
11	Αγ.Δημήτριος	40	31	36	0,86	0,78
12	Αγ.Γεώργιος	40	32	37	0,86	0,80
13	Αγ.Ελένη	40	31	37	0,84	0,78
14	Αγ.Πελαγία	20	15	17	0,88	0,75
15	Αγ.Νικόλαος	40	33	38	0,87	0,83
16	Αγ.Βασίλειος	40	31	37	0,84	0,78
17	Αγ.Αντώνιος	30	20	27	0,74	0,67
18	Αγ.Ευθύμιος	25	18	23	0,78	0,72
19	Αγ.Θωμάς	35	26	32	0,81	0,74
20	Αγ.Μηνάς	20	14	17	0,82	0,70

Πίνακας 5.3: Αξιολόγηση συστήματος με τη χρήση των μέτρων ορθότητας και ανάκλησης

Κεφάλαιο 6

Ασφάλεια Πολυμεσικού Περιεχομένου

6.1 Εισαγωγή

Λόγω της ταχύτατης ανάπτυξης της τεχνολογίας εξ αποστάσεως εκπαίδευσης είναι επιτακτική η ανάγκη ασφαλούς ανταλλαγής πληροφοριών και ιδιαίτερα προστασίας της πνευματικής ιδιοκτησίας ψηφιακών αρχείων όπως φωτογραφίες, εικόνες, βίντεο και αρχεία ήχου.

Σε αυτό το πλαίσιο ασχοληθήκαμε με τους τρόπους που θα μπορούσαμε να προστατεύσουμε την ιδιοκτησία πολυμεσικών αρχείων. Θελήσαμε να δημιουργήσουμε ένα σύστημα το οποίο θα είναι ικανό να προσφέρει προστασίας σε πολυμεσικά αρχεία χωρίς να γίνεται αντιληπτό από τρίτους και παράλληλα σε περίπτωση υποκλοπής αυτών των αρχείων να μπορεί να αποδειχθεί η γνησιότητά του ακόμη και αν έχουν αλλοιωθεί τα δεδομένα.

Η έρευνά μας οδηγήθηκε στη μελέτη τεχνικών απόκρυψης δεδομένων και συγκεκριμένα εξετάσαμε τεχνικές στενογραφίας, κρυπτογραφίας και υδατογράφηση πολυμεσικού υλικού.

Οι τεχνικές στεγανογραφίας (στεγανό + γράφειν) ουσιαστικά επιδιώκουν την απόκρυψη της πληροφορίας μέσα στα δεδομένα, προφυλάσσοντάς την, μέσα σε κάποιο "στεγανό", τέτοιο ώστε να μη κινεί υποψίες για την ύπαρξη κρυμμένου μηνύματος. Στόχος αυτών των μεθόδων είναι η διατήρηση της κρυφής επικοινωνίας δίχως να λαμβάνουν υπόψη το ενδεχόμενο επίθεσης σε αυτήν. Αλλά τα ψηφιακά δεδομένα μπορούν πολύ εύκολα να διαστρεβλωθούν (π.χ.να συμπιεστούν), δυσκολεύοντας τα συστήματα στεγανογραφία στο να ανταποκριθούν αξιόπιστα σε τέτοια ενδεχόμενα.

Οι τεχνικές κρυπτογραφίας (κρυπτός + γράφειν) ασχολούνται με την προστασία της πληροφορίας μέσω κωδικοποίησης η οποία καθιστά ακατανόητη την πληροφορία δίχως τη χρήση ειδικού κλειδιού. Αρχικά, οι τεχνικές κρυπτογραφίας περιορίζονταν στην επεξεργασία της γλωσσική δομή της πληροφορίας που περιείχε ένα κείμενο. Αργότερα, άρχισε να δίνεται έμφαση σε διάφορα πεδία των μαθηματικών, όπως διακριτά μαθηματικά, θεωρία αριθμών, θεωρία πληροφορίας, υπολογιστική πολυπλοκότητα, στατιστική και συνδυαστική ανάλυση με σκοπό την κρυπτογράφηση δεδομένων κατά τέτοιο τρόπο ώστε να διασφαλίζεται το ότι η πληροφορία που θα διαβαστεί από μη εξουσιοδοτημένους χρήστες θα είναι άχρηστη και ακατανόητη ή παραπλανητική. Οι εξουσιοδοτημένοι χρήστες έχουν τη δυνατότητα να προσπελάσουν την προστατευμένη πληροφορία με τη χρήση κλειδιού αποκρυπτογράφησης. Δυστυχώς όμως η μη προστατευμένη πληροφορία στην οποία έχουν πρόσβαση οι εξουσιοδοτημένοι χρήστες μπορεί να τη διοχετεύσουν και σε

μη εξουσιοδοτημένους χρήστες. Άρα η πληροφορία είναι προστατευμένη μόνο όσο είναι κρυπτογραφημένη. Αυτή η αδυναμία της κρυπτογραφίας αντιμετωπίστηκε με την υδατογραφία η οποία είναι σε θέση να προστατεύει την πληροφορία ακόμη και μετά από την αποκρυπτογράφηση της.

Η υδατογράφηση (watermarking) έχει ως στόχο την απόκρυψη πληροφορίας μέσα σε πολυμεσικά αρχεία μέσω ανεπαίσθητης τροποποίησης, είτε της αρχικής τους μορφής, είτε κάποιας μετασχηματισμένης εκδοχής τους. Το υδατογράφημα είναι μόνιμα συνδεδεμένο με την αρχική πληροφορία του πολυμεσικού αρχείου και κάθε πιθανή προσπάθεια αλλοίωσης του θα πρέπει να οδηγεί σε αλλοίωση όλου του πολυμεσικού αρχείου. Τα κύρια χαρακτηριστικά ενός υδατογραφήματος είναι η διαφάνεια, η χωρητικότητα και η ευρωστία του. Ένα σύστημα υδατογράφησης θεωρείται αξιόπιστο όταν το υδατογράφημα του δε γίνεται αντιληπτό από τον παρατηρητή.

Η διαφάνεια (perceptual transparency) αφορά στην απαίτηση ή μη του να είναι αντιληπτό το υδατογράφημα που ενσωματώνεται στα αρχικά δεδομένα. Σε κάποιες εφαρμογές είναι επιθυμητή η αναγνωρισιμότητα ενός υδατογραφήματος (π.χ. αποκλειστικά πλάνα ενός τηλεοπτικού σταθμού) και σε άλλες όχι (π.χ. σε ένα ψηφιακό πίνακα ζωγραφικής). Η χωρητικότητα (capacity) αναφέρεται στην ποσότητα της πληροφορίας που μπορεί να ενσωματωθεί στην κύρια πληροφορία δηλαδή το μέγεθος σε bits του υδατογραφήματος που μπορεί να ενσωματωθεί στην κύρια πληροφορία-υλικό που θέλουμε να υδατογραφήσουμε.

Η ευρωστία (robustness) ενός υδατογραφήματος έχει σχέση με το ποσοστό επιβίωσης του στα πολυμεσικά αρχεία στα οποία έχει ενσωματωθεί. Η ανθεκτικότητα κρίνεται έπειτα τόσο από κοινές ενέργειες επεξεργασίας σήματος όσο και από κακόβουλες επιθέσεις. Τα εμφανή (non-blind) σχήματα υδατογράφησης είναι αυτά που προϋποθέτουν τη γνώση των αρχικών δεδομένων για την εξαγωγή του υδατογραφήματος και χαρακτηρίζονται από μεγάλη ανθεκτικότητα σε επιθέσεις αλλά όχι και από πρακτικότητα. Σε αντίθεση, τα τυφλά (blind) σχήματα υδατογράφησης είναι πιο ευέλικτα αφού μόνο που χρειάζεται είναι η γνώση του υδατογραφήματος.

Η διαφάνεια, η χωρητικότητα και η ευρωστία είναι τρεις αντικρουόμενες παράμετροι που χρειάζεται να αντισταθμιστούν (trade off) ανάλογα με τις ιδιαιτερότητες της εφαρμογής. Για παράδειγμα, το υδατογράφημα σε ένα έργο τέχνης δε θα θέλαμε να επηρεάζει την πιστότητα του έργου άρα θα θέλαμε διαφάνεια και ευρωστία και συνεπώς περιορισμό στη χωρητικότητα ενώ το υδατογράφημα στα αποκλειστικά πλάνα ενός τηλεοπτικού σταθμού θα θέλαμε να είναι εμφανές δηλαδή θα θέλαμε μεγαλύτερη χωρητικότητα και ευρωστία ενώ δε θα μας ενδιέφερε καθόλου η διαφάνεια.

Η ασφάλεια είναι ένα επιπλέον χαρακτηριστικό ενός συστήματος υδατογράφησης που αναφέρεται στην ικανότητά του να αντιστέκεται σε κακόβουλες επιθέσεις. Οι επιθέσεις αυτές συνήθως προσπαθούν να αφαιρέσουν το υδατογράφημα, επιθέσεις όπως συμπύκνωση, γραμμικό και μη γραμμικό φιλτράρισμα, βελτίωση εικόνας, κβαντισμός (quantization), κοπή τμήματος (cropping), περιστροφή (rotation) και αλλαγή μεγέθους (resizing).

Τα χαρακτηριστικά αυτά καθιστούν τη ψηφιακή υδατογράφηση ιδιαίτερα πολύτιμη για την εξασφάλιση της εγκυρότητας στη διακίνηση ψηφιακών δεδομένων πολυμεσικού τύπου μέσω του Παγκόσμιου Ιστού και επιπλέον θεωρείται ως μια σύγχρονη μέθοδος για την προστασία πολυμεσικών αρχείων. Σε αυτό το πλαίσιο, στη συνέχεια θα παρουσιασθεί η μέθοδος υδατογράφησης που προτείναμε για την προστασία

σημασιολογικού υλικού.

6.2 Ανασκόπηση τεχνικών υδατογράφησης πολυμεσικών αρχείων

Λόγω της ευρείας χρήσης της υδατογράφησης ως τεχνικής προστασίας των δεδομένων και ιδίως εικόνων από κακόβουλες επιθέσεις, στη βιβλιογραφία υπάρχει πλούσια συλλογή σχετικών εργασιών. Όλες οι τεχνικές ακολουθούν ένα συγκεκριμένο μοντέλο που περιλαμβάνει το στάδιο της ενσωμάτωσης του υδατογραφήματος στα δεδομένα και το στάδιο της εξαγωγής του. Στο στάδιο της ενσωμάτωσης, εισάγεται το υδατογράφημα στα δεδομένα.

Στο παρελθόν έχουν προταθεί πολλοί αλγόριθμοι που επικεντρώνονται σε εικόνες ή πλαίσια (frames) από βίντεο. Αυτοί οι αλγόριθμοι μπορούν να χωριστούν σε αυτούς που εφαρμόζονται στο πεδίο του χώρου και σε αυτούς που εφαρμόζονται στο πεδίο της συχνότητας. Οι μέθοδοι στο πεδίο του χώρου ενσωματώνουν το υδατογράφημα απευθείας στις τιμές των εικονοστοιχείων και χαρακτηρίζονται από χαμηλή διαφάνεια, ευρωστία και χωρητικότητα όπως επίσης και από χαμηλή ασφάλεια. Σε αντίθεση, οι μέθοδοι που εφαρμόζονται στο πεδίο της συχνότητας, προσθέτουν το υδατογράφημα στις τιμές που προκύπτουν από το μετασχηματισμό. Χαρακτηρίζονται από ευρωστία, διαφάνεια και αρκετά μεγάλη χωρητικότητα ενώ η ασφάλεια βρίσκεται σε υψηλό επίπεδο.

Στις πρώτες τεχνικές τα υδατογραφήματα εισάγονταν στα λιγότερο σημαντικά ψηφία (LSBs) των εικονοστοιχείων των εικόνων [136]. Παρόλα αυτά, αυτή η τεχνική καθώς και κάποιες άλλες προτεινόμενες βελτιώσεις [97], [143] παρουσιάζουν σχετικά χαμηλή χωρητικότητα και δεν είναι ανθεκτική σε συμπίεση, σε κατάτμηση τμήματος της εικόνας (cropping) και άλλες επιθέσεις.

Μια από τις πιο διαδεδομένες τεχνικές είναι αυτή που πρότειναν ο Cox και οι συνεργάτες του [38]. Εφάρμοσαν διακριτό μετασχηματισμό συνημιτόνου DCT σε ολόκληρη την εικόνα και έπειτα πρόσθεσαν στους 1000 συντελεστές χαμηλής συχνότητας το υδατογράφημα που αποτελείται από μία ακολουθία τύπου i.i.d (independent and identically distributed) με Γκαουσιανή κατανομή 1000 bits. Η ανίχνευση του υδατογραφήματος προέκυψε από τη συσχέτιση της ακολουθίας αυτής με τους υδατογραφημένους συντελεστές DCT αφού αφαιρεθούν από αυτούς οι DCT συντελεστές της αρχικής εικόνας. Στην εργασία [65] οπτικά αναγνωρίσιμα υδατογραφήματα εισάγονται στην εικόνα, τροποποιώντας επιλεκτικά τους DC συντελεστές του DCT μετασχηματισμού σε περιοχές 8x8.

Στην εργασία [78] εφαρμόζεται αρχικά σε τυχαίες περιοχές 8x8 της εικόνας μετασχηματισμός DCT, ακολουθεί κβαντοποίηση όπως αυτή που εφαρμόζεται στο jpeg πρότυπο και τέλος επιλέγεται ένα ζεύγος συντελεστών DCT μεσαίων συχνότητων για ενσωμάτωση 1 bit υδατογραφήματος. Στην εργασία [128] χρησιμοποιείται ένας αλγόριθμος σημασιολογικής κατάτμησης βασισμένος στην πληροφορία κίνησης. Το υδατογράφημα αντιμετωπίζεται ως μία ψευδοτυχαία ακολουθία που εισάγεται στους DCT συντελεστές μιας ακολουθίας video με ανάλυση του εικονοπλαισίου σε περιοχές μεγέθους 8x8 εικονοστοιχείων. Βέβαια πρέπει να σημειωθεί ότι τα αντικείμενα που εντοπίζονται είναι κινούμενες περιοχές, με όλες τις συνέπειες αστοχίας εντοπισμού (μη-συμπαγή αντικείμενα, κίνηση του φόντου κλπ.). Επίσης η τεχνική αυτή δεν μπορεί να εφαρμοστεί σε εικόνες (έλλειψη κίνησης) και σε

σκηές video με ελάχιστη κίνηση.

Άλλες προσεγγίσεις [64], [155], χρησιμοποιούν το διακριτό μετασχηματισμό Wavelet (Discrete Wavelet Transform - DWT) για να αποκρύψουν υδατογραφήματα στο συχνοτικό περιεχόμενο της εικόνας έχοντας ως υδατογράφημα ψευδοτυχαίες (pseudorandom) ακολουθίες. Τα προτεινόμενα συστήματα είναι προσανατολισμένα σε ολόκληρες εικόνες και δεν λαμβάνουν υπόψη τα σημασιολογικά αντικείμενα τα οποία συνήθως συνθέτουν το περιεχόμενο των πλαισίων. Ο Xia και οι συνεργάτες του πρότειναν την υδατογράφιση των συντελεστών DWT των υψηλών και μεσαίων συχνοτήτων μιας και αυτοί οι συντελεστές αντιστοιχούν σε περιοχές σύνθετης υψής όπου οι μικρές μεταβολές δεν είναι αντιληπτές από το μάτι. Το υδατογράφημα αποτελείται από μια ψευδοτυχαία ακολουθία Gauss αλλά για την ανίχνευσή του απαιτείται η γνώση της αρχικής εικόνας.

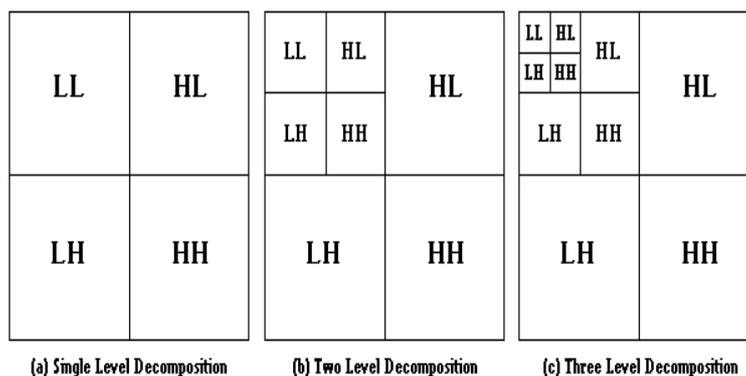
Οι Kundur και Hatzinakos [81] πρότειναν μια τεχνική σύντηξης τόσο των συντελεστών DWT της εικόνας όσο και του υδατογραφήματος, ρυθμίζοντας την ένταση βάσει της ευαισθησία του ματιού. Για την ανίχνευση του απαιτείται τη γνώση της αρχικής εικόνας. Στην εργασία [142] προτείνεται ένα σχήμα υδατογράφισης όπου ο καθορισμός των σημαντικών συντελεστών κυματιδίων βασίζεται σε ένα σχήμα διαδοχικής κβαντοποίησης υποζωνών και με χρήση συσχέτισης γίνεται ο εντοπισμός του υδατογραφήματος. Οι Zeng και Liu [152] πρότειναν η χρήση ψηφιακής πιστοποίησης των αρχείων που υδατογραφούνται με κατάλληλα επιλεγμένη ψευδοτυχαία ακολουθία. Στην εργασία [147] προτείνεται ένα τυφλό υδατογραφικό σχήμα το οποίο χρησιμοποιεί DWT μετασχηματισμό σε συνδυασμό με μια τεχνική διαδοχικής κβαντοποίησης υποζωνών και η ανίχνευση του υδατογραφήματος πραγματοποιείται με χρήση συσχέτισης.

Παρόλα αυτά στη βιβλιογραφία είναι λίγα τα προτεινόμενα συστήματα που είναι προσανατολισμένα σε υδατογράφιση σημασιολογικών αντικειμένων. Το σύστημα που παρουσιάζεται στην εργασία [148] προσθέτει ψηφιακά υδατογραφήματα σε αντικείμενα, χρησιμοποιώντας το διδιάστατο ή τρισδιάστατο προσαρμοσμένο στο σχήμα διακριτό μετασχηματισμό wavelet (shape adaptive discrete wavelet transform - SA-DWT) με υποστήριξη περιοχών αυθαίρετης δομής (Arbitrary Regions Of Support - AROS). Το υδατογράφημα σε αυτή την περίπτωση είναι μία i.i.d ακολουθία με Γκαουσιανή κατανομή, το οποίο προστίθεται σε όλες τις υπερπερατές περιοχές συχνοτήτων του αντικειμένου στο συχνοτικό χώρο των wavelets.

Στην εργασία [19], η μέθοδος εισαγωγής του υδατογραφήματος εκμεταλλεύεται το σχήμα του αντικειμένου. Το υδατογράφημα μπορεί να είναι μία ψευδοτυχαία ακολουθία μετασχηματισμένη έτσι ώστε να ταιριάζει στην κλίμακα και τον προσανατολισμό των αντικειμένων. Εντούτοις και οι δύο προηγούμενες τεχνικές δεν χρησιμοποιούν κάποιον αλγόριθμο σημασιολογικής κατάτμησης, θεωρώντας ότι τα αντικείμενα είναι διαθέσιμα. Στην πράξη βέβαια κάτι τέτοιο δεν ισχύει, στις περισσότερες περιπτώσεις τα σημασιολογικά αντικείμενα δεν είναι διαθέσιμα, ενώ η σημασιολογική κατάτμηση αντικειμένων με επίβλεψη είναι μια χρονοβόρα διαδικασία.

Τέλος στην εργασία [69] προτείνεται ένα ανάμεικτο σύστημα υδατογράφισης (cocktail watermarking) όπου πάλι το υδατογράφημα είναι μια i.i.d Γκαουσιανή κατανομή. Το προτεινόμενο σύστημα πραγματοποιεί χαμηλού επιπέδου κατάτμηση υψής, η οποία στις περισσότερες περιπτώσεις δεν παρέχει σημασιολογικά αντικείμενα. Άλλωστε σε όλες τις προηγούμενες εργασίες υδατογράφισης σημασιολογικών αντικειμένων το υδατογράφημα είναι μία ψευδοτυχαία ακολουθία.

Γενικά είναι πλούσια η βιβλιογραφία που αφορά στα συστήματα υδατογράφισης



Σχήμα 6.1: Πυραμιδική αποσύνθεση μιας εικόνας

εικόνων και υπό αυτή την έννοια ο κάθε ερευνητής καλείται να εστιάσει σε τεχνικές που είναι ικανές να ανταποκριθούν στα χαρακτηριστικά του προβλήματός του. Το δικό μας ενδιαφέρον στράφηκε σε τεχνικές που θα εξασφάλιζαν την ασφάλεια στη διακίνηση πολυμεσικού υλικού και συγκεκριμένα εικόνων καθώς και ψηφιακών έργων τέχνης σε συστήματα ηλεκτρονικής μάθησης. Σε αυτό το πλαίσιο παρουσιάζεται το προτεινόμενο σύστημα σημασιολογικής υδατογράφησης και πιστοποίησης πολυμεσικού υλικού.

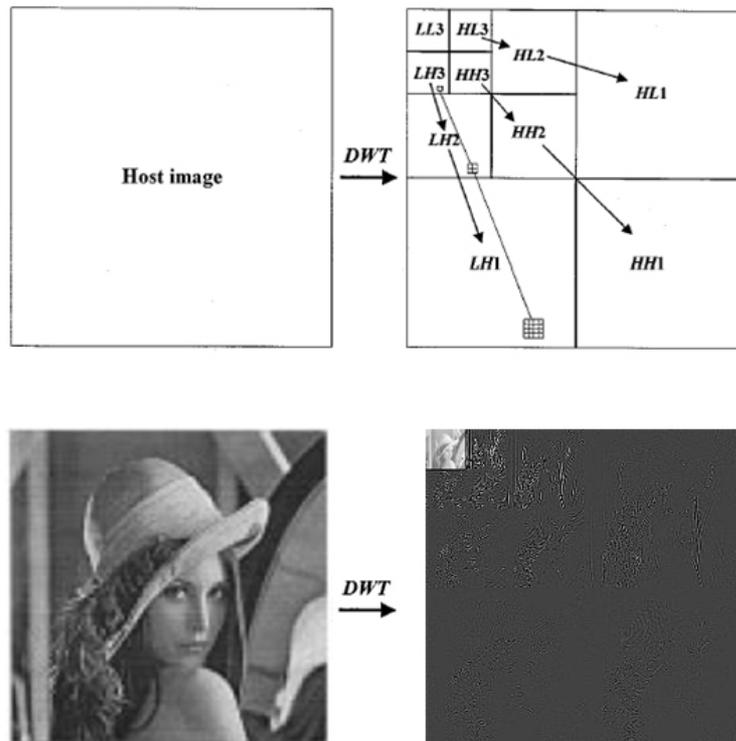
6.3 Μέθοδος σημασιολογικής υδατογράφησης πολυμεσικού περιεχομένου

Η προτεινόμενη μέθοδος υδατογράφησης αφορά στην υδατογράφηση σημασιολογικών αντικειμένων και στην πιστοποίηση τους με σκοπό την ασφαλή διακίνησή τους μέσα από διαδικτυακά περιβάλλοντα. Αρχικά σε κάθε σημασιολογικό αντικείμενο εφαρμόζεται μετασχηματισμός DWT αποσυνθέτοντάς το σε τρία επίπεδα δέκα υποζωνών συχνοτήτων. Έπειτα επιλέγεται το ζεύγος των υποζωνών που περιέχουν το υψηλό ενεργειακό περιεχόμενο και υπολογίζεται το σημαντικό δέντρο συντελεστών κυματιδίων για το συγκεκριμένο ζεύγος. Τέλος στους συντελεστές του σημαντικού δέντρου συντελεστών κυματιδίων προστίθεται το υδατογράφημα. Στη συνέχεια παρουσιάζεται αναλυτικά η προτεινόμενη μέθοδος.

6.3.1 Διακριτός μετασχηματισμός κυματιδίων

Ο διακριτός μετασχηματισμός κυματιδίων (Discrete Wavelet Transform) χρησιμοποιείται για αποτελεσματική κωδικοποίηση εικόνων. Ο DWT μετασχηματίζει τα δεδομένα της εικόνας, σε συντελεστές στο χώρο των υποζωνών συχνότητας, ενώ ταυτόχρονα διατηρεί τη χωρική συσχέτιση μεταξύ των υποζωνών συχνότητας.

Εφαρμόζοντας τον DWT μία φορά στην εικόνα, παράγονται τέσσερις υποζώνες συχνότητας όπως φαίνεται στο Σχήμα 6.1.α. Η υποζώνη συχνότητας LL_1 (HH_1) περιέχει χαμηλές (υψηλές) συχνοτικές συνιστώσες τόσο στην οριζόντια όσο και στην κατακόρυφη διεύθυνση, ενώ η υποζώνη συχνότητας HL_1 (LH_1) περιλαμβάνει υψηλές (χαμηλές) συχνοτικές κατά την οριζόντια διεύθυνση και χαμηλές (υψηλές) συχνοτικές κατά την κατακόρυφη διεύθυνση. Η υποζώνη συχνότητας LL_1 μπορεί με παρόμοιο τρόπο να αποσυντεθεί περαιτέρω σε τέσσερις νέες υποζώνες συχνότητας οι οποίες συμβολίζονται ως LL_2 , HL_2 , LH_2 και HH_2 . Αυτή η διαδικασία



Σχήμα 6.2: Η πυραμιδική αποσύνθεση τριών στρώματος μιας εικόνας

μπορεί να επαναληφθεί πολλές φορές με βάση την υπό εξέταση εφαρμογή. Ένα παράδειγμα αποσύνθεσης μιας εικόνας σε τρία επίπεδα με δέκα υποζώνες συχνότητας χρησιμοποιώντας το μετασχηματισμό DWT παρουσιάζεται στο Σχήμα 6.2.

Η ανακατασκευή της εικόνας μπορεί να πραγματοποιηθεί από τους συντελεστές DWT. Αυτή η ανακατασκευή καλείται αντίστροφος μετασχηματισμός IDWT. Αν με $I(m, n)$ παραστήσουμε την εικόνα, οι συντελεστές DWT και IDWT μπορούν να ληφθούν εφαρμόζοντας τους σε κάθε διάσταση m και n ξεχωριστά.

Τα Υψηλής Σημαντικότητας Δέντρα Συντελεστών Κυματιδίων (Qualified Significant Wavelet Trees) παράγονται κατά την εφαρμογή ενός DWT μετασχηματισμού σε μία εικόνα. Αυτά βασίζονται στην κωδικοποίηση Embedded Zerotree Wavelet (EZW) [118] που έχει παίξει σημαντικό ρόλο σε πρότυπα συμπίεσης εικόνας-video, όπως το JPEG2000 ή το MPEG4.

Η σχέση μεταξύ συντελεστών wavelet διαφορετικών κλιμάκων, που αντιστοιχούν στην ίδια θέση μπορούν να ορισθούν ως σχέση "γονέα-παιδιού". Οι υποζώνες συχνότητας LL3, LL2, LL1 ακολουθούν σχέση "γονέα-παιδί" όπως απεικονίζεται στο wavelet δέντρο του Σχήματος 6.1. Ο συντελεστής στο υψηλότερο επίπεδο καλείται "γονέας" και όλοι οι συντελεστές των χαμηλότερων επιπέδων που αντιστοιχούν στην ίδια χωρική θέση και έχουν παρόμοιο προσανατολισμό καλούνται "παιδιά". Για ένα δεδομένο "γονέα", το σύνολο όλων των συντελεστών σε όλα τα χαμηλότερα επίπεδα με παρόμοιο προσανατολισμό που αντιστοιχούν στην ίδια χωρική θέση καλούνται "απόγονοι".

Οι βασικοί ορισμοί των Υψηλής Σημαντικότητας Δέντρα Συντελεστών Κυματιδίων δίνονται στη συνέχεια:

Ορισμός 1: Ένας συντελεστής wavelet $x_n(i, j) \in D$ είναι "πατέρας" του συντελεστή $x_{n-1}(p, q) \in D$, όπου D είναι μία από τις υποζώνες συχνότητας HL_n , HL_n , HL_n , $p = 2i - 1$ ή $p = 2i$ και $q = 2j - 1$ ή $q = 2j$, όπου $n > 1$, $i > 1$ και

$j > 1$. Οι συντελεστές $x_{n-k}(p, q)$ καλούνται "απόγονοι" του $x_n(i, j)$, για $1 < k < n$.

Σύμφωνα με τον ορισμό των ZeroTrees [118] τίθεται ένα κατώφλι συχνότητας T και ελέγχεται αν ένας wavelet συντελεστής $x_n(i, j)$ ικανοποιεί τη σχέση $|x_n(i, j)| < T$. Έπειτα ελέγχεται αν και οι απόγονοί του ικανοποιούν αντίστοιχες σχέσεις δεδομένου του κατώφλιου. Αν ικανοποιούνται αυτές οι συνθήκες τότε αυτός ο συντελεστής θεωρείται ρίζα του ZeroTree. Τα ZeroTrees βασίζονται στην υπόθεση ότι αν ένας wavelet συντελεστής στην ισχυρότερη κλίμακα είναι ασήμαντος σύμφωνα με το δεδομένο κατώφλι, τότε και όλοι οι απόγονοι του θα είναι ασήμαντοι αν συγκριθούν με το δεδομένο κατώφλι.

Ορισμός 2: Αν ένας συντελεστής wavelet $x_n(i, j)$ και όλοι οι "απόγονοί" του $x_{n-k}(p, q)$ για $1 < k < n$, ικανοποιούν τις συνθήκες $|x_n(p, q)| < T$, $|x_{n-k}(p, q)| < T$ για δεδομένο κατώφλι T , όπου $p = 2i - 1$ ή $p = 2i$ και $q = 2j - 1$ ή $q = 2j$, όπου $n > 1$, $i > 1$ και $j > 1$, τότε το δέντρο $x_n \rightarrow x_{n-1} \rightarrow \dots \rightarrow x_{n-k}$ καλείται μηδενικό δέντρο συντελεστών wavelet.

Ορισμός 3: Αν ένας συντελεστής wavelet $x_n(i, j)$ ικανοποιεί τη συνθήκη $|x_n(p, q)| < T$, για δεδομένο κατώφλι T , τότε ο συντελεστής $x_n(i, j)$ καλείται σημαντικός wavelet συντελεστής.

Ορισμός 4: Αν ένας συντελεστής wavelet $x_{n-1}(p, q) \in D$, όπου D είναι μία από τις υποζώνες συντελεστών wavelet HL_n, LH_n, HH_n , ικανοποιεί τη συνθήκη $|x_n(p, q)| > T_1$ και τα "παιδιά" του $x_{n-k}(p, q)$ τα οποία είναι συντελεστές wavelet ικανοποιούν τη συνθήκη $|x_{n-k}(p, q)| > T_2$, για δεδομένα κατώφλια T_1 και T_2 , τότε οι συντελεστής wavelet $x_n(i, j)$ και $x_{n-k}(p, q)$ καλούνται υψηλής σημαντικότητας δέντρο συντελεστών wavelet (QSWT).

Τα Υψηλής Σημαντικότητας Δέντρα Συντελεστών Κυματιδίων μπορούν να χρησιμοποιηθούν για απόκρυψη πληροφορίας σε αρχεία πολυμέσων, αφού παρέχουν ευρωστία στις διαφόρου τύπου επιθέσεις όπως επίσης προκαλούν μη αντιληπτές μεταβολές στο υλικό καθώς η πληροφορία τοποθετείται στις θέσεις πολύ ισχυρών ακμών της εικόνας. Οι συντελεστές αυτών των δέντρων αντιπροσωπεύουν τους συντελεστές με το υψηλότερο ενεργειακό περιεχόμενο.

6.3.2 Δέντρα συντελεστών κυματιδίων

Για τον υπολογισμό Υψηλής Σημαντικότητας Δέντρα Συντελεστών Κυματιδίων προτείνουμε την εξής μέθοδο: αρχικά αποσυνθέτουμε την εικόνα σε τρία επίπεδα, με τη χρήση του διακριτού μετασχηματισμού κυματιδίων, σχηματίζοντας τρία ζευγάρια υποζωνών συχνοτήτων. Το πρώτο ζευγάρι περιέχει τις υποζώνες συχνοτήτων (HL_3, HL_2) και θα το συμβολίζουμε με P_1 , το δεύτερο ζευγάρι περιέχει τις υποζώνες συχνοτήτων (LH_3, LH_2) και θα το συμβολίζουμε με P_2 , ενώ το τρίτο ζευγάρι περιέχει τις υποζώνες συχνοτήτων (HH_3, HH_2) και θα το συμβολίζουμε με P_3 .

Στόχος της μεθόδου είναι η επιλογή του ζευγαριού υποζωνών συχνοτήτων που περιέχει το υψηλότερο ενεργειακό περιεχόμενο, σε σχέση με τα άλλα δύο ζευγάρια. Για αυτό το λόγο, υπολογίζεται η ενέργεια του κάθε ζεύγους. Επιλέγοντας το ζεύγος P_k , όπου $k = 1, 2, 3$, η ενέργεια του ζεύγους υπολογίζεται ως εξής:

$$E_{p_k} = \sum_i \sum_j [x_3(i, j)]^2 + \sum_p \sum_q [x_2(p, q)]^2, k = 1, 2, 3 \quad (6.1)$$

όπου $x_3(i, j) \in R_k, k = 1, 2, 3$ είναι ένας συντελεστής wavelet της αντίστοιχης

υποζώνης συχνότητας με $R_1 = HL_3$, $R_2 = LH_3$ και $R_3 = HH_3$. Όμοια για τους συντελεστές $x_2(p, q) \in S_k, k = 1, 2, 3$ με $S_1 = HL_2$, $S_2 = LH_2$ και $S_3 = HH_2$. Υπολογίζοντας κατά αυτό το τρόπο τις ενέργειες για όλα τα ζεύγη, επιλέγουμε το ζεύγος που μεγιστοποιεί την ενέργεια ως εξής:

$$\hat{k} = \operatorname{argmax}_{k=1,2,3} E_{P_k} \quad (6.2)$$

Έχοντας επιλέξει το ζεύγος υποζωνών συχνότητας με την υψηλότερη ενέργεια, υπολογίζεται το Υψηλής Σημαντικότητας Δέντρο Συντελεστών Κυματιδίων (QSWT). Αρχικά υπολογίζονται τα κατώφλια βάσει των οποίων θα επιλέξουμε τους ισχυρού περιεχομένου συντελεστές wavelet του ζεύγους υποζωνών συχνότητας με την υψηλότερη ενέργεια. Έτσι η αρχικοποίηση των τιμών των κατωφλίων για κάθε υποζώνη υπολογίζεται ως εξής:

$$\begin{aligned} T_1 &= \frac{1}{N_{P_3} * M_{P_3}} \sum_{i=1}^{M_{P_3}} \sum_{j=1}^{N_{P_3}} (x_3^2(i, j)), x_3(i, j) \in R_k \\ T_2 &= \frac{1}{2N_{P_2} * 2M_{P_2}} \sum_{i=1}^{2M_{P_2}} \sum_{j=1}^{2N_{P_2}} (x_2^2(i, j)), x_2(i, j) \in S_k \end{aligned} \quad (6.3)$$

όπου N_{P_2}, M_{P_2} είναι οι διαστάσεις της υποζώνης. Για να εξαχθούν οι συντελεστές QSWTs εφαρμόζεται αλγόριθμος του οποίου ο ψευδοκώδικας είναι:

```

t = 0
QSWT[t] = ∅
for i = 1 to NP2
for j = 1 to MP2
if x3(i, j) ≥ T1
if x2(2i - 1, 2j - 1) ≥ T2 and x2(2i - 1, 2j) ≥ T2 and x2(2i, 2j - 1) ≥ T2 and
x2(2i, 2j) ≥ T2 or [x2(2i - 1, 2j - 1) + x2(2i - 1, 2j) + x2(2i, 2j - 1) + x2(2i, 2j)]/4 ≥
T2
QSWT[t] = x3(i, j), x2(2i - 1, 2j - 1) + x2(2i - 1, 2j) + x2(2i, 2j - 1) + x2(2i, 2j)
t = t + 1
end if
end if
end for j
end for i
    
```

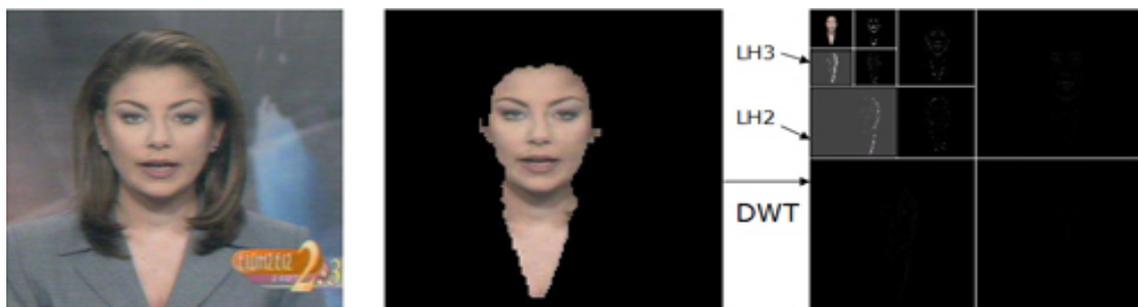
Το αποτέλεσμα αυτού του αλγορίθμου είναι η παραγωγή ενός πίνακα όπου στην πρώτη στήλη του περιέχει τους συντελεστές wavelet (γονέας) οι οποίοι ανήκουν στην υποζώνη R_k και έχουν τιμή $\geq T_1$ ενώ στις υπόλοιπες στήλες της ίδιας γραμμής περιέχονται οι συντελεστές wavelet (παιδιά) οι οποίοι ανήκουν στην υποζώνη S_k και έχουν τιμή $\geq T_2$.

6.3.3 Μέθοδος εισαγωγή υδατογραφήματος σε πολυμεσικό υλικό

Στην προτεινόμενη μέθοδο χρησιμοποιείται ως υδατογράφημα μία εικόνα που είναι οπτικά αναγνωρίσιμη, στην προκειμένη περίπτωση μια bitmap εικόνα η οποία απεικονίζει το λογότυπο του εργαστηρίου Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων (Image Video Multimedia Laboratory). Το πλεονέκτημα στη χρήση μιας οπτικά αναγνωρίσιμης εικόνας είναι η αναγνωρισιμότητά της ακόμα και μετά από επιθέσεις.

IVML

Σχήμα 6.3: Εικόνα-Υδατογράφημα



Σχήμα 6.4: Στάδια αποσύνθεσης σημασιολογικής εικόνας. Αρχικά εξάγεται το σημασιολογικό τμήμα της εικόνας, το οποίο αποσυνθέτεται με χρήση του μετασχηματισμού κυματιδίου

Αρχικά, εφαρμόζεται πάνω στην εικόνα ή στο πλαίσιο ενός βίντεο, η μέθοδος εντοπισμού χρώματος δέρματος (Ενότητα 4.4.2.1.1) με σκοπό την εξαγωγή της περιοχής της εικόνας που περιέχει σημαντικό περιεχόμενο για προστασία. Είσοδο αυτής της μεθόδου μπορεί να αποτελέσει η εικόνα(α) του Σχήματος 6.4. Ως έξοδο της εικόνας θα ληφθεί η εικόνα (β) του Σχήματος 6.4.

Εφαρμόζοντας τώρα στη σημασιολογική εικόνα του σχήματος 6.4, το DWT μετασχηματισμό πραγματοποιείται η αποσύνθεση της αρχικής εικόνας σε τρία ζεύγη υποζωνών συχνότητας. Έστω ότι έχει επιλεγεί ως ζεύγος υποζωνών συχνότητας με την υψηλότερη ενέργεια (σύμφωνα με τη διαδικασία που περιγράφηκε στην Ενότητα 6.3.2), το ζεύγος P2 (LH_3, LH_2) των υποζωνών δηλαδή όλα τα παιδιά $x_2(p, q)$ του γονέα $x_3(i, j)$ και έχει σχηματισθεί το Υψηλής Σημαντικότητας Δέντρο Συντελεστών Κυματιδίων (QSWT) για αυτό το ζεύγος.

Θεωρώντας τώρα ότι το υδατογράφημα έχει μέγεθος $a \times b$, θα επιλεγούν οι $a \times b$ υψηλότεροι συντελεστές QSWTs με σκοπό την εισαγωγή του υδατογραφήματος σε αυτούς. Για το λόγο αυτό η εικόνα-υδατογράφημα ταξινομείται κατά φθίνουσα σειρά.

Στη συνέχεια εφαρμόζεται ο ακόλουθος αλγόριθμος με σκοπό την εισαγωγή των τιμών του υδατογραφήματος στους συντελεστές των γραμμών $a \times b$ του QSWT:

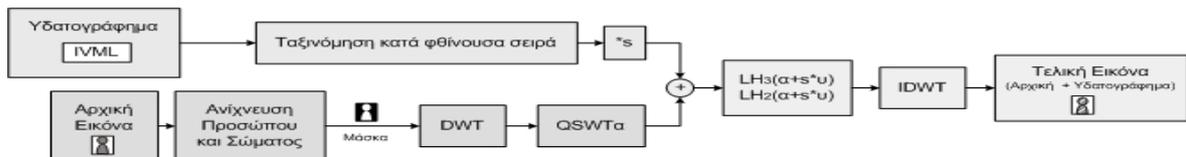
```

for  $i = 1$  to  $a \times b$ 
 $x'_3(i, j) = x_3(i, j) + s \cdot w(k, l)$ 
 $x'_2(i, j) = x_2(i, j) + s \cdot w(k, l)$ 
endfor

```

όπου $x_3(i, j) \in R_k$ είναι ένας συντελεστής wavelet της υψηλής ενέργειας υποζώνης συχνότητας και $x_2(i, j) = \max(x_2(2i - 1, 2j - 1), x_2(2i - 1, 2j), x_2(2i, 2j - 1), x_2(2i, 2j)), x_2(i, j) \in S_k$ ενώ $w(k, l)$ είναι η εικόνα-υδατογράφημα και s είναι μια σταθερά κλίμακας.

Τέλος ο αντίστροφος διακριτός μετασχηματισμός κυματιδίων (IDWT) εφαρμόζεται σε όλα τα ζεύγη των υποζωνών συχνότητων έτσι ώστε να κατασκευασθεί η τελική εικόνα η οποία θα περιέχει και την υδατογραφημένη περιοχή προσώπου. Ένα διάγραμμα στο οποίο φαίνεται η μέθοδος που περιγράφηκε απεικονίζεται στο Σχήμα 6.5.



Σχήμα 6.5: Διάγραμμα μεθόδου εισαγωγής υδατογραφήματος

6.3.4 Μέθοδος εξαγωγής υδατογραφήματος από πολυμεσικό υλικό

Η μέθοδος εξαγωγής λαμβάνει ως είσοδο την αρχική εικόνα, την υπό εξέταση υδατογραφημένη εικόνα και το συντελεστή κλίμακας. Στόχος της μεθόδου είναι η εξαγωγή της εικόνας-υδατογράφημα και απόδειξη της γνησιότητας της υπό εξέταση υδατογραφημένης εικόνας. Προς αυτή την κατεύθυνση πραγματοποιούνται τα εξής βήματα:

Βήμα 1ο

Γίνεται η ανίχνευση προσώπου και σώματος από την αρχική εικόνα. Εξάγεται μια μάσκα η οποία χρησιμοποιείται για να προσδιορίσει και να εξάγει την περιοχή προσώπου και σώματος της υπό εξέταση εικόνας.

Βήμα 2ο

Η μάσκα που εξάγεται στο Βήμα 1 χρησιμοποιείται ως οδηγός για την υπό εξέταση εικόνα με σκοπό να εξαχθεί η περιοχή της εικόνας στην οποία το υδατογράφημα έχει εισαχθεί κατά τη διαδικασία της μεθόδου εισαγωγής του υδατογραφήματος.

Βήμα 3ο

Αποσυντίθενται τόσο η σημασιολογική εικόνα που εξάγεται από την αρχική εικόνα όσο κι αυτή που εξάγεται από την υπό εξέταση εικόνα σε τρία ζεύγη παρέχοντας δέκα υποζώνες συχνοτήτων με τη χρήση του DWT.

Βήμα 4ο

Εν συνεχεία, υπολογίζεται το QSWT της αρχικής εικόνας. Οι θέσεις μέσα στην αρχική εικόνα των συντελεστών με υψηλή στάθμη ενέργειας κατευθύνουν τη διαδικασία στο να επιλέγουν από τις αντίστοιχες θέσεις της υπό εξέτασης εικόνας, οι συντελεστές συχρότητας οι οποίοι εμπεριέχουν το υδατογράφημα.

Βήμα 5ο

Οι τιμές από τα δύο QSWT αφαιρούνται και το αποτέλεσμα της αφαίρεσης διαιρείται με την τιμή που είχε ο συντελεστής κλίμακας στη διαδικασία εισαγωγής του υδατογραφήματος στην αρχική εικόνα. Συγκεκριμένα οι τιμές του υδατογραφήματος εξάγονται ως εξής:

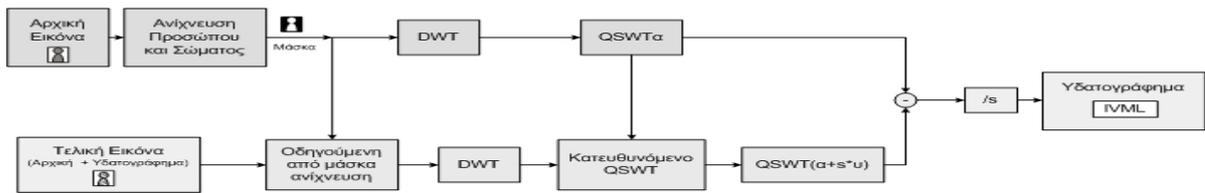
```

for  $i = 1$  to  $r \times s$ 
 $w_i^3 = (x_i'^3 - x_i^3)/s$ 
 $w_i^2 = (x_i'^2 - x_i^2)/s$ 
endfor

```

Για τις υπολογισμένες τιμές υδατογραφήματος (grayscale values) \hat{w}_3 και \hat{w}_2 αρχικά προσδιορίζεται ο μέσος όρος και έπειτα γίνεται αναδιάταξη για σχηματισμό της εικόνας υδατογραφήματος. Η αναδιάταξη πραγματοποιείται επειδή οι τιμές του υδατογραφήματος ταξινομούνται με βάση την τιμή προτού εισαχθούν στην εικόνα. Αυτό το βήμα μας οδηγεί στη λήψη του υδατογραφήματος και στην εξακρίβωση της γνησιότητας της υπό εξέταση εικόνας.

Μια γραφική αναπαράσταση της μεθόδου εξαγωγής υδατογραφήματος δίνεται στο Σχήμα 6.6



Σχήμα 6.6: Διάγραμμα μεθόδου εξαγωγής υδατογραφήματος

Σχήμα 6.7: Στάδια μεθόδου υδατογράφησης α) Αρχική εικόνα, β) Σημειολογικό αντικείμενο, γ) Εικόνα υδατογράφησης, δ) Υδατογραφημένη εικόνα

6.4 Πειραματικά Αποτελέσματα Μεθόδου Υδατογράφησης

Η αποτελεσματικότητα και ευρωστία του προτεινόμενου συστήματος υδατογράφησης αντικειμένων ακολουθιών βίντεο ή εικόνων δοκιμάστηκε κάτω από διαφορετικού τύπου επιθέσεις επεξεργασίας εικόνας, χρησιμοποιώντας το σημειολογικό αντικείμενο (Σχήμα 6.7.β) το οποίο έχει εξαχθεί από την αρχική εικόνα διαστάσεων 300x300 εικονοστοιχείων (Σχήμα 6.7.α). Στα πειράματα που διενεργήθηκαν χρησιμοποιήθηκε σαν υδατογράφημα ασπρόμαυρη (grayscale) εικόνα με μέγεθος 10x20 εικονοστοιχεία που περιείχε τη λέξη "IVML" και εισήχθη στο αντικείμενο που εξάχθηκε από την αρχική εικόνα. (Σχήμα 6.7.γ).

Στη συνέχεια και με βάση το μέγεθος το υδατογραφήματος τα 200 καλύτερα βασισμένα στο αντικείμενο υψηλής σημαντικότητας δέντρα συντελεστών wavelet επιλέγονται για εισαγωγή του υδατογραφήματος. Στα υδατογραφημένα αντικείμενα τα εισαχθέντα υδατογραφήματα δεν γίνονται αντιληπτά ούτε προκαλούν αντιληπτές παραμορφώσεις στα αντικείμενα όπως μπορεί κανείς να παρατηρήσει στο Σχήμα 6.7.δ.

Οι πίνακες που ακολουθούν περιλαμβάνουν τα αποτελέσματα εξαγωγής των υδατογραφημάτων από τα αντικείμενα (χωρίς επιθέσεις) χρησιμοποιώντας την προτεινόμενη μέθοδο. Στον ίδιο πίνακα δίνονται οι τιμές του PSNR (Peak Signal to Noise Ratio) των αντικειμένων μετά την εισαγωγή των υδατογραφημάτων. Στα επιτελεσθέντα πειράματα ο συντελεστής PSNR ορίζεται ως:

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{\alpha(V_o)} \sum_i \sum_j (V_o(i, j) - V_q(i, j))^2} \quad (6.4)$$

όπου $\alpha(\cdot)$ είναι συνάρτηση που επιστρέφει το εμβαδό (αριθμός εικονοστοιχείων) περιοχών αυθαιρέτου σχήματος και $V_o(i, j), V_q(i, j)$ συμβολίζουν τα εικονοστοιχεία των αντικειμένων V_o και V_q αντίστοιχα.

Στη συνέχεια, ερευνήθηκε η ευρωστία του προτεινόμενου συστήματος κάτω από διαφορετικές επιθέσεις όπως συμπίεση JPEG με απώλειες, θάμπωμα (blurring), όξυνση (sharpening) καθώς και συνδυασμό τους. Επιπρόσθετα, χρησιμοποιείται ένα αντικειμενικό κριτήριο συσχέτισης του εξαγόμενου με το εισαχθέν υδατογράφημα. Το κριτήριο αυτό είναι ο συντελεστής συσχέτισης. Γι' αυτό το λόγο ας συμβολίσουμε με w το διάνυσμα που περιέχει τις τιμές (grayscale values) του αυθεντικού υδατογραφήματος και με w' το διάνυσμα που περιέχει τις τιμές του εξαχθέντος υδατογραφήματος. Τότε ο κανονικός συντελεστής συσχέτισης ορίζεται ως:

Πίνακας 6.1: Εξαγωγή υδατογραφήματος από την περιοχή προσώπου και σώματος έπειτα από συμπίεση JPEG με απώλειες

Λόγος Συμπίεσης JPEG	13.2	9.5	7.9	3.2
PSNR μετά από την επίθεση	24.64	27.94	31.21	34.62
Συσχέτιση	0.9208	0.9408	0.9446	0.9997

IVML IVML IVML IVML

Σχήμα 6.8: Εξαγόμενα υδατογραφήματα για διαφορετικούς λόγους συμπίεσης(13.2,9.5,7.9,3.2)

$$p = \frac{\sum(w - \bar{w})(w' - \bar{w}')}{\sqrt{\sum(w - \bar{w})^2} \sqrt{\sum(w' - \bar{w}')^2}} \quad (6.5)$$

όπου \bar{w} είναι η μέση τιμή του διανύσματος \mathbf{w} και \bar{w}' είναι η μέση τιμή του διανύσματος \mathbf{w}' .

Αυτός ο συντελεστής μπορεί να χρησιμοποιηθεί ως συμπληρωματικό κριτήριο προς την υποκειμενική ερμηνεία της εξαγόμενης οπτικά αναγνωρίσιμης εικόνας και είναι χρήσιμος για αυτόματη ανίχνευση υδατογραφημένων αντικειμένων βίντεο ή εικόνων (χωρίς ανάγκη ανθρώπινης ερμηνείας των εξαγόμενων σχημάτων).

Ανοχή σε συμπίεση jpeg με απώλειες

Στον Πίνακα 6.1 παρουσιάζονται τα αποτελέσματα εξαγωγής των υδατογραφημάτων από το υδατογραφημένο αντικείμενο το οποίο υπέστη συμπίεση JPEG με λόγους συμπίεσης 3.2, 7.9, 9.5 και 13.2. Όπως μπορούμε να παρατηρήσουμε το εξαγόμενο υδατογράφημα είναι καλής ποιότητας (καθαρά αναγνωρίσιμο) ακόμα και στην περίπτωση των υψηλότερων λόγων συμπίεσης.

Τα εξαγόμενα υδατογράφηματα για τους διαφορετικούς λόγους συμπίεσης σύμφωνα με τον Πίνακα 6.1 φαίνονται στο Σχήμα 6.8.

Ανοχή σε θόρυβο και επιθέσεις επεξεργασίας εικόνας

Κατά τη μετάδοση των υδατογραφημένων εικόνων μέσω δικτύων με ανοχή σφαλμάτων, μπορεί να προστεθεί θόρυβος στα αντικείμενα. Οι διαδικασίες όξυνσης πραγματοποιούνται πολλές φορές για βελτίωση της ποιότητας των αυθεντικών αντικειμένων, ενώ θάμπωμα χρησιμοποιείται σε μερικές περιπτώσεις για τη μείωση των ασυνεχειών που προκαλούνται από τα κανάλια μετάδοσης.

Στον πίνακα 6.2 παρουσιάζονται τα αποτελέσματα εξαγωγής του υδατογραφήματος για το αντικείμενο του Σχήματος 6.7.δ. Στους δύο αυτούς πίνακες και για όλες τις περιπτώσεις τα εξαγόμενα υδατογραφήματα παρουσιάζουν υψηλή συσχέτιση με τα αυθεντικά υδατογραφήματα και είναι καθαρά αναγνωρίσιμα. Σε όλες τις περιπτώσεις τα εξαγόμενα υδατογραφήματα παρουσιάζουν υψηλή συσχέτιση με τα αυθεντικά υδατογραφήματα, ενώ οι περιεχόμενοι χαρακτήρες στις περισσότερες περιπτώσεις αναγνωρίζονται με ευκολία.

Τα εξαγόμενα υδατογράφηματα για τους διαφορετικούς τρόπους επεξεργασίας εικόνας φαίνονται στο Σχήμα 6.8

Η προτεινόμενη μέθοδος έχει εφαρμοσθεί σε εικόνες (still images) και σε χαρακτηριστικά πλαίσια ασυμπίεστου βίντεο. Επιπλέον η προτεινόμενη μέθοδος έχει χρησιμοποιηθεί και για την υδατογράφιση έργων τέχνης. Συγκεκριμένα, σε

Πίνακας 6.2: Εξαγωγή υδατογραφήματος από την περιοχή προσώπου και σώματος έπειτα από επιθέσεις επεξεργασίας εικόνας

Επεξεργασία Εικόνας	Θάμπωμα (Blur)	Όξυνση (Sharpen)	Θάμπωμα + Όξυνση	Όξυνση + Θάμπωμα
PSNR μετά από την επίθεση	29.61	26.59	29.60	29.58
Συσχέτιση	0.840	0.971	0.9403	0.928

IVME IVML IVME IVML

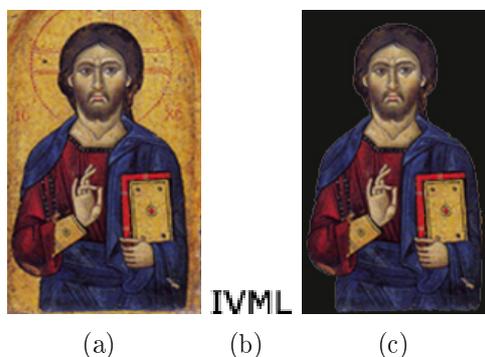
(a) (b) (c) (d)

Σχήμα 6.9: Εξαγόμενα υδατογραφήματα για τρόπους επεξεργασίας εικόνας όπως αναφέρονται στον Πίνακα 6.2

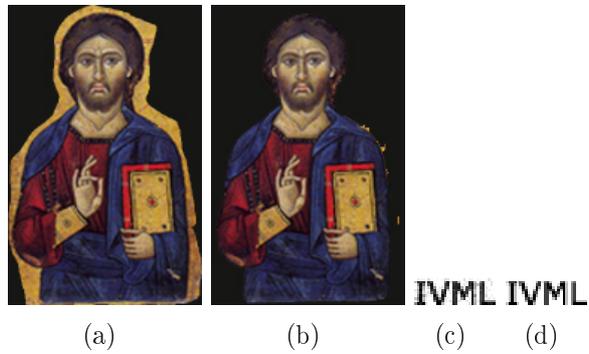
αντικείμενα μορφών Αγίων όπως αυτά εξάγονται, σύμφωνα με την προτεινόμενη μέθοδο του κεφαλαίου 4, εισάγεται το υδατογράφημα (Σχήμα 6.10) παράγοντας τη νέα υδατογραφημένη μορφή.

Θεωρούμε ότι η υδατογραφημένη μορφή ενσωματώνεται πάλι, πίσω στο φόντο της αρχικής εικόνα, και ότι διανέμεται πάνω από ένα δίκτυο π.χ. ηλεκτρονικής μάθησης σε εξουσιοδοτημένους χρήστες. Έστω ότι ένα κακόβουλος εξουσιοδοτημένος χρήστης θελήσει να χρησιμοποιήσει τη μορφή της εικόνας και έστω ότι προσπαθεί με διάφορα εργαλεία να αποκόψει τη μορφή από το φόντο. Εξετάσαμε την περίπτωση επίθεσης κοψίματος στην πρώτη εικόνα του Σχήματος 6.11 - έχει γίνει αποκοπή του αντικειμένου μαζί με το φόντο. Η εξαγωγή του υδατογραφήματος ακόμη και σε αυτή την περίπτωση είναι εφικτή μιας και η εξαγωγή του υψηλής σημαντικότητας δέντρου συντελεστών κυματιδίων δεν έχει υποστεί κάποια αλλοίωση από αυτήν την "επίθεση" (δεν έχουν αλλοιωθεί οι τιμές των υψηλών συχνοτήτων των εικονοστοιχείων). Έπειτα εξετάστηκε η περίπτωση όπου ο χρήστης έχει κόψει κάποιο τμήμα του αντικειμένου (Εικόνα 2 Σχήμα 6.11), όπου και πάλι η ευρωστία του συστήματος είναι ικανοποιητική.

Το σύνολο των πειραματικών αποτελεσμάτων αποδεικνύουν την ασφάλεια που προσφέρει το προτεινόμενο σύστημα υδατογράφησης καθώς και την ευελιξία του για εφαρμογή του σε οποιοδήποτε αντικείμενο εξαχθεί με κάποια μέθοδο από μια εικόνα.



Σχήμα 6.10: Ευρωστία συστήματος υδατογράφησης σε βυζαντινές εικόνες. Στην αρχική εικόνα εφαρμόζεται η μέθοδος εξαγωγή μορφών Αγίων από Βυζαντινές εικόνες βασισμένο σε γνώση και στην εξαγόμενη μορφή ενσωματώνεται το υδατογράφημα



Σχήμα 6.11: α. Επίθεση 1η: αποκοπή όλου του αντικειμένου και τμήματος του φόντου, β. Επίθεση 2η: αποκοπή μέρους του αντικειμένου, γ. Περίπτωση 1η: Εξαγωγή υδατογραφήματος και δ. Περίπτωση 2η: Εξαγωγή υδατογραφήματος

Κεφάλαιο 7

Εφαρμογές

7.1 Εισαγωγή

Η ιδέα του εμπλουτισμού του προτύπου του IEEE με το προφίλ των χρηστών του, έδωσε τη δυνατότητα ανάπτυξης συστημάτων μάθησης προσαρμοσμένα στις ανάγκες του εκάστοτε χρήστη. Βασιζόμενοι στην προτεινόμενη επέκταση του συστήματος μαθησιακής τεχνολογίας, σχεδιάσαμε και υλοποιήσαμε τρία σενάρια χρήσης.

Στο πρώτο σενάριο ως χρήστες συμμετείχαν εκπαιδευτικοί με διαφορετικό επίπεδο γνώσεων πάνω σε θέματα νέων τεχνολογιών. Στόχος του σεναρίου ήταν η εξαγωγή του προφίλ των χρηστών με τη βοήθεια ερωτηματολογίων και η κατάταξής τους σε κατάλληλο επίπεδο με σκοπό την επιμόρφωσή τους σε τεχνολογίες πληροφορικής και υπολογιστών (ΤΠΕ).

Το θέμα του δεύτερου σεναρίου ήταν η δημιουργία ενός εργαλείου υποβοήθησης ανάγνωσης για παιδιά με προβλήματα δυσλεξίας. Η σχεδίαση αυτού του συστήματος στηρίχθηκε πάνω στην προτεινόμενη αρχιτεκτονική (Ενότητα 2.5). Σε αυτό το σενάριο πέρα από την χρήστη ερωτηματολογίου, για τον καθορισμό του αρχικού προφίλ του χρήστη, αναλύεται και η συμπεριφορά του χρήστη τροποποιώντας αυτόματα το αρχικό προφίλ του. Το προφίλ του χρήστη καθορίζει τον τρόπο παρουσίασης του εκπαιδευτικού υλικού.

Το τρίτο σενάριο περιλαμβάνει ως χρήστες φοιτητές σε μάθημα Βυζαντινής τέχνης και στόχος του σεναρίου είναι οι φοιτητές αυτοί να καταλάβουν τις τεχνικές που χρησιμοποιούνται στην αγιογραφία τόσο μέσα από την ανάλυση όσο και από τη σύνθεση των μορφών αγίων. Σε αυτό το σενάριο χρησιμοποιείται η αυτόματη κατηγοριοποίηση του υλικού, των Βυζαντινών εικόνων.

Στη συνέχεια παρουσιάζονται αναλυτικά τα τρία σενάρια.

7.2 1ο σενάριο: Σύστημα εξ αποστάσεως εκπαίδευσης σε θέματα νέων τεχνολογιών

Τα υπάρχοντα συστήματα ηλεκτρονικής μάθησης, που είναι συνήθως αυτοδίδακτα, δεν είναι κατάλληλα για τους αρχαίους που δεν κατέχουν ήδη κάποια γνώση για το θέμα που τους ενδιαφέρει. Για τέτοιους αρχαίους, η αλληλεπίδραση με έναν εκπαιδευτικό παραμένει αναπόφευκτη, δεδομένου ότι τα αυτοματοποιημένα συστήματα ανοικτής γραμμής δεν είναι ικανά να προσφέρουν το υλικό εκμάθησης που προσαρμόζεται στις ανάγκες εκμάθησης κάθε ατόμου [25].

Στο πρώτο σενάριο υιοθετείται η προτεινόμενη επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE 2.5 όπου η ομάδα χρηστών απαρτίζεται από εκπαιδευτικούς της ειδικής αγωγής με διαφορετικό επίπεδο γνώσεων πάνω σε θέματα νέων τεχνολογιών. Σε αυτό το πλαίσιο υιοθετείται η δυναμική αξιολόγησης χρηστών με τη βοήθεια κατάλληλων σχεδιασμένων ηλεκτρονικών ερωτηματολογίων με απώτερο σκοπό την προσαρμοσμένη στις ανάγκες του χρήστη εκπαίδευση. Η προσέγγισή μας είναι βασισμένη στη χρήση ηλεκτρονικού ερωτηματολογίου το οποίο σχεδιάστηκε από ερευνητές και στοχεύει να ανιχνεύσει το γνωστικό επίπεδο του χρήστη και τις προτιμήσεις εκμάθησης πριν από την έναρξη της διαδικασίας ηλεκτρονικής εκμάθησης, καθώς επίσης και μετά από την ολοκλήρωσή της.

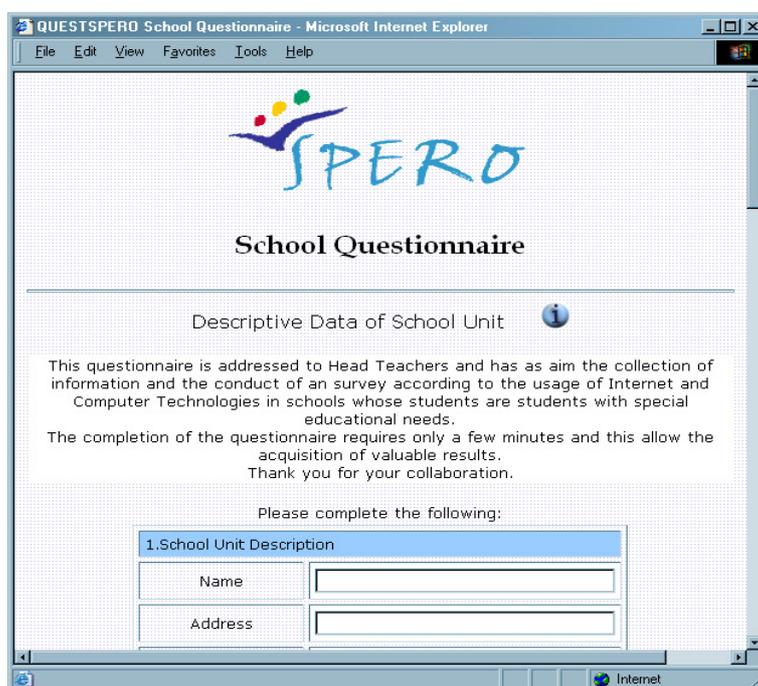
Αρχικά η ομάδα των ειδικών, που απαρτίζονταν από ερευνητές και εκπαιδευτικούς ειδικής αγωγής, σχεδίασε τα ερωτηματολόγια (Παράρτημα Β) με σκοπό την ανάκτηση πληροφοριών. Στόχος αυτών των ερωτηματολογίων ανακτούνται πληροφορίες σχετικά με τα περιγραφικά στοιχεία της σχολικής μονάδας (εξοπλισμός, πλήθος μαθητών κ.τ.λ.) όπου ο χρήστης- δάσκαλος ειδικής αγωγής διδάσκει καθώς και στοιχεία σχετικά με το επίπεδο του χρήστη σε ΤΠΕ (γνωστικό επίπεδο χρήστη σε ΤΠΕ, παιδαγωγική χρήση των ΤΠΕ, επαγγελματική εξέλιξη, επίπεδο γνώση νέων τεχνολογιών).

Οι ποικίλες πτυχές που θέλησαν να παρουσιαστούν με τη χρήση των ηλεκτρονικών ερωτηματολογίων είναι η ταυτότητα και ο ρόλος του εκπαιδευτικού στην Κοινωνία της Πληροφορίας, η ταυτότητα και ο ρόλος των εκπαιδευτικών ιδρυμάτων στην Κοινωνία της Πληροφορίας, ιδιαίτερα των "καθηγητικών" σχολών, η δομή των επιμορφωτικών σεμιναρίων και προγραμμάτων που απευθύνονται σε εκπαιδευτικούς όλων των βαθμίδων, η στάση των εκπαιδευτικών απέναντι σε καινοτόμους μεθόδους και εργαλεία, οι μηχανισμοί εκπαίδευσης και αξιολόγησης εκπαιδευτικών προγραμμάτων στις ευρωπαϊκές χώρες.

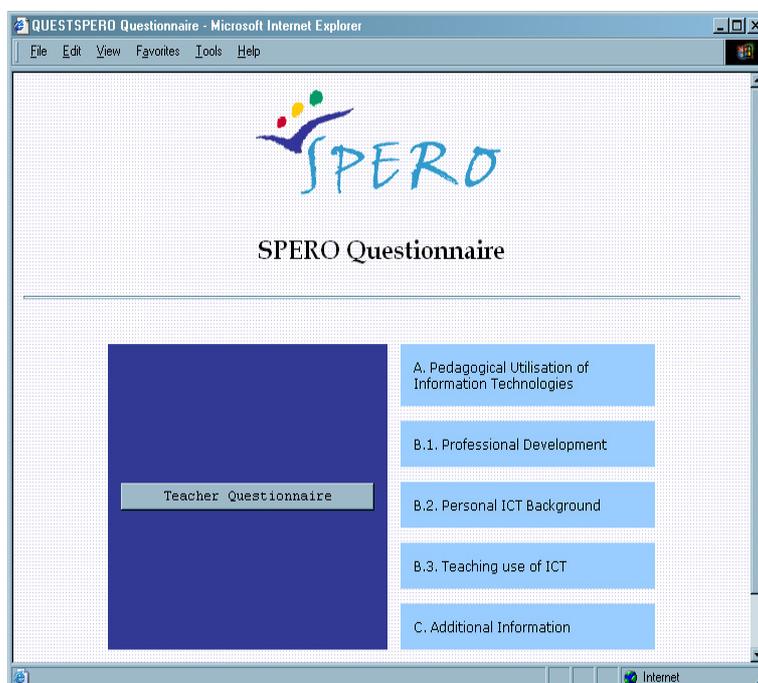
Θέλοντας το ερωτηματολόγιο να αποκτήσει ευέλικτη δομή, κρίθηκε ουσιαστικός ο διαχωρισμός του σε δύο τομείς. Ένας τομέας περιέχει ερωτήσεις σχετικές με την σχολική μονάδα στην οποία ανήκει ο εκπαιδευτικός και ο άλλος τομέας περιέχει ερωτήσεις, τις οποίες καλείται ο εκπαιδευτικός να απαντήσει.

Το σχετικό με την σχολική μονάδα ερωτηματολόγιο (Σχήμα 7.1 και Παράρτημα Β) μπορεί να συλλέξει πληροφορίες σχετικές με τα στοιχεία σχολικής μονάδας, την πληθυσμιακή κατανομή διδασκόντων και μαθητών, τις ειδικές εκπαιδευτικές ανάγκες σχολικής μονάδας (όπως πλήθος μαθητών με σοβαρές ή σύνθετες μαθησιακές δυσκολίες, φυσικές ανικανότητες, αυτισμό, δυσκολίες επικοινωνίας, μαθητές με προβλήματα κώφωσης, όρασης, πλήθος μαθητών με διαταραγμένη συμπεριφοράς, με πολιτιστικά ή γλωσσικά προβλήματα, συγκεκριμένες μαθησιακές δυσκολίες (π.χ. δυσλεξία) και άλλες δυσκολίες), τα επιμορφωτικά προγράμματα που πραγματοποιούνται από τις κυβερνητικές υπηρεσίες ή από την Ευρωπαϊκής Ένωσης στο πλαίσιο εκπαίδευση των εκπαιδευτών ή των μαθητών της σχολικής μονάδας.

Από το ερωτηματολόγιο που απευθύνεται στον εκπαιδευτικό (ερωτηματολόγιο χρήστη - Σχήμα 7.2) μπορούν να εξαχθούν πληροφορίες σχετικές με το προφίλ του εκπαιδευτικού (Επίπεδο βασικών σπουδών, μετεκπαιδύσεις, σεμινάρια), το επίπεδο γνώσης τεχνολογιών πληροφορικής και επικοινωνιών, την παιδαγωγική χρήση των ΤΠΕ (π.χ. αλλαγές στην καθημερινή δραστηριότητα του διδακτικού προσωπικού που οφείλονται στη χρήση των ΤΠΕ, διδακτική διαδικασία και χρήση των ΤΠΕ κ.τ.λ.), τη χρήση ΤΠΕ στη διάρκεια διδασκαλίας των μαθητών του.



Σχήμα 7.1: Ερωτηματολόγιο σχετικό με τη σχολική μονάδα



Σχήμα 7.2: Ερωτηματολόγιο για τον χρήστη

	Beginners	Advanced	Experts
Professional Development	84	223	168
Personal ICT Background	197	215	63
Teaching use of ICT	395	11	69

	User ID	Professional Development	Personal ICT Background	Teaching use of ICT
1.	1937	advanced	beginner	expert
2.	2199	advanced	expert	advanced
3.	996	beginner	advanced	beginner

Σχήμα 7.3: Αποτελέσματα κατηγοριοποίησης όλων των χρηστών στα βασικά προφίλ και παρουσίαση των προφίλ του κάθε χρήστη ανάλογα με το κάθε ερωτηματολόγιο

Η ανάλυση των απαντήσεων των ερωτηματολογίων πραγματοποιήθηκε με τη χρήση της τεχνικής που αναπτύχθηκε στην Ενότητα 3.3. Οι απαντήσεις των χρηστών στα ερωτηματολόγια (Παράρτημα Β) θεωρούνται ως διάνυσμα πολυδιάστατων δεδομένων από τα οποία θέλουμε να εξάγουμε πληροφορία. Αρχικά, ζητούμενο είναι η αντιστοίχιση του διανύσματος σε κάποια από τις στατικές συστάδες του συστήματος (Ενότητα 3.4.1). Έτσι έχοντας αυτά τα διανύσματα, χρησιμοποιούμε την προτεινόμενη επέκταση της ιεραρχικής ομαδοποίησης σε συνδυασμό με μια ασαφή επιλογή χαρακτηριστικών που λαμβάνονται υπόψη κατά τη διαδικασία σύγκρισης των δεδομένων καθώς και στην ασαφοποίηση αυτών. Τα αποτελέσματα αυτής της αρχικής ομαδοποίησης βελτιώνονται μέσω ενός βήματος επαναταξινόμησης χωρίς επίβλεψη όπως περιγράφηκε στην Ενότητα 3.3.2.

Με αυτό τον τρόπο, κάθε τελικός χρήστης ταξινομείται σε ένα αρχικό, στατικό προφίλ (Σχήμα 7.3), που χαρακτηρίζει τη συμπεριφορά του, τα ενδιαφέροντά του. Αυτός ο ιδιαίτερος χαρακτηρισμός προφίλ αποτελεί τη βάση της επόμενης διαδικασίας συσταδοποίησης, η οποία περιλαμβάνει την έννοια της δυναμικής εξαγωγής και της ολοκλήρωσης του προφίλ του χρήστη, όπως αυτή περιγράφηκε στην ενότητα 3.4.2.

Σε κάθε γνωστικό επίπεδο αντιστοιχεί κατάλληλο εκπαιδευτικό υλικό (σειρά ηλεκτρονικών μαθημάτων), το οποίο και παρέχεται στον εκπαιδευόμενο αμέσως μετά την κατάταξή του. Ενδεικτικά οι σειρές ηλεκτρονικών μαθημάτων καλύπτουν την εξής θεματολογία: Βασικές αρχές πληροφορικής και υπολογιστών (π.χ. ορισμοί bit, byte, παρουσιάσεις υλικού υπολογιστών, παρουσιάσεις ειδών λογισμικού), χρήση λειτουργικών συστημάτων (π.χ. σειρά μαθημάτων για χρήση Microsoft Windows και Linux), χρήση προγραμμάτων εφαρμογών (π.χ. χρήση προγράμματος κειμενογράφου, προγράμματος λογιστικών φύλλων, προγράμματος παρουσιάσεων, σχεδιαστικών προγραμμάτων κ.τ.λ.), βασικές υπηρεσίες διαδικτύου (π.χ. χρήση φυλλομετρητή, μηχανές αναζήτησης, αποστολή και παραλαβή ηλεκτρονικού ταχυδρομείου κ.τ.λ.).

Το προφίλ φυλάσσεται στην αποθήκη Προφίλ Εκπαιδευομένου όπως

παρουσιάστηκε στην Ενότητα 2.5 και ανακτάται από εκεί κάθε φορά που επιβεβαιωθεί η πιστοποιημένη είσοδος του χρήστη στο σύστημα. Όταν ένας χρήστης ο οποίος έχει περάσει τη διαδικασία αξιολόγησης (μετά τη συμπλήρωση του ερωτηματολογίου) επισκεφθεί ξανά το προτεινόμενο σύστημα τότε το προφίλ του ανακτάται αυτόματα από την αποθήκη Προφίλ Χρήστη.

Αλλαγές στο προφίλ του χρήστη μπορούν να πραγματοποιηθούν, με βάση τη διαδικασία εκπαίδευσης, μέσω των τεστ αξιολόγησης. Κατά τη διάρκεια όλης της διαδικασίας, καταγράφεται και αξιολογούνται οι επιδόσεις του εκπαιδευόμενου με τη βοήθεια τεστ αξιολόγησης. Τα αποτελέσματα των τεστ συμβάλλουν στην αλλαγή του προφίλ του χρήστη και στην παράδοση του κατάλληλου εκπαιδευτικού υλικού βάσει αυτού του προφίλ, καθώς και στην τροποποίηση του ρυθμού κατάρτισης του και του τρόπου παρουσίασης του υλικού εκμάθησης. Κατά αυτόν τον τρόπο, ένα νέο προφίλ χρήστη μπορεί να δημιουργηθεί βάσει των απαντήσεων του χρήστη στα ερωτηματολόγια ή ένα υπάρχον μπορεί να τροποποιηθεί βασιζόμενο στις επιδόσεις του στα τεστ αξιολόγησης.

7.2.1 Στατιστική ανάλυση των ερωτηματολογίων

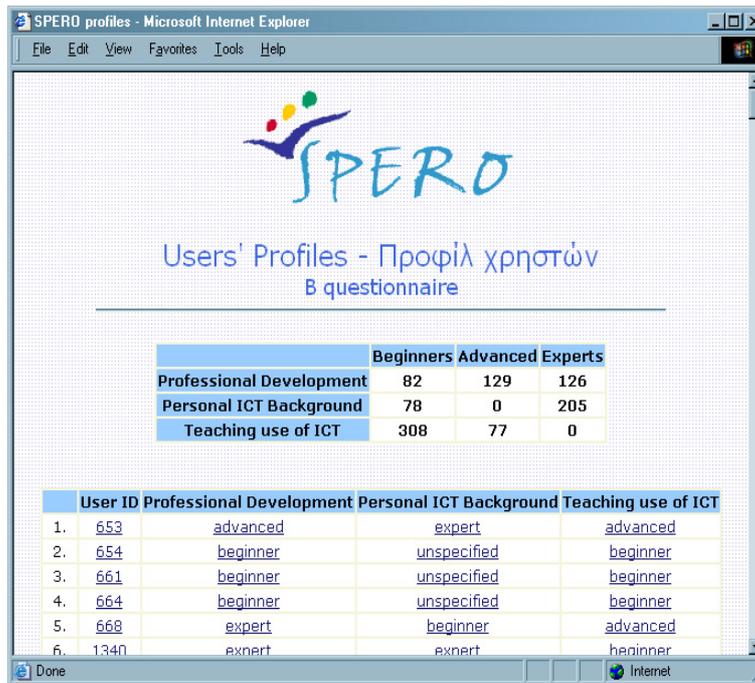
Στο προτεινόμενο σύστημα που βασίζεται στην επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE (Ενότητα 2.5) έχει προστεθεί η διαδικασία Ηλεκτρονική Έρευνα. Σκοπός αυτής της διαδικασίας είναι η στατιστική επεξεργασία των απαντήσεων των ερωτηματολογίων και των τεστ αξιολόγησης των χρηστών, τα οποία έχουν αναλυθεί από τη διαδικασία Αξιολόγηση και έχουν αποθηκευτεί στην αποθήκη Αρχεία Χρηστών.

Το εκπαιδευτικό σενάριο που περιγράψαμε (Ενότητα 7.2) στηρίζεται σε αυτήν την αρχιτεκτονική και άρα είναι εφοδιασμένο με τεχνική για αυτόματη στατιστική ανάλυση των απαντήσεων των χρηστών στα ερωτηματολόγια και στα τεστ αξιολόγησης. Στον ιστόχωρο που φιλοξενεί το προτεινόμενο σύστημα [120] μπορεί οποιοσδήποτε ενδιαφερόμενος να δει τα τρέχοντα αποτελέσματα από την στατιστική ανάλυση των προφίλ των χρηστών [121].

Σε αυτό το εκπαιδευτικό σενάριο συμμετείχαν σχολεία από επτά ευρωπαϊκές χώρες (Ελλάδα, Δανία, Ηνωμένο Βασίλειο, Ισπανία, Ισλανδία, Ουγγαρία και Ρουμανία) γιαυτό και η στατιστική ανάλυση παρουσιάζει ιδιαίτερο ενδιαφέρον. Μέσα από τη στατιστική ανάλυση μπορεί κάποιος να λάβει πλούσια πληροφόρηση σχετικά με τα εκπαιδευτικά συστήματα των χωρών, το επίπεδο των εκπαιδευτικών στις νέες τεχνολογίες, τη εισροή των νέων τεχνολογιών στα σχολεία και ειδικά στα σχολεία ειδικής αγωγής.

Από το κεντρικό μενού της ιστοσελίδας υπάρχει η επιλογή στατιστικά όπου είναι διαθέσιμα τα εξής: στατιστική ανάλυση βασισμένη τα προφίλ χρηστών και αριθμητική στατιστική ανάλυση. Επιλέγοντας την αριθμητική στατιστική ανάλυση δίνονται οι εξής επιλογές: α) στατιστικά συνολικής συμμετοχής (Παράρτημα Γ, Σχήμα 1) β) στατιστικά συμμετοχής σχολείων (Παράρτημα Γ, Σχήμα 2) και γ) στατιστικά συμμετοχής χρηστών (Παράρτημα Γ, Σχήμα 3).

Η πρώτη επιλογή από την αριθμητική ανάλυση δίνει στατιστικά συμμετοχής ανά χώρα για το σύνολο των σχολείων που συμμετείχαν και για το πλήθος των χρηστών (Παράρτημα Γ, Σχήμα 1). Τα στατιστικά εξάγονται δυναμικά λαμβάνοντας υπόψη τα τρέχοντα στοιχεία που υπάρχουν στην βάση δεδομένων του συστήματος. Η επιλογή Στατιστικά συμμετοχής σχολείων, μας πληροφορεί για το ποσοστό των χρηστών



Σχήμα 7.4: Αποτελέσματα κατηγοριοποίησης όλων των χρηστών στα βασικά προφίλ και παρουσίαση των προφίλ του κάθε χρήστη ανάλογα με το κάθε ερωτηματολόγιο

που επέλεξε την κάθε πιθανή απάντηση της ερώτησης (Παράρτημα Γ, Σχήμα 2) και αντίστοιχα είναι και τα στατιστικά συμμετοχής χρηστών (Παράρτημα Γ, Σχήμα 3). Στο Παράρτημα Γ, στα Σχήματα 5, 14, 17 παρουσιάζονται τα αριθμητικά στατιστικά που εξάχθηκαν από το ερωτηματολόγιο που αφορά στους εκπαιδευτικούς.

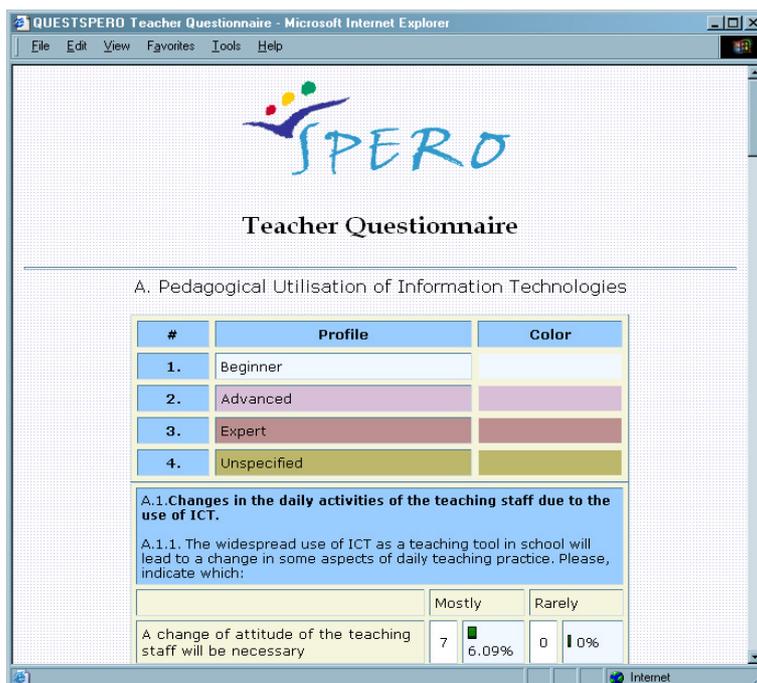
Η δεύτερη επιλογή, η στατιστική ανάλυση βασισμένη τα προφίλ χρηστών [106], παρέχονται από τον ίδιο ιστόχωρο. Επιλέγοντας αυτήν την ανάλυση δίνονται οι εξής επιλογές: α) πληροφορίες για το προφίλ χρηστών ανά χώρα σε κάθε ερώτηση του Β ερωτηματολογίου (δες Παράρτημα Β) και β) συνδυαστικά στατιστικά προφίλ ερωτηματολογίων Α και Β.

Στο σχήμα 7.4 εμφανίζονται πληροφορίες για το προφίλ χρηστών ανά χώρα όπως αυτό εξάγεται για κάθε χρήστη που απάντησε στις ερωτήσεις του Β ερωτηματολογίου (το οποίο συμπεριλαμβάνει τα ερωτηματολόγια Β1. Επαγγελματική εξέλιξη, Β2. Γνώση της Τεχνολογίας, Β3. Χρήση των τεχνολογιών στη διδασκαλία)

Όπως φαίνεται στο Σχήμα 7.4 εμφανίζεται αρχικά ένας πίνακας, στο οποίο το πλήθος των χρηστών που εντάσσεται στις τρεις βασικές κατηγορίες προφίλ. Μια ανώνυμη λίστα χρηστών ακολουθεί όπου εμφανίζεται το προφίλ τους ανά ερωτηματολόγιο. Επιπλέον δίνεται η δυνατότητα εμφάνισης των απαντήσεων του ανά ερωτηματολόγιο.

Τα συνδυαστικά στατιστικά προφίλ ερωτηματολογίων Α και Β δίνουν πληροφορίες για το πως απάντησαν στις ερωτήσεις του ερωτηματολογίου Α χρήστες με διαφορετικά προφίλ σύμφωνα με τα ερωτηματολόγια Β. Για παράδειγμα στο Σχήμα 7.5 φαίνεται το ποσοστό των απαντήσεων που αντιστοιχεί σε κάθε προφίλ. Ομοίως, και για τα άλλα ερωτηματολόγια μπορούμε να πάρουμε αντίστοιχες πληροφορίες.

Στην [121] μπορεί κάποιος να δει αναλυτικά την αριθμητική στατιστική ανάλυση όλων των ερωτηματολογίων ενώ στην [106] παρουσιάζονται τα συνδυαστικά, σε σχέση με το προφίλ, στατιστικά.



Σχήμα 7.5: Κατηγοριοποίηση απαντήσεων στο ερωτηματολόγιο A σύμφωνα με τα προφίλ που ορίστηκαν ανά ερωτηματολόγιο B

7.3 2ο σενάριο: Σύστημα εκπαίδευσης βασισμένη σε ανάλυση μη λεκτικών στοιχείων και με λειτουργία υποβοήθησης ανάγνωσης

Το 2ο σενάριο αναφέρεται σε ένα σύστημα εκπαίδευσης που στηρίχθηκε πάνω στην προτεινόμενη επέκταση της αρχιτεκτονικής της IEEE (Ενότητα 2.5) και στο οποίο έχει ενσωματωθεί η τεχνική ανίχνευσης και παρακολούθησης των χαρακτηριστικών του προσώπου και εκτίμηση βλέμματος και στάσης (Ενότητα 3.4.3.2). Στόχος του συστήματος είναι η ανάλυση της συμπεριφοράς του χρήστη και η προσαρμογή του περιβάλλοντος διεπαφής ανάλογα με τις ανάγκες του. Το σύστημα έχει σχεδιαστεί για άτομα με προβλήματα δυσλεξίας θέλοντας να τα βοηθήσει στην ανάγνωση και στην κατανόηση κειμένου. Προορίζεται για να αποτελέσει ένα καθημερινό εργαλείο αντιμετώπισης δυσκολιών ανάγνωσης και κατανόησης κειμένου από άτομα με δυσλεξία και όχι μόνο. Αυτό το σύστημα έχει αναπτυχθεί στο πλαίσιο του ευρωπαϊκού ερευνητικού προγράμματος (IST) Agent-DYSL [6] στοχεύοντας στο να αναπτύξει ένα επόμενης γενιάς βοηθητικό σύστημα ανάγνωσης. Η βελτίωση της αναγνωστικής ικανότητας του χρήστη, έτσι ώστε αυτός να μπορεί να ενταχθεί στο κανονικό πρόγραμμα της τάξης, χωρίς να έχει την ανάγκη ενισχυτικής διδασκαλίας είναι το ζητούμενο από αυτήν τη διαδικασία.

Από τα μαθητικά του χρόνια ο καθένας από εμάς αναπτύσσει δεξιότητες ανάγνωσης και γραφής καθώς και τις βασικές του ικανότητες οι οποίες και αποτελούν βασικό παράγοντα για τη μετέπειτα πορεία του μέσα στην κοινωνία. Από την άλλη πλευρά, διάφορες μαθησιακές δυσκολίες, όπως η δυσλεξία, δυσχεραίνουν την ομαλή πορεία των ατόμων. Συχνά τα δυσλεκτικά άτομα έχουν δύσκολες μαθησιακές εμπειρίες γιατί δεν έχει γίνει έγκαιρα η διάγνωση της ιδιαίτερης κατάστασής τους. Σύμφωνα με το International Dyslexia Association [10], η δυσλεξία είναι

μια χρόνια κατάσταση με άγνωστη αιτιολογία, η οποία εμπλέκει νευρολογικούς παράγοντες και δεν θεραπεύεται ως ασθένεια αλλά χρίζει αναλυτικής αξιολόγησης και εξατομικευμένης παρέμβασης με στόχο τη βελτίωση της μαθησιακής λειτουργικότητας του ατόμου.

Η έγκαιρη, ολιστική και ακριβής διάγνωση του συγκεκριμένου τύπου δυσλεξίας που ενδεχομένως έχει το παιδί, μέσω μιας πλήρους ψυχοεκπαιδευτικής αξιολόγησης είναι δίχως άλλο η απαρχή της σωστής αντιμετώπισής της [10]. Η έγκαιρη διάγνωση είναι ικανή να δώσει τη δυνατότητα για ειδική εκπαίδευση, να βελτιώσει τις μαθησιακές αδυναμίες του παιδιού ακόμα και να απομακρύνει τις δευτερογενείς συνέπειες που εμφανίζονται όταν ξεκινήσουν οι σχολικές προσδοκίες και έχουν ως συνέπεια πλήγματα στην αυτοπεποίθηση του παιδιού.

Τα οφέλη από τη δημιουργία και τη χρήση επικουρικού λογισμικού, κατάλληλο για άτομα με δυσλεξία, έχουν αναγνωρισθεί ([140], [60]) και σήμερα τίτλοι λογισμικού που απευθύνονται σε χρήστες με δυσλεξία είναι διαθέσιμοι στο εμπόριο ([48],[84], [85], [130], [119]). Αυτά τα βοηθητικά λογισμικά συνήθως παρέχουν οπτικοακουστική αλληλεπίδραση με το χρήστη, θέτοντας ως στόχο την ενίσχυση της ικανότητας ανάγνωσης και κατανόησης. Η βασική λειτουργία του επικουρικού λογισμικού, που συνήθως καλείται "μηχανή ανάγνωσης", συνίσταται στη χρήση συνθετικής φωνής η οποία εκφωνεί τις λέξεις του κειμένου καθώς και στην εμφάνιση του κειμένου ανάλογα με τις ανάγκες του χρήστη.

Αλλά κανένα από αυτά τα προγράμματα δεν έχουν λάβει υπόψη τους ούτε την τρέχουσα κατάσταση του χρήστη ως αποτέλεσμα αυτά τα προγράμματα δεν παρέχουν προσωποποιημένη πρόσβαση στους χρήστες τους.

Τα περισσότερα συστήματα ως προς την εκτίμηση της κατάστασης του χρήστη χρησιμοποιούν συνήθως φωτεινή υπογράμμιση και σύστημα ανάγνωσης του κειμένου χωρίς να λαμβάνουν κάποια είσοδο από τον χρήστη. Υιοθετώντας την προτεινόμενη αρχιτεκτονική (Ενότητα 2.5) προτείναμε ένα συνδυασμό διαφορετικών τεχνολογιών με έναν ευφυή τρόπο τέτοιο ώστε να παράσχει λογισμικό προσαρμοστικό στις ανάγκες του χρήστη. Το λογισμικό αυτό [6] περιλαμβάνει τόσο λεκτική αναγνώριση των λεγόμενων του χρήστη όσο και αναγνώριση της κατάστασης του χρήστη (Ενότητα 3.4.3.1) με σκοπό να αποκτηθούν οι πληροφορίες για τον αρχάριο, οι οποίες είναι μια προϋπόθεση για ένα εξατομικευμένο μαθησιακό περιβάλλον, το οποίο λαμβάνει υπόψη τις ιδιαίτερες ανάγκες κάθε χρήστη με δυσλεξία.

Η χρήση αυτού του συστήματος μπορεί να πραγματοποιηθεί είτε μέσα σε σχολικό ή σε οικιακό περιβάλλον. Στην περίπτωση του σχολικού περιβάλλοντος, η χρήση του συστήματος γίνεται μέσω υπολογιστών που υπάρχουν στις αίθουσες ένταξης στα δημοτικά σχολεία και με την επίβλεψη και τη βοήθεια του ειδικού δασκάλου. Στην περίπτωση του οικιακού περιβάλλοντος, το σύστημα εγκαθίσταται στον οικιακό υπολογιστή και το παιδί θα μπορεί να το χρησιμοποιεί είτε μόνο του ή με τη βοήθεια των γονέων του.

Τα βασικά χαρακτηριστικά του συστήματος περιλαμβάνουν τα εξής:

- Αυτόματη παρακολούθηση και εκτίμησης της λεκτικής απόδοσης του χρήστη,
- Αυτόματη παρακολούθηση και εκτίμησης της τρέχουσας συναισθηματικής και ψυχολογικής κατάστασης,
- Κατασκευή προσωπικό προφίλ του χρήστη,
- Παροχή προσωποποιημένης βοήθειας για το χρήστη ανάλογα με το προφίλ του.

Τα δύο πρώτα χαρακτηριστικά του συστήματος ενσωματώνονται στη ροή συμπεριφορά του προτεινόμενου συστήματος της Ενότητας 2.5. Το χαρακτηριστικό κατασκευή προσωπικού προφίλ του χρήστη αναφέρεται ουσιαστικά στη λειτουργία της διαδικασίας αξιολόγησης του προτεινόμενου συστήματος ηλεκτρονικής μάθησης και η φύλαξή του γίνεται στην αποθήκη προφίλ χρηστών. Η παροχή προσωποποιημένης βοήθειας ενσωματώνεται στη διαδικασία δάσκαλος του προτεινόμενου συστήματος ηλεκτρονικής μάθησης (Ενότητας 2.5).

7.3.1 Γενική περιγραφή ευφυούς συστήματος ανάγνωσης

Το προτεινόμενο σύστημα έχει ως στόχο την ανάπτυξη ενός ευφυούς συστήματος ανάγνωσης, το οποίο θα χρησιμοποιηθεί αρχικά από μαθητές δημοτικού που αντιμετωπίζουν προβλήματα δυσλεξίας με σκοπό τη βελτίωση της μαθησιακής τους αντίληψης καθώς και της ικανότητας κατανόησης κειμένων. Η ευφυΐα του συστήματος οφείλεται στη δυνατότητα που αυτό θα έχει να προσαρμόζει το περιβάλλον (interface) του στις ανάγκες του εκάστοτε χρήστη. Ο χρήστης, όταν προσπελάσει για πρώτη φορά το σύστημα μπορεί να καθορίσει το προφίλ του είτε απαντώντας σε ερωτηματολόγιο σχετικό με τα προβλήματα δυσλεξίας που έχει ή αυτόματα αφού πρώτα χρησιμοποιήσει το υποσύστημα αξιολόγησης.

Το υποσύστημα αξιολόγησης παρουσιάζει προτάσεις στην οθόνη του χρήστη (από το διαθέσιμο ειδικό σύνολο προτάσεων) τις οποίες θα καλείται να διαβάσει. Οι λέξεις που περιλαμβάνονται σε αυτές τις προτάσεις αποτελούν το λεξιλόγιο του συστήματος. Οι λέξεις αυτές έχουν επιλεγεί από ομάδα ειδικών σε θέματα δυσλεξία και αποτελούν λέξεις που κατά το σύνολό τους δυσκολεύουν τους δυσλεκτικούς αναγνώστες.

Κατά τη διάρκεια της ανάγνωσης, το υποσύστημα αναλύει τη φωνή του χρήστη, αναγνωρίζοντας και καταγράφοντας τα λάθη του. Ανάλογα με τις επιδόσεις του χρήστη, δημιουργείται το προφίλ του, το οποίο και φυλάσσεται στο σύστημα. Το υποσύστημα αξιολόγησης εκτός από τη δυνατότητα ανάλυσης φωνής και καταγραφής λαθών, αναλύει τη συμπεριφορά του χρήστη κατά τη διάρκεια της ανάγνωσης. Υιοθετώντας τεχνικές ανάλυσης εικόνας, το σύστημα εκτιμά αυτόματα τόσο τη συναισθηματική κατάσταση του χρήστη, όσο και τη σωματική κατάστασή του, αντλώντας πληροφορίες από τη στάση του σώματος καθώς και τις χειρονομίες του και από τις εκφράσεις του προσώπου του 3.4.3.1. Τα αποτελέσματα των αναλύσεων θα αποτελούν άλλη μια παράμετρο για τον καθορισμό του προφίλ του χρήστη.

Στο πλαίσιο της καθημερινής χρήσης του συστήματος, παρέχεται η δυνατότητα στο χρήστη να "φορτώνει" κάποιο κείμενο σε μορφή pdf (η μορφή που είναι τα αρχεία των εκπαιδευτικών ηλεκτρονικών βιβλίων). Αμέσως, το περιβάλλον του συστήματος θα προσαρμοστεί σύμφωνα με το προφίλ του χρήστη, όπως αυτό καθορίστηκε από το υποσύστημα αξιολόγησης. Ο χρήστης θα καλείται ανά τακτά διαστήματα να αξιολογείται από το υποσύστημα αξιολόγησης, με σκοπό το προφίλ του να είναι πάντα έγκυρο.

Το προτεινόμενο σύστημα αναπτύχθηκε στο πλαίσιο του προγράμματος AGENT-DYSL [6] και προσδοκά να αποτελέσει ένα ουσιαστικό βοήθημα για τα δυσλεκτικά παιδιά και κατ' επέκταση και για άλλες πληθυσμιακές ομάδες που αντιμετωπίζουν προβλήματα ανάγνωσης (παλλινοστούντες, αλλοδαποί, δυσλεκτικοί μεγαλύτερης ηλικίας, άτομα με μαθησιακά προβλήματα κ.α.). Το σύστημα αυτό αρχικά υποστηρίζει τρεις γλώσσες (Ελληνική, Αγγλική και Δανέζικη) και έχει ήδη εγκατασταθεί και λειτουργεί σε δημοτικά σχολεία ένταξης στην Ελλάδα, στη Μεγάλη Βρετανία και

στη Δανία καθώς και σε κέντρα διάγνωσης, αξιολόγησης και υποστήριξης και στις προαναφερθείσες χώρες.

Επειδή τα λάθη που μπορεί να κάνει ένας δυσλεκτικός αναγνώστης ποικίλλουν από γλώσσα σε γλώσσα [10], δε μπορεί να γίνει κοινή καταγραφή των λαθών και για τις τρεις γλώσσες. Γιαυτό το λόγο, στο πλαίσιο του προγράμματος AGENTDYSL, κρίθηκε απαραίτητο να συμμετέχουν στην ομάδα ειδικών, ειδικοί σε θέματα δυσλεξίας από κάθε χώρα. Ο βασικός ρόλος των ειδικών ήταν η ομαδοποίηση και η καταγραφή των πιθανών λαθών, ανά γλώσσα, που μπορεί να κάνει ένας δυσλεκτικός.

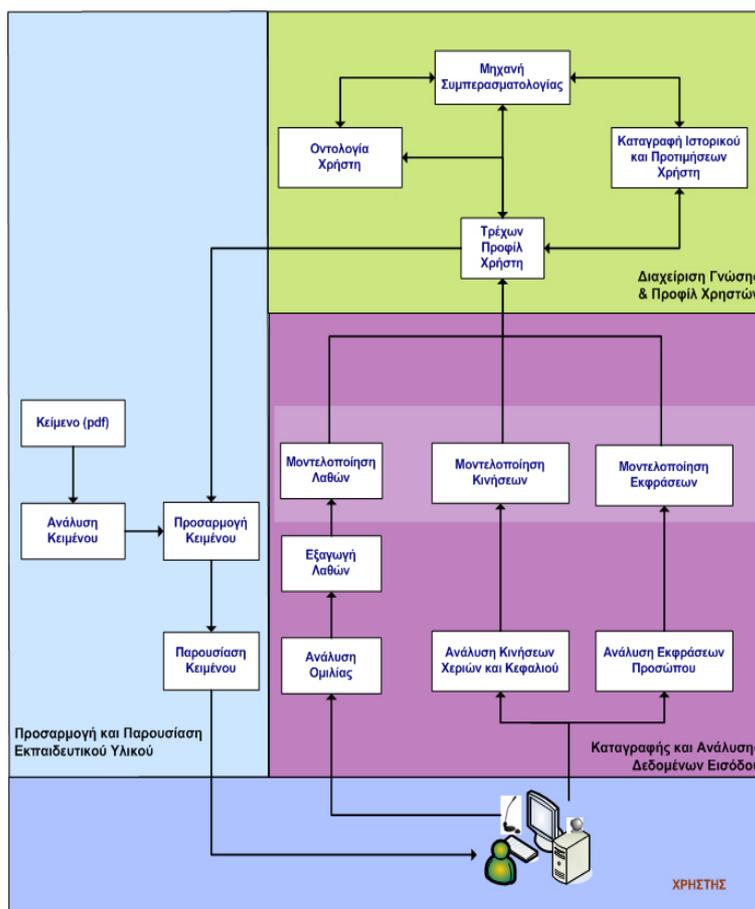
Στόχος του συστήματος είναι η δημιουργία ενός προσαρμοζόμενου εκπαιδευτικού περιβάλλοντος, το οποίο λαμβάνοντας υπόψη του το γενικό πλαίσιο εκπαίδευσης (όπως αλληλεπιδράσεις μεταξύ δυσλεκτικών και μη ατόμων μέσα σε μια κοινή αίθουσα) και το οποίο θα επιτρέπει σε ανθρώπους με δυσλεξία να εντάσσονται ομαλά μέσα στην κοινωνία όπου ζουν. Η επιτυχία ενός τέτοιου σκοπού είναι σημαντική μιας και θα βοηθήσει στην οικοδόμηση της γνώσης γύρω από το πεδίο κατανόησης της δυσλεξίας. Γι' αυτό το λόγο, στην ομάδα των ειδικών που αναπτύσσουν αυτό το σύστημα δεν εντάσσονται μόνο τεχνολόγοι, αλλά και ειδικοί σε θέματα δυσλεξίας οι οποίοι οριοθετούν το πλαίσιο, παρέχοντας κατάλληλες τεχνικές για την αντιμετώπιση προβλημάτων δυσλεξίας.

7.3.2 Περιγραφή υποσυστημάτων ευφυούς συστήματος ανάγνωσης

Το ευφύες σύστημα σνάγνωσης μπορεί να θεωρηθεί ότι αποτελείται από τα εξής δομικά υποσυστήματα (Σχήμα 7.6): το υποσύστημα αξιολόγησης το οποίο περιλαμβάνει τις λειτουργίες της καταγραφής και ανάλυσης δεδομένων εισόδου (ήχου και εικόνας) καθώς και τη μηχανή διαχείρισης γνώσης και προφίλ χρηστών και τη μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού. Το υποσύστημα αξιολόγησης ταυτίζεται στη προκειμένη περίπτωση με τη διαδικασία αξιολόγησης του προτεινόμενου συστήματος ηλεκτρονικής μάθησης (Ενότητα 2.5). Η μηχανή διαχείρισης γνώσης αποτελεί τμήμα της διαδικασίας δάσκαλος και το προφίλ χρηστών τμήμα της αποθήκης προφίλ χρηστών. Τέλος η μηχανή προσαρμογής ενσωματώνεται στη διαδικασία δάσκαλος και παρουσίασης εκπαιδευτικού υλικού αντιστοιχεί στην ροή παράδοση υλικού.

7.3.2.0.1 Καταγραφή και ανάλυση δεδομένων εισόδου Η μονάδα Καταγραφής και Ανάλυσης Δεδομένων Εισόδου του υποσυστήματος αξιολόγησης λαμβάνει ως είσοδο τόσο τη φωνή όσο και την εικόνα του χρήστη του συστήματος, χρησιμοποιώντας το μικρόφωνο και την κάμερα που είναι εγκατεστημένα στο ίδιο υπολογιστή με το λογισμικό του προτεινόμενου συστήματος. Έπειτα, η λειτουργία αυτόματης ανάλυσης ομιλίας είναι υπεύθυνη για την ανάλυση της φωνής του χρήστη, ενώ η λειτουργία της ανάλυσης της συμπεριφοράς είναι υπεύθυνη για την ανάλυση τόσο τη συναισθηματική κατάσταση όσο και τη σωματική κατάσταση του χρήστη, στοιχεία που παρέχουν σημαντικές πληροφορίες για το ποσοστό εστίασης, προσοχής και κατανόησης του κειμένου το οποίο διαβάζει ο χρήστης.

Η λειτουργία της αυτόματης ανάλυσης ομιλίας ενεργοποιείται, προς το παρόν, μόνο από το υποσύστημα αξιολόγησης. Μελλοντικά, όταν η υπολογιστική ισχύς των συστημάτων που διαθέτουν τα σχολεία καθώς και οι προσωπικοί υπολογιστές



Σχήμα 7.6: Δομικά υποσυστήματα του ευφυούς συστήματος ανάγνωσης

αυξηθεί, θα μπορούσε να λειτουργήσει και για το υποσύστημα μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού. Σε αυτό το πλαίσιο η αυτόματης ανάλυσης ομιλίας θα μπορούσε να έχει ενσωματώσει μεγαλύτερο λεξιλόγιο, με σκοπό την αναγνώριση περισσοτέρων λέξεων που ο χρήστης εκφωνεί. Προς το παρόν χρησιμοποιεί ένα λεξικό 600 λέξεων (για κάθε μία γλώσσα -Ελληνικά, Αγγλικά, Δανέζικα) κατάλληλα επιλεγμένο για τον εντοπισμό προβλημάτων που αντιμετωπίσει ένας δυσλεκτικός αναγνώστης. Η δημιουργία λεξιλογίου του συστήματος εξηρητεεί στην εξαγωγή του προφίλ των χρηστών του δηλαδή στην μοντελοποίηση της γνώσης, η οποία είναι σημαντική για το σύστημα, καθώς θα μπορεί να βασιστεί σε αυτή και να πετύχει ακριβή εξαγωγή του προφίλ του χρήστη.

Αναλυτικότερα, αρχικά στο χρήστη δίνεται μέσω του συστήματος ένα κείμενο (το οποίο αποτελείται από λέξεις που ανήκουν στο λεξιλόγιο των 600 λέξεων) και καλείται να το διαβάσει δυνατά. Η ανάλυση της φωνής του χρήστη πραγματοποιείται με τη χρήση ενός συστήματος ανάλυσης ομιλίας [107]. Το σύστημα αυτό είναι ικανό να αναλύσει το σήμα της φωνής του χρήστη, φτάνοντας σε επίπεδο φθόγγων και για τις τρεις γλώσσες (Αγγλική, Ελληνική και Δανέζικη), και καταγράφει λεπτομερώς τα λάθη συνάρθρωσης και επιτονισμού που κάνει ο χρήστης καθώς διαβάζει. Τα λάθη του χρήστη κατά την ανάγνωση θα αποτελούν μια από τις βασικές παραμέτρους για τον καθορισμό του προφίλ του. Επιπλέον, τα χαρακτηριστικά γνωρίσματα της ομιλίας κάθε χρήστη ανιχνεύονται και μοντελοποιούνται, αποτελώντας επιπλέον παραμέτρους για το προφίλ του χρήστη.

Η λειτουργία ανάλυσης της συμπεριφοράς λαμβάνει ως είσοδο το σήμα της κάμερας που καταγράφει τις κινήσεις του χρήστη. Στόχος είναι η ανάλυση και ο καθορισμός της συναισθηματική κατάσταση του χρήστη. Αυτό πραγματοποιείται χρησιμοποιώντας τεχνική ανάλυσης εκφράσεων προσώπου [71] (Ενότητα 3.4.3.1), η οποία λαμβάνει υπόψη της τη θέση και το σχήμα του στόματος, των ματιών και των χειλιών του χρήστη και, στηριζόμενη σε ψυχολογικές μελέτες, είναι ικανή να εξάγει την ψυχολογική κατάσταση του χρήστη. Οι βασικές κατηγορίες (βασικά συναισθήματα) που χαρακτηρίζουν τη ψυχολογική κατάσταση του χρήστη είναι οι εξής έξι: θυμός, λύπη, έκπληξη, χαρά απέχθεια και φόβος, αλλά το σύστημα είναι ικανό να εξάγει οποιοδήποτε συνδυασμό των βασικών συναισθημάτων παράγοντας κατά αυτό τον τρόπο και "ενδιάμεσα" συναισθήματα.

Επιπλέον, χρησιμοποιούνται τεχνικές [11] ανάλυσης της σωματικής κατάστασης του χρήστη που βασίζονται σε ανάλυση κινήσεων των χεριών του, των κινήσεων και της θέσης του σώματος και δίνουν πληροφορίες σχετικά με το αν ο χρήστης είναι ενεργός ή παθητικός, απασχολημένος ή όχι, υπερκινητικός ή ήρεμος, έχει θετική αντιμετώπιση ή όχι (Ενότητα 3.4.3.1). Τα αποτελέσματα της ανάλυσης της κατάστασης του χρήστη μοντελοποιούνται, μετατρέποντας σε επιπλέον παραμέτρους για το προφίλ του χρήστη.

Η λειτουργία ανάλυσης της συμπεριφοράς χρησιμοποιείται επιπλέον και από τη μηχανή διαχείρισης γνώσης και προφίλ χρηστών και από το υποσύστημα μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού. Γενικότερα, η συμπεριφορά του χρήστη επηρεάζει την ταχύτητα της φωτεινής υπογράμμισης ενώ επηρεάζει και το προφίλ του, αφού αυτό εμπλουτίζεται με την τρέχουσα κατάσταση του χρήστη.

7.3.2.0.2 Μηχανή διαχείρισης γνώσης και προφίλ Η μηχανή διαχείρισης γνώσης και προφίλ λαμβάνει τις παραμέτρους που εξάγονται από τη μονάδα καταγραφής και ανάλυσης δεδομένων εισόδου και μία από τις αρμοδιότητές της

είναι η δημιουργία του προφίλ κάθε χρήστη. Το προφίλ αυτό καθορίζει τις ιδιαιτερότητες του χρήστη και βοηθά το σύστημα να προσαρμόζεται στις ανάγκες του, παρέχοντας του, ένα πρωτότυπο, προσαρμοσμένο στις ανάγκες του, εκπαιδευτικό περιβάλλον. Το σύστημα, γνωρίζοντας το προφίλ του χρήστη, θα είναι, για παράδειγμα, ικανό να προβλέπει τα πιθανά λάθη που ο χρήστης θα μπορούσε να κάνει και χρησιμοποιώντας ποικίλες τεχνικές παρουσίασης (Μηχανή Προσαρμογής και Παρουσίασης Εκπαιδευτικού Υλικού), θα προειδοποιεί το χρήστη, ώστε να αποφεύγει αυτά τα λάθη. Αυτά τα επιπρόσθετα χαρακτηριστικά μπορούν να αποτελέσουν αντικείμενο μελλοντικής έρευνας.

Μετά την ανάλυση της συμπεριφοράς του χρήστη, το σύστημα είναι σε θέση να γνωρίζει αν αυτός έχει κουρασθεί ή είναι αγχωμένος - κάτι που είχε ως αποτέλεσμα τη μείωση της μαθησιακής του αντίληψης και έτσι το περιβάλλον του συστήματος (interface) προσαρμόζεται στο χρήστη (τα γραφικά που θα χρησιμοποιούνται θα είναι κατάλληλα για την ηλικία των χρηστών καθώς και το μέγεθος - χρώμα), βελτιώνοντας την μαθησιακή του αντίληψη. Κάθε φορά που το σύστημα πραγματοποιεί νέα εκτίμηση για το χρήστη, αυτή αποθηκεύεται στο προφίλ του χρήστη και θεωρείται ως η τρέχουσα κατάσταση του, ενώ παράλληλα υπάρχει η δυνατότητα καταγραφής του ιστορικού του χρήστη (διατηρώντας τις αλλαγές στο προφίλ του), καθώς και τυχόν προτιμήσεις του χρήστη. Με σκοπό να επιτευχθεί πραγματική προσαρμογή του συστήματος στο χρήστη, είναι απαραίτητο το σύστημα να κατανοήσει τα στοιχεία της συμπεριφοράς και της απόδοσης του χρήστη και να τα καταγράψει σε οντολογίες.

Αυτά τα δεδομένα αποτελούν μια προσωπική οντολογία για το χρήστη, η οποία θα μπορεί ανά πάσα στιγμή να ανακτηθεί από το σύστημα. Η σχεδίαση και η υλοποίηση μεθόδων και εργαλείων για την έξυπνη διαχείριση τόσο του εκπαιδευτικού υλικού, όσο και του χρήστη, θεωρείται το κλειδί για τη δημιουργία του προτεινόμενου ευφυούς συστήματος. Η έξυπνη διαχείριση επιτεύχθηκε υιοθετώντας εργαλεία αυτόματης εκπαίδευσης, αξιολόγησης και διαχείρισης της οντολογίας του χρήστη, με σκοπό την αυτόματη διαχείριση της γνώσης που εξάγεται κατά την αλληλεπίδραση του χρήστη με το σύστημα.

Σε αυτό το πλαίσιο, χρησιμοποιήθηκαν τεχνικές μηχανικής μάθησης καθώς και έξυπνοι αλγόριθμοι [104], [115], [114] με στόχο τη χρήση μηχανής συμπερασματολογίας (Ενότητα 5.3.2), η οποία αξιολογεί τα υπάρχοντα προφίλ καθώς και το εκπαιδευτικό υλικό του συστήματος και εμφανίζει, σε πραγματικό χρόνο, το εκπαιδευτικό υλικό στην κατάλληλη, σύμφωνα με την τρέχουσα κατάσταση και πρόοδο του χρήστη, μορφή.

Για να μοντελοποιήσουμε τη γνώση που διατίθεται από το σύστημα πρέπει να επιλέξουμε τον κατάλληλο φορμαλισμό της οντολογίας (Ενότητα 5.3). Η αξιολόγηση και η επιλογή τόσο της γλώσσα ορισμού της οντολογίας, των εργαλείων, της πλατφόρμας και των τεχνικών είναι από τα βασικά βήματα που έγιναν για το σωστό καθορισμό της οντολογίας του προτεινόμενου συστήματος [114]. Εκτός από την επιλογή του κατάλληλου φορμαλισμού, πραγματοποιηθεί και ο καθορισμός τόσο του λεξιλογίου που θα χρησιμοποιηθεί για την περιγραφή της κατάστασης του χρήστη καθώς και του εκπαιδευτικού περιεχομένου όσο και της περιγραφής του γενικού περιεχομένου του συστήματος. Έτσι, αρχικά θα πραγματοποιηθεί ο ορισμός και η κατασκευή της οντολογίας του συστήματος καθώς και ο καθορισμός των σημασιολογικών σχέσεων ανάμεσα στα μεταδεδομένα και στους περιγραφείς του συγκεκριμένου πεδίου εφαρμογής.

Συγκεκριμένα, η οντολογία περιλαμβάνει γνώση σχετική με άτομα με γνωστικές

δυσκολίες, γνώση σχετική με το εκπαιδευτικό υλικό, γνώση για την ολοκλήρωση της οντολογίας, συμπερασματικούς κανόνες και περιορισμούς που θα χρησιμοποιηθούν από τη μηχανή συμπερασματολογίας.

Δύο τύποι μηχανών συμπερασματολογίας εφαρμόστηκαν στο πλαίσιο του προτεινόμενου συστήματος. Ο πρώτος τύπος διαπραγματεύεται με τον έλεγχο της συνοχής καθώς και με τις διαδικασίες υποβολής ερωτήσεων και ανάκτησης δεδομένων. Ο δεύτερος τύπος παράγει την on-line συμπερασματολογία, χρησιμοποιώντας προκαθορισμένους κανόνες. Και στους δύο τύπους μηχανών συμπερασματολογίας, τεχνικές ασαφούς λογικής θα χρησιμοποιηθούν έτσι ώστε να επιτευχθεί η αντιμετώπιση των αβέβαιων καταστάσεων (π.χ. καταστάσεις όπου είτε η αναγνωστική ικανότητα του χρήστη ή η κατάσταση του δεν είναι βέβαιες). Τελικός σκοπός είναι η αυτόματη (ή η ημιαυτόματη) δόμηση της γνώσης που σχετίζεται με το προτεινόμενο σύστημα, χρησιμοποιώντας τα εξαγόμενα μεταδεδομένα από τους χρήστες του συστήματος (δυσλεκτικούς ή μη). Οι οντολογίες θα παρέχουν ένα κοινό λεξιλόγιο και μία κοινή αναπαράσταση μεταξύ των μονάδων του συστήματος, υποστηρίζοντας τη διαμοιραζόμενη γνώση και την επαναχρησιμοποίησή της μεταξύ διαφορετικών σχολείων, ερευνητικών κέντρων κ.τ.λ. Οι οντολογίες και οι σημασιολογικοί χαρακτηρισμοί αυτών αναμένεται να παίζουν το βασικό ρόλο, υποστηρίζοντας την ολοκλήρωση και την ένωση των πληροφοριών.

Ο εμπλουτισμός της οντολογίας πραγματοποιείται μέσω της μηχανής συμπερασματολογίας, χρησιμοποιώντας πληροφορίες που αυτή εξάγει. Πιο συγκεκριμένα, η διαδικασία εξαγωγή δίνει δείγματα από διαφορετικές έννοιες, μαζί με τα χαρακτηριστικά τους καθώς επιπλέον θα παρέχει μη ταξινομημένες οντότητες οι οποίες οδηγούν στον εμπλουτισμό της οντολογίας με νέες έννοιες και σχέσεις. Αυτές οι νέες έννοιες και σχέσεις εξάγονται μέσω μιας έξυπνη διαδικασίας ανίχνευσης. Κατά τη διάρκεια δημιουργίας, οποιαδήποτε αλλαγή μπορεί να δημιουργήσει ασυμβατότητα με τα άλλα μέρη της οντολογίας. Οπότε κρίνεται απαραίτητη η ανάπτυξη μοντέλων, τεχνικών και εργαλείων για τη σημασιολογική συμβατότητα, ελέγχοντας το περιεχόμενο της οντολογίας μέσω μιας διαδικασίας αξιολόγησης.

Οι λειτουργίες της μηχανής συμπερασματολογίας είναι ικανές να εξάγουν γνώση, χρησιμοποιώντας την οντολογία. Επιπλέον, μπορούν να ελέγχουν τη συμβατότητα και την εγκυρότητα του παραγόμενου δείγματος, εξάγουν πλήρη γνώση χρησιμοποιώντας υποθέσεις και σχέσεις ισότητας όπως αυτές έχουν δηλωθεί στην οντολογία, και τέλος να λαμβάνουν αποφάσεις χρησιμοποιώντας συμπερασματικούς κανόνες.

Διαφορετικές προσεγγίσεις υπάρχουν για την κατασκευή μηχανών συμπερασματολογίας ανάλογα με τον λογικό φορμαλισμό που έχει χρησιμοποιηθεί για την κατασκευή της οντολογίας. Οι περιγραφικές λογικές έχουν αποδειχθεί κατάλληλες για πολλές εφαρμογές μιας και έχουν μεγάλη εκφραστικότητα και μικρή υπολογιστική πολυπλοκότητα. Αλλά, λόγω της μεγάλης αβεβαιότητας που υπάρχει στην εφαρμογές πραγματικού χρόνου, η υιοθέτηση της ασαφής αναπαράστασης της πληροφορίας κρίνεται απαραίτητη (Ενότητα 5.3.1.2). Κατά αυτόν τον τρόπο, η μηχανή συμπερασματολογίας μπορεί και διαχειρίζεται τους περιορισμούς και τους κανόνες.

7.3.2.0.3 Μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού

Η Μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού διαχειρίζεται την παρουσίαση του εκπαιδευτικού υλικού σύμφωνα με το προφίλ του εκάστοτε χρήστη. Η

μηχανή προσαρμογής υποστηρίζει την απευθείας προσαρμογή του τρόπου παρουσίασης του εκπαιδευτικού υλικού (μεγαλύτερη - μικρότερη γραμματοσειρά, φωτιζόμενο κείμενο κ.τ.λ.) κατά τη διάρκεια ανάγνωσης κειμένου από το χρήστη, ανάλογα με το τρέχον προφίλ του ([139], [45]).

Το προτεινόμενο σύστημα, λοιπόν, μπορεί να προσαρμόσει το στυλ της παρουσίασης του εκπαιδευτικού υλικού και το "ποσοστό βοήθειας" βασιζόμενο στην τρέχουσα κατάσταση του χρήστη. Για παράδειγμα, η ταχύτητα ανάγνωσης και η αντίθεση μεταξύ των φωτισμένων λέξεων και του φόντου θα πρέπει να συμβαδίζουν. Έτσι, το ποσοστό της βοήθειας που παρέχεται στο χρήστη προσαρμόζεται ενώ και η μοντελοποίηση του χρήστη ενημερώνεται συνεχώς από την τρέχουσα κατάσταση του.

Ανακεφαλιώνοντας, θα κάνουμε μια ανασκόπηση της λειτουργίας του προτεινόμενου συστήματος, βλέποντας το από την πλευρά του χρήστη. Η μηχανή προσαρμογής και παρουσίασης εκπαιδευτικού υλικού αλληλεπιδρά με το χρήστη, μεταβάλλοντας τις ρυθμίσεις για την παρουσίαση του υλικού (ταχύτητα ανάγνωσης, μέγεθος γραμματοσειράς, φωτεινή υπογράμμιση). Από τη μεριά του χρήστη, η αλληλεπίδραση πραγματοποιείται στο πλαίσιο καταγραφής τόσο της φωνής, όσο και των κινήσεων του χρήστη. Επιπλέον, ο χρήστης μπορεί να λάβει ανά πάσα στιγμή την προσωπική του οντολογία.

7.3.3 Δοκιμαστικά σενάρια χρήσης

Όσον αφορά τη μοντελοποίηση του συστήματος και την ανάλυση των απαιτήσεων των χρηστών, ορίστηκαν δοκιμαστικά σενάρια χρήσης, τα οποία διαφέρουν ως προς το βαθμό εμπλοκής τους στο σύστημα καταγραφής και ανάλυσης των δεδομένων εισόδου.

Λειτουργία "Εκπαίδευση"

Κατά τη διάρκεια της λειτουργίας "Εκπαίδευση" του ευφυούς συστήματος ενεργοποιείται το υποσύστημα αξιολόγησης. Ο χρήστης καλείται να διαβάσει το κείμενο. Τα δεδομένα που καταγράφονται από τη μονάδα Καταγραφής και Ανάλυσης Δεδομένων Εισόδου περιέχουν τη φωνή του χρήστη της οποίας η ανάλυση οδηγεί σε προσδιορισμό των λαθών που ο χρήστης έχει κάνει. Επιπλέον, αυτά τα δεδομένα περιέχουν και την καταγραφή του χρήστη από την κάμερα και αναλύονται από το ίδιο υποσύστημα, εξάγοντας την τρέχουσα κατάσταση του χρήστη.

Τα αποτελέσματα των αναλύσεων συλλέγονται από τη Μηχανή Διαχείρισης Γνώσης και Προφίλ Χρηστών, η οποία είναι υπεύθυνη για τη δημιουργία/ ενημέρωση του προφίλ του χρήστη. Όταν η ικανότητα ανάγνωσης καθώς και η κατάσταση του χρήστη διαφοροποιηθεί από αυτή που είναι καταγεγραμμένη στο τρέχον προφίλ του, η Μηχανή Διαχείρισης Γνώσης και Προφίλ Χρηστών προσαρμόζει το προφίλ του.

Λειτουργία "Μόνο Ανάγνωση"

Αν επιλεγεί η λειτουργία "Μόνο Ανάγνωση", τότε η καταγραφή δεδομένων είναι ανενεργή. Ο χρήστης, σε αυτή την περίπτωση, διαβάζει το κείμενο που του δίνεται. Ο τρόπος παρουσίασης του κειμένου καθορίζεται από το υπάρχον προφίλ του χρήστη. Καθώς διαβάζει ο χρήστης, έχει τη δυνατότητα να προσαρμόζει τον τρόπο παρουσίασης του κειμένου (ταχύτητα ανάγνωσης, μέγεθος γραμματοσειράς, φωτεινή υπογράμμιση). Επίσης, ο χρήστης μπορεί να ενεργοποιεί ή να απενεργοποιεί

τη λειτουργία "μετατροπή κειμένου σε ήχο" που το σύστημα μπορεί να διαθέτει.

Μια άλλη εφαρμογή της λειτουργίας "Μόνο Ανάγνωση" μπορεί να είναι όταν το σύστημα χρησιμοποιείται είτε σε παρωχημένο Η/Υ ή σε υπολογιστή χειρός. Το σύστημα πρέπει να εκπαιδευτεί έτσι ώστε ένα προσωπικό προφίλ να "φορτώνεται". Αυτή η λειτουργία κατά ένα τρόπο απενεργοποιεί τη Μηχανή Διαχείρισης Γνώσης και Προφίλ Χρηστών, αφού την "αναγκάζει" να φορτώνει το προφίλ του χρήστη και βάση αυτού να πραγματοποιείται η παρουσίαση του εκπαιδευτικού υλικού, χωρίς να λαμβάνονται υπόψη οι εναλλαγές της κατάστασης του χρήστη ή η ικανότητα ανάγνωσης που αυτός έχει τη συγκεκριμένη στιγμή.

7.4 3ο σενάριο: Σύστημα εκπαίδευσης σε αγιογραφίες της Βυζαντινής τέχνης

Στο πλαίσιο του 3ου σεναρίου αναπτύχθηκε ένα σύστημα εκπαίδευσης σε θέματα τέχνης και συγκεκριμένα σε ζωγραφική της Βυζαντινής τέχνης. Το σύστημα βασίστηκε στην προτεινόμενη επέκταση της αρχιτεκτονικής του IEEE, όπως αυτή παρουσιάστηκε στην Ενότητα 2.5. Την ομάδα των ειδικών του συστήματος απαρτίζουν ιστορικοί τέχνης, μηχανικοί σε τεχνολογίες γνώσης και ερευνητές, ενώ την ομάδα των χρηστών απαρτίζουν ομάδες μαθητών, φοιτητών, μελετητών, ερευνητών, ακαδημαϊκών καθώς και φίλων της τέχνης.

Σκοπός του συστήματος μάθησης είναι η μοντελοποίηση ιστάμενων μορφών, έτσι ώστε να υποστηρίζονται διαδικασίες όπως η εκπαίδευση σχετικά με τις τεχντροπίες έργων τέχνης, η τεκμηρίωση ιστάμενων μορφών, καθώς και η αναζήτηση και εξόρυξη μορφών και παραστάσεων με βάση το περιεχόμενο. Η μοντελοποίηση αφορά στις αναλογίες των μορφών (λόγοι μεγεθών κορμού - κεφαλιού), το σχήμα του προσώπου των μορφών καθώς και τμημάτων αυτού, στοιχεία τα οποία παρέχουν πολύ σημαντική πληροφορία για τη μορφή, τη συσχέτιση τους καθώς και για χαρακτηριστικά στοιχεία τεχντροπίας.

Το βασικό πολυμεσικό υλικό που χρησιμοποιήθηκε αποτελείται από τα σιναιϊτικά μνημεία [100], τα οποία λόγω της μακράς ιστορικής τους διαδρομής, παρέχουν επαρκές δείγμα για τη μελέτη τόσο της αρχαίας παραδόσεως όσο και της νεότερης. Οι σιναιϊτικές συλλογές έργων τέχνης είναι από τις σημαντικότερες διεθνώς και περιλαμβάνουν κορυφαία έργα της παγκόσμιας πολιτιστικής κληρονομιάς. Επιπλέον, σε αυτές αντιπροσωπεύονται κατά μοναδικό τρόπο και σε αδιάσπαστη συνέχεια, από την παλαιохριστιανική εποχή μέχρι σήμερα, έργα τέχνης διαφόρων κατηγοριών (εικόνες, τοιχογραφίες, ψηφιδωτά, μικρογραφίες χειρογράφων, κεντήματα, έργα μικροτεχνίας-μικρογλυπτικής κ.λπ.), υλικών, τεχντροπιών και γεωγραφικής προέλευσης.

Στο πλαίσιο αυτό η γνώση του πεδίου όπως αυτή αποδόθηκε από τους ιστορικούς τέχνης, στηριζόμενοι στο εγχειρίδιο του Διονυσίου εκ Φουρνά [132] (Ενότητα 4.3) οργανώθηκε σε οντολογία (5.4.1, Παράρτημα Α). Αναλυτικότερα, εξετάστηκαν οι τρόποι με τους οποίους κατά κανόνα οι ιστορικοί τέχνης ονοματίζουν το σύνολο και τα επί μέρους των προσώπων των Μορφών, όταν αυτές απεικονίζονται ιστάμενες (ολόσωμες ή ημίτομες). Η οντολογία αυτή φιλοξενεί τα επιμέρους πεδία του προσώπου και περιέχει τις δομικές τους πληροφορίες (Παράρτημα Α) και βάσει αυτών περιέχει πληροφορίες σχετικά με συγκεκριμένες μορφές Αγίων. Την οντολογία αυτή χειρίζεται η διαδικασία Δάσκαλος της προτεινόμενης επέκτασης του προτύπου (Ενότητα 2.5).

Η διαδικασία Δάσκαλος έχει εμπλουτιστεί και με το σύστημα αυτόματης

κατηγοριοποίησης Βυζαντινών έργων τέχνης, όπως αυτό παρουσιάστηκε στην Ενότητα 4.4.2. Αποτέλεσμα αυτού του συστήματος, που λαμβάνει ως είσοδο μια εικόνα που περιέχει μια μορφή, είναι η δημιουργία ενός σώματος ισχυρισμού (Ενότητα 4.4.3.1), το οποίο δίνει πληροφορίες για το περιεχόμενο της εικόνας και φυλλάσσεται μαζί με την εικόνα στην αποθήκη *Εκπαιδευτικό Υλικό* της προτεινόμενης επέκτασης του προτύπου του IEEE (Ενότητα 2.5). Στη διαδικασία *Δάσκαλος* πλέον περιλαμβάνει και τη μονάδα αναπαράστασης γνώσης και συλλογιστική όπως αυτή παρουσιάστηκε στην Ενότητα 5.4.1. Κατά αυτόν τον τρόπο η προτεινόμενη επέκταση της αρχιτεκτονική συστήματος μαθησιακής τεχνολογίας του IEEE έχει εμπλουτιστεί με υπηρεσίες συμπερασματολογίας και είναι σε θέση να κατηγοριοποιήσει το πολιτιστικό υλικό.

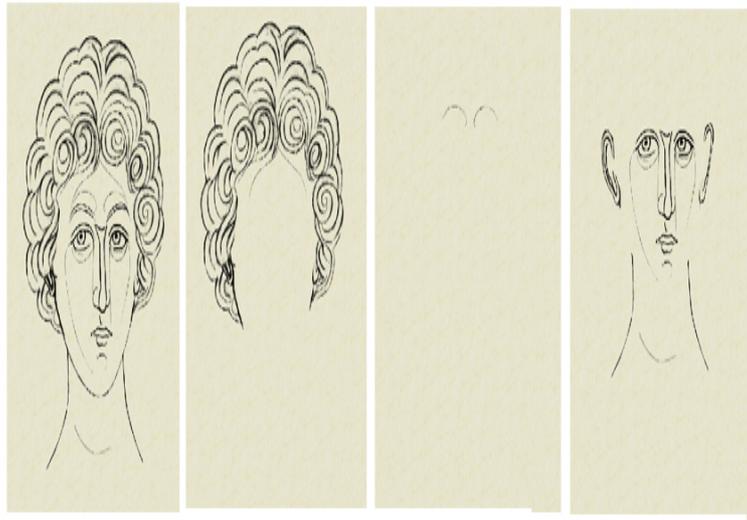
Σε αυτό το σενάριο, όλο το πολιτιστικό υλικό έχει αυτόματα κατηγοριοποιηθεί και συνοδεύεται από το σώμα ισχυρισμών του. Το προφίλ του χρήστη θεωρείται προφίλ αρχαρίου σε αυτό το σενάριο. Το σενάριο αποτελείται από δύο τμήματα το τμήμα της ανάλυσης και το τμήμα της σύνθεσης, τα οποία παρουσιάζονται στο χρήστη μέσω της ροής *Παρουσίαση*. Αρχικά ο χρήστης έχει τη δυνατότητα να επιλέξει μια εικόνα και να την 'αναλύσει' δηλαδή να δει τα χαρακτηριστικά τμήματά της, όπως αυτά εξήχθησαν από σύστημα αυτόματης κατηγοριοποίησης Βυζαντινών έργων τέχνης (Σχήμα 4.4.γ) καθώς και το παραγόμενο σώμα ισχυρισμών (ABox). Με αυτό τον τρόπο ο χρήστης 'μαθαίνει' να ξεχωρίζει τα χαρακτηριστικά του προσώπου της κάθε μορφής.

Έπειτα, μαζί με το σώμα ισχυρισμών παρουσιάζεται στο χρήστη και το σώμα ορολογίας με σκοπό την εξοικείωσή του τόσο με τα αξιώματα όσο και με την οντολογία. Στόχος του τμήματος ανάλυσης είναι η μεθοδική εκμάθηση ορολογίας σχετιζόμενης με τις ιστάμενες (ολόσωμες ή ημίτομες) μορφές, που είναι απαραίτητη στις σπουδές και τις εφαρμογές της αρχαιολογίας και της ιστορίας της τέχνης.

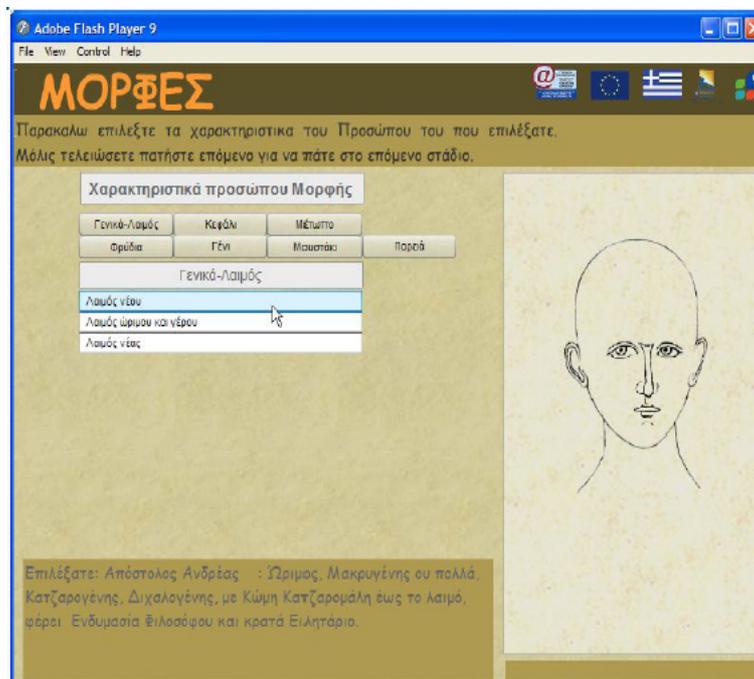
Το τμήμα της σύνθεσης (Σχήμα 7.8) αποτελεί το επόμενο στάδιο του σεναρίου όπου ζητείται από το χρήστη βάσει αυτών που έμαθε στο στάδιο της ανάλυσης να συνθέσει τη μορφή ενός συγκεκριμένου Αγίου. Το τμήμα αυτό περιέχει συνολικά 23 ψηφιοποιημένα ανθίβολα προσώπων διαφορετικών αγίων τα οποία έχουν καταταμηθεί σε σημασιολογικά τμήματα, όπως αυτά ορίστηκαν στην Ενότητα 4.2, δημιουργώντας 95 διαφορετικά τμήματα (Σχήμα 7.7). Τα ανθίβολα, όπως ονομάζονται ήδη από τα βυζαντινά χρόνια, είναι σχέδια σε χαρτί για την αντιγραφή και τη μεταφορά παραστάσεων από τοιχογραφίες και εικόνες σε νέα έργα. Με τον τρόπο αυτό διαλεγμένα εικονογραφικά θέματα επαναλαμβάνονται και μεταδίδονται από ζωγράφο σε ζωγράφο, από ναό σε ναό, από περιοχή σε περιοχή. Τα ανθίβολα, σύμφωνα με τις τεχνικές παραγωγής τους, διακρίνονται σε έκτυπα, διάτρητα ή και ελεύθερα σχέδια. Τα ανθίβολα, που χρησιμοποιήθηκαν στην παρούσα εργασία, ανήκουν στο 'Ίδρυμα του Όρους Σινά [100].

Τα ανθίβολα για κάποιο νέο Άγιο φαίνονται στο Σχήμα 7.7 στην πρώτη εικόνα εμφανίζεται το συνολικό ανθίβολο, ενώ στις εικόνες που ακολουθούν εμφανίζονται τα σημασιολογικά τμήματα του προσώπου (μαλλιά, μέτωπο, παρειές, μύτη στόμα, αυτιά και λαιμός).

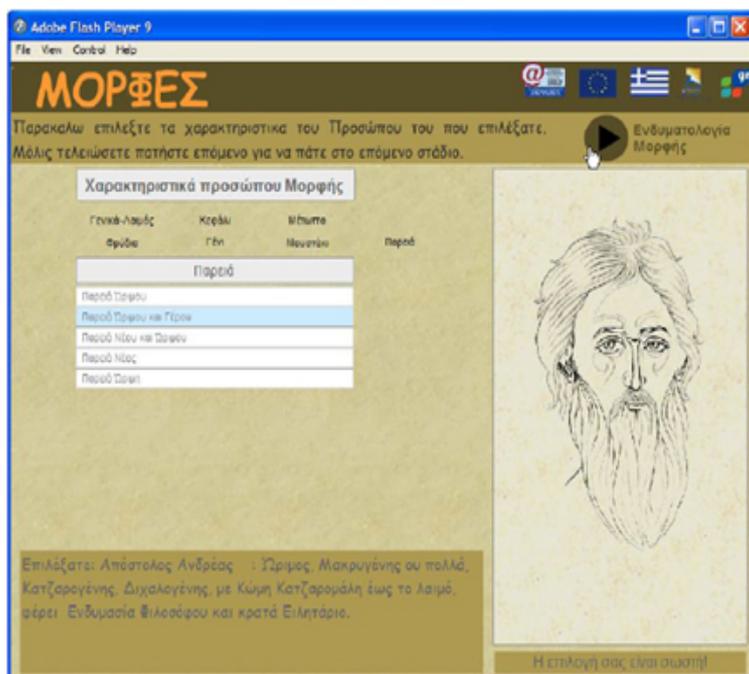
Για παράδειγμα, έστω ότι έχει ζητηθεί από το χρήστη να αποδοθεί η μορφή του Αποστόλου Ανδρέα. Αρχικά εμφανίζεται στο χρήστη μόνο μία γυμνή κεφαλή (Σχήμα 7.8) και του ζητείται να συνθέσει τη ζητούμενη μορφή, επιλέγοντας τα κατάλληλα τμήματα των ανθιβόλων. Παράλληλα του δίνεται η χαρακτηριστική οντολογία της μορφής, στο κάτω τμήμα του παραθύρου (Ώριμος, Μακρυγένης ου πολλά, Κατσαρογένης, Διχαλογένης, με ΚώμηΚατσαρομάλλη).



Σχήμα 7.7: Αντίβολα μορφής νεαρού Αγίου



Σχήμα 7.8: Τμήμα σύνθεσης- Περιβάλλον διεπαφής - αρχική οθόνη



Σχήμα 7.9: Τμήμα σύνθεσης- Περιβάλλον διεπαφής - τελική οθόνη

Πραγματοποιώντας τις σωστές επιλογές ο χρήστης οδηγείται στη μορφή του παραθύρου στο Σχήμα 7.9.

Σε επόμενη μελέτη μπορεί να μετρηθεί η απόδοσή του χρήστη και η μέτρηση αυτή να επηρεάζει το προφίλ του. Ως μελλοντική εργασία θα μπορούσε να θεωρηθεί και η ανάλυση του ενδύματος και των παρελκόμενων στοιχείων που κρατά μια Άγια μορφή.

Κεφάλαιο 8

Συμπεράσματα και κατευθύνσεις μελλοντικής έρευνας

Η καταγραφή των προβλημάτων που παρουσιάζουν τα υπάρχοντα συστήματα μαθησιακής τεχνολογίας ήταν το πρώτο μέλημά μας, στην παρούσα διατριβή. Ένα τέτοιο βασικό πρόβλημα ήταν η ενιαία αντιμετώπιση των χρηστών με ταυτόχρονη ανάλυση των ιδιαιτεροτήτων του κάθε χρήστη.

Με κύριο στόχο την ενοποίηση και την από κοινού εκπαίδευση διαφορετικών χρηστών, η διατριβή επικεντρώθηκε στην αυτόματη κατηγοριοποίηση τόσο των χρηστών όσο και του εκπαιδευτικού υλικού. Από την έρευνά μας, έχουν προκύψει συνεισφορές στην κατηγοριοποίηση των χρηστών και στην κατηγοριοποίηση του περιεχομένου που χρησιμοποιείται για την εκπαίδευση, ενώ εξετάστηκε και η προστασία του πολυμεσικού εκπαιδευτικού υλικού.

Απώτερος σκοπός της διατριβής ήταν ο εμπλουτισμός των συστημάτων μαθησιακής τεχνολογίας με τεχνικές μηχανικής μάθησης, ανάλυσης ψηφιακού περιεχομένου, αναπαράστασης γνώσης και η δημιουργία "ευφυών" εκπαιδευτικών περιβαλλόντων, τα οποία προσαρμόζονται δυναμικά στις ανάγκες του ατόμου, ενισχύοντας παράγοντες όπως η παραγωγικότητα, η μάθηση, η εκφραστικότητα και η δημιουργικότητα. Ουσιαστικά, θέλουμε να πιστεύουμε ότι, μέσω αυτής της διατριβής, "έγιναν εμφανείς" οι τρόποι εμπλουτισμού των προτύπων μαθησιακής τεχνολογίας, με σκοπό τη δημιουργία συστημάτων μάθησης με νοημοσύνη.

Τα κύρια καινοτόμα αποτελέσματα και οι μελλοντικές τους επεκτάσεις μπορούν να συνοψιστούν στα επόμενα:

Ξεκινώντας από ένα ηλεκτρονικό ερωτηματολόγιο συλλέχθηκαν πληροφορίες για τους χρήστες ενός συστήματος μαθησιακής τεχνολογίας. Η ανάλυση αυτών των πληροφοριών οδήγησε στη δημιουργία δυναμικών προφίλ των χρηστών (Ενότητα 3.3). Πολλές φορές αυτά τα ερωτηματολόγια δείχνουν να μην επαρκούν, επειδή είτε είναι ελλιπή είτε δεν είναι απολύτως σωστές οι απαντήσεις των χρηστών. Η τεχνική αυτόματης αξιολόγησης της κατάστασης της συμπεριφοράς των χρηστών (Ενότητα 3.4.3) μπορεί σε αυτή την περίπτωση να λειτουργήσει συμπληρωματικά με σκοπό τον εμπλουτισμό του προσδιορισμού του προφίλ του χρήστη.

Οι τεχνολογίες γνώσης έχουν τη δυνατότητα να συμβάλουν ουσιαστικά στην επέκταση της έρευνας στο πεδίο της κατηγοριοποίησης των χρηστών. Η δημιουργία σώματος ισχυρισμών για τον κάθε χρήστη, καθώς και του αντίστοιχου σώματος ορολογίας μπορούν, με χρήση ασαφών μηχανών συμπερασματολογίας, να εξάγουν το προφίλ των χρηστών. Επιπλέον, μέσα σε ένα περιβάλλον με νοημοσύνη, θα μπορούσε

το σώμα ισχυρισμών για το χρήστη να γνωστοποιείται σε όλες τις έξυπνες συσκευές του περιβάλλοντος αυτού, με αποτέλεσμα τη δημιουργία ενός περιβάλλοντος πλήρως προσαρμοσμένου στις ανάγκες του χρήστη.

Η περιβάλλουσα νοημοσύνη θα πρέπει να έχει κάποια χαρακτηριστικά, έτσι ώστε να είναι δυνατή η κοινωνική της αποδοχή. Σε αυτή την κατεύθυνση, στόχος των συστημάτων με περιβάλλουσα νοημοσύνη θα πρέπει να είναι η διευκόλυνση της ανθρώπινης επαφής καθώς και η απόκτηση γνώσεων και δεξιοτήτων ικανών να βελτιώσουν τις καθημερινές δραστηριότητες του ατόμου. Τα συστήματα αυτά θα πρέπει να εμπνέουν εμπιστοσύνη και αξιοπιστία στους χρήστες του. Βέβαια, εδώ, θα πρέπει να αναφερθούμε στην ηθική πλευρά τέτοιων εφαρμογών και να τονίσουμε τη σπουδαιότητα παραχώρησης του πλήρους ελέγχου των συστημάτων στον ίδιο το χρήστη. Επιπλέον, τα συστήματα αυτά θα πρέπει να είναι ικανά για μακροπρόθεσμη υποστήριξη (προσωπική και κοινωνική) των χρηστών τους και να συμβάλλουν στη δια βίου εκπαίδευση αυτών.

Όσον αφορά στο εκπαιδευτικό υλικό και την κατηγοριοποίησή του, στο πλαίσιο της προταθείσας αρχιτεκτονικής συστήματος μαθησιακής τεχνολογίας, η έρευνα μελλοντικά πρέπει να κινηθεί προς την κατεύθυνση της αυτόματης τεκμηρίωσης πολιτιστικού πολυμεσικού υλικού από διάφορα πεδία τέχνης, με σκοπό τη χρήση τους μέσα σε ένα ενιαίο σύστημα ηλεκτρονικής μελέτης και εξοικείωσης με εικαστικές τέχνες. Κατά παρόμοιο τρόπο, η έρευνα για την τεκμηρίωση υλικού μέσα από ένα ενιαίο πρίσμα, ικανό να χρησιμοποιηθεί για εκπαίδευση διαφορετικών ειδών χρηστών και τύπων περιεχομένου, σε συνδυασμό με τις τεχνολογίες γνώσης, θα οδηγήσει σε ένα ενιαίο σύστημα ηλεκτρονικής μάθησης.

Κατάλογος δημοσιεύσεων

Περιοδικά

1. Paraskevi Tzouveli, Nikos Simou, Giorgios Stamou, and Stefanos Kollias. Semantic classification of byzantine icons. *IEEE Intelligent Systems*, 24(2):35–43, Mar./Apr. 2009.
2. S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias. Estimation of behavioral user state based on eye gaze and head pose-application in an e-learning environment. *Multimedia Tools and Applications, Springer*, 41(3):469–493, 2009.
3. P. Tzouveli, Ph. Mylonas, and S. Kollias. An intelligent e-learning system based on learner profiling and learning resources adaptation. *Computers & Education, Elsevier*, 51(1):224–238, August 2008.
4. Ph. Mylonas, P. Tzouveli, and S. Kollias. An integrated approach towards intelligent educational content adaptation. *Emerging Artificial Intelligence Applications in Computer Engineering, Frontiers in Artificial Intelligence and Applications, IOS Press*, 160, October 2007.
5. Ph. Mylonas, P. Tzouveli, and S. Kollias. E-learning and intelligent content adaptation: an integrated approach. *International Journal of Continuing Engineering Education and Life-Long Learning, Inderscience*, 17(4/5):273–293, September 2007.
6. P. Tzouveli and S. Kollias. Using e-questionnaire as an improvement to a learners' adaptation system. *Learning Technology Newsletter*, 6, 10 2004.
7. P. Tzouveli, K. Ntalianis, and S. Kollias. Protection of Real and Artwork Human Objects based on a Moments Modulation Method *Journal of Information Assurance and Security*, June 2009, *under review*.
8. G. Caridakis, P. Tzouveli, A. Raouzaïou, K. Karpouzis and S. Kollias Analysis of Learners' State in an Integrated E-learning System *Affective, Interactive and Cognitive Methods for E-Learning Design: Creating an Optimal Education Experience*, 2009, *accepted*.

Βιβλία

9. P. Tzouveli and S. Kollias. *E-questionnaire for innovative adaptive-learning scheme*, volume 2. Encyclopedia of E-Commerce, E-Government, and Mobile Commerce, information resources management association, idea group reference edition, 3, 2006.
10. P. Tzouveli, G. Andreou, G. Tsechpenakis, Y. Avrithis, and S. Kollias. *Intelligent Visual Descriptor Extraction from Video Sequences*, volume 3094 of *Lecture Notes in Computer Science ∪ Adaptive Multimedia Retrieval*. Springer-Verlag, 2004.
11. G. Caridakis, P. Tzouveli, A. Raouzaïou, K. Karpouzis and S. Kollias. Analysis of Learners' State in an Integrated E-learning System., *Affective, Interactive and Cognitive Methods for E-Learning Design: Creating an Optimal Education Experience*, March 2009, *accepted*.

Συνέδρια

12. P. Tzouveli, N. Simou, G. Stamou, S. Kollias, D. Kalomoirakis, G. Foukaneli, and N. Fyssas. Sacred figure recognition based on byzantine iconography knowledge. In *Proceedings of the International Conference of Hellenic Ministry of Cultural on Digital Culture in a new knowledge enviroment*, November 2008.
13. P. Tzouveli, K. Ntalianis, and S. Kollias. Semantic region protection using hu moments and a chaotic pseudo-random number generator. In *Proceedings of the International IEEE Workshop on Computational Intelligence in Security for Information Systems (CISIS'08)*, October 2008.
14. K. Raftopoulos, P. Tzouveli, K. Ntalianis, S. Kollias, D.Kalomoirakis, N. Fyssas, and G. Foukaneli. Saint classification in byzantine art. In *Proceedings of the 14th International Conference on Virtual Systems and MultiMedia (VSMM'08)*, October 2008.
15. P. Tzouveli, A. Schmidt, M.Schneider, A. Symvonis, and S. Kollias. Adaptive reading assistance for the inclusion of students with dyslexia: The agent-dysl approach. In *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008)*, July 2008.
16. V. Giannekou, P. Tzouveli, Y. Avrithis, and S.Kollias. Affine normalization curve matching method using curvaturescale-space and normalization. In *Proceedings of the 6th International Workshop on Content-Based Multimedia Indexing (CBMI 2008)*, June 2008.

17. S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias. Non-verbal feedback on user interest based on gaze direction and head pose. In *Proceedings of the 2nd International Workshop on Semantic Media Adaptation and Personalization (SMAP 2007)*, December 2007.
18. P. Tzouveli, E. Mitropoulou, K. Ntalianis, S. Kollias, and A. Symvonis. Design of an accommodative intelligent educational environments for dyslexic learners. In *Proceedings of the 11th Conference on Learning Difficulties in the Framework of School Education*, November 2007.
19. P. Tzouveli, K. Ntalianis, A. Symvonis, and S. Kollias. An intelligent adaptable reading system targeting people with learning difficulties. In *Proceedings of the 4th Greek conference on ICT and Education*, May 2007.
20. Paraskevi Tzouveli, Klimis Ntalianis, and Stefanos Kollias. Video object watermarking using hu moments. In *Proceedings of the 13th International Conference on Systems, Signals and Image Processing (IWSSIP'06)*, September 2006.
21. P. Tzouveli, K. Ntalianis, and S. Kollias. Confronting the synchronization problem of semantic region under geometric attacks. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2006)*, July 2006.
22. P. Tzouveli, K. Ntalianis, and S. Kollias. Chaotic encryption driven watermarking of human video objects based on hu moments. In *Proceedings of the 3rd IFIP Conference on Artificial Intelligence Applications and Innovations (AIAI 2006)*, June 2006.
23. P. Tzouveli, Y. Avrithis, and S. Kollias. Fast video object tracking using affine invariant normalization. In *Proceedings of the 3rd IFIP Conference on Artificial Intelligence Applications and Innovations (AIAI 2006)*, June 2006.
24. P. Tzouveli, K. Ntalianis, and S. Kollias. Human face watermarking based on zernike moments. In *Proceedings of the IEEE Symposium on Signal Processing and Information Technology (ISSPIT'05)*, December 2005.
25. P. Tzouveli, K. Ntalianis, and S. Kollias. Human video object watermarking based on hu moments. In *Proceedings of the IEEE Workshop on Signal Processing Systems (SiPS'05)*, November 2005.
26. P.Tzouveli, K. Ntalianis, and S. Kollias. Video objects watermarking: Hu versus zernike moments. In *Proceedings of the International Workshop on Very Low Bitrate Video Coding (VLBV'05)*, September 2005.
27. P. Tzouveli, Ph. Mylonas, and S. Kollias. Spero - a personalized integrated e-learning system. In *Proceedings of the International Association for Development of the Information Society (IADIS WWW/Internet)*, October 2004.
28. P.Tzouveli, K.Ntalianis, and S.Kollias. Security of human video objects by incorporating a chaos-based feedback cryptographic scheme. In *Proceedings of the ACM Multimedia 2004*, October 2004.

29. P.Tzouveli and S. Kollias. User adaptation in e-learning environments. In *Proceedings of the IFIP World Computer Congress, Technology Enhanced Learning (TEL'04)*, August 2004.
30. Ph. Mylonas, P. Tzouveli, and S. Kollias. Towards a personalized e-learning scheme for teachers. In *Proceedings of the 4th IEEE International Conference on Advanced Learning Technologies (ICALT'04)*, August 2004.
31. K. Karpouzis, A. Raouzaïou, P. Tzouveli, and S. Kollias. Mpeg-4: One multimedia standard to unite all. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME'04)*, July 2004.
32. P. Tzouveli, K. Ntalianis, and S. Kollias. Automatic videoconference objects watermarking using object adapted qualified significant wavelet trees. In *Proceedings of the International Workshop Very Low Bitrate Video Coding (VLBV03)*, September 2003.
33. P. Tzouveli, G. Andreou, G. Tsechpenakis, Y. Avrithis, and S. Kollias. Intelligent visual descriptor extraction from visual sequences. In *Proceedings of the International Workshop on Adaptive Multimedia Retrieval (KI'03)*, September 2003.
34. P.Tzouveli, N. Tsapatsoulis, S. Kollias, and M. Michaelidou. On the development of user adapted e-learning schemes. In *Proceedings of the International Conference on New Horizons in Industry and Education (NHIE'03)*, August 2003.
35. P.Tzouveli, K. Ntalianis, N. Tsapatsoulis, and S.Kollias. Automatic face region watermarking using qualified significant wavelet trees. In *Proceedings of the 9th International Workshop on Systems, Signal and Image Processing (IWSSIP'02)*, November 2002.

□

Βιβλιογραφία

- [1] Καριώτογλου Π., Ψύλλος Δ., Κουμαράς Π. Εποικοδόμηση της γνώσης στην τάξη με συνέρευνα δασκάλου και μαθητή. *Σύγχρονη Εκπαίδευση*, 70:34–42, 1993.
- [2] Ν. Τσέλιος Ν. Αβούρης, Κ. Σολωμός. Το διαδίκτυο ως εργαλείο παροχής ανοικτής και εξ-αποστάσεως εκπαίδευσης: εναλλακτικοί μηχανισμοί ελέγχου ποιότητας. In *1ο Πανελλήνιο Συνέδριο στην ανοικτή και εξ αποστάσεως εκπαίδευση*, 2001.
- [3] Γεώργιος Μπαμπινιώτης. Η μαγεία της γραμματικής. *Το Βήμα*, (15349):1, 5 2008.
- [4] YS Abu-Mostafa and D. Psaltis. Image normalization by complex moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):46–55, 1985.
- [5] NI Achieser. Lectures on the theory of approximation, moscow, 1947. *English Translation, Theory of Approximation, Ungar, New York*, 1956.
- [6] AGENT-DYSL. Accommodative intelligent educational environments for dyslexia learners.
- [7] M. Alghoniemy and AH Tewfik. Geometric invariance in image watermarking. *IEEE Transactions on Image Processing*, 13(2):145–153, 2004.
- [8] Y. Amit and D. Geman. A computational model for visual selection. *Neural Computation*, 11(7):1691–1715, 1999.
- [9] J.M. Applefield, R. Huber, and M. Moallem. Constructivism in theory and practice: Toward a better understanding. *High School Journal*, 84(2):35–53, 2001.
- [10] International Dyslexia Association.
- [11] S. Asteriadis, K. Karpouzis, and S. Kollias. A neuro-fuzzy approach to user attention. In *Proceedings of the 18th International Conference on Artificial Neural Networks (ICANN 2008)*, November 2008.
- [12] S. Asteriadis, N. Nikolaidis, A. Hajdu, and I. Pitas. An eye detection algorithm using pixel to edge information. In *Proceedings of Second IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing (ISCCSP 2006)*, Marrakech, Morocco, 2006.

- [13] S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias. Estimation of behavioral user state based on eye gaze and head pose application in an e-learning environment. *Multimedia Tools and Applications*, 41(3):469–493, Springer 2009.
- [14] P. Aygeriou. *Μια αρχιτεκτονική αναφοράς για ανοικτά συστήματα διαχείρισης της μάθησης*. PhD thesis, 2003.
- [15] F. Baader and W. Nutt. Basic description logics. In Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors, *The Description Logic Handbook: Theory, Implementation, and Applications*, pages 43–95. Cambridge University Press, 2003.
- [16] Franz Baader, Ian Horrocks, and Ulrike Sattler. Description Logics for the Semantic Web. *KI – Künstliche Intelligenz*, 16(4):57–59, 2002.
- [17] DH Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern recognition*, 13(2):111–122, 1981.
- [18] S. Baron-Cohen. *Mindblindness: An essay on autism and theory of mind*. MIT press, 1995.
- [19] P. Bas and B. Macq. A new video-object watermarking scheme robust to objectmanipulation. 3, 2001.
- [20] J. Beckwith and R. Krautheimer. *Early Christian and Byzantine Art*. Yale Univ Pr, 1986.
- [21] SO Belkasim, M. Shridhar, and M. Ahmadi. Pattern recognition with moment invariants: a comparative study and new results. *Pattern Recognition*, 24(12):1117–1138, 1991.
- [22] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. pages 34–43, May 2001.
- [23] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284(5):34–43, 2001.
- [24] D. Beymer, M. Flickner, I.B.M.A.R. Center, and CA San Jose. Eye gaze tracking using an active stereo head. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*, volume 2, 2003.
- [25] B. Bixler and J. Spotts. Screen design and levels of interactivity in web-based training. In *IVLA Conference (International Visual Literacy Association)*, 1998.
- [26] M. Bohme, A. Meyer, T. Martinetz, and E. Barth. Remote eye tracking: State of the art and directions for future development. In *Proc. of the 2006 Conference on Communication by Gaze Interaction (COGAIN)*, pages 12–17, 2006.

- [27] R.M. Bottino. The evolution of ict-based learning environments: which perspectives for the school of the future? *British Journal of Educational Technology*, 35(5):553–567, 2004.
- [28] J.Y. Bouguet. Pyramidal implementation of the lucas kanade tracker. OpenCV Documentation, 2000.
- [29] Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In *Proceedings of the International Conference on Computer Vision (ICCV 2001)*, September 2001.
- [30] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
- [31] I. Bruha and F. Franek. Comparison of various routines for unknown attribute value processing: the covering paradigm. *International Journal of Pattern Recognition and Artificial Intelligence*, 10:939–956, 1996.
- [32] C.J.C Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
- [33] G. Caridakis, K. Karpouzis, and S. Kollias. User and context adaptive neural networks for emotion recognition. 2008.
- [34] S.F. Chang, T. Sikora, and A. Puri. Overview of the mpeg-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6):688–695, 2001.
- [35] M. Chatzidakis. *Hellines Zographoi Meta Thn Allosi: 1450-1830*. Kentro Neoellinikon Ereunon, 1987.
- [36] G. Collier. E-learning application infrastructure. *Sun Microsystems White Paper*, 2002.
- [37] T. Coppeto, J. Kahn, S. Thorne, and J. Merriman. Integrating e-learning applications using oki. 2006.
- [38] J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6:1673–1687, December 1997.
- [39] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(3):131–156, 1997.
- [40] JG Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1161, 1993.
- [41] J.Y. Deng and F. Lai. Region-based template deformation and masking for eye-feature extraction and description. *Pattern recognition*, 30(3):403–419, 1997.

- [42] P. Dollar, Z. Tu, H. Tao, and S. Belongie. Feature mining for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR'07*, pages 1–8, 2007.
- [43] P. Domingos and M. Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Machine Learning*, 29(2):103–130, 1997.
- [44] Francesco M. Donini, Maurizio Lenzerini, Daniele Nardi, and Andrea Schaerf. Deduction in concept languages: From subsumption to instance checking. *Journal of Logic and Computation*, 4(4):423–452, 1994.
- [45] W3C Working Draft. How people with disabilities use the web, December 2004.
- [46] A.T. Duchowski. A breadth-first survey of eye-tracking applications. *Behavior Research Methods Instruments and Computers*, 34(4):455–470, 2002.
- [47] R.O. Duda and P.E Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.
- [48] J. Elkind. Computer reading machines for poor readers. *Perspectives*, 24(2):9–13, 1998.
- [49] F. Fleuret and D. Geman. Coarse-to-fine face detection. *International Journal of Computer Vision*, 41(1):85–107, 2001.
- [50] Y. Freund and R.E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory: Second European Conference, EuroCOLT'95, Barcelona, Spain, March 13-15, 1995: Proceedings*, page 23. Springer, 1995.
- [51] DR Fulkerson and LR Ford. *Flow in Networks*. Princeton University Press, 1962.
- [52] AH Gee and R. Cipolla. Determining the gaze of faces in images. *Image and Vision Computing*, 1994.
- [53] N. Gourier, J. Maisonnasse, D. Hall, and J.L. Crowley. Head pose estimation on low resolution images. *Lecture Notes in Computer Science*, 4122:270, 2007.
- [54] J. C. Gower. A general coefficient of similarity and some of its properties. *Biometrics*, 23:623–637, 1971.
- [55] DM Greig, BT Porteous, and AH Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 271–279, 1989.
- [56] GUNet. Προτυποποίηση μαθησιακών τεχνολογιών, 2003.
- [57] J.M. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. *Unpublished manuscript*, 3, 1971.
- [58] L.A. Harwood. A chrominance demodulator ic with dynamic flesh correction. *IEEE Transaction Consumer Electronics*, 22:111–117, 1976.

- [59] S. Haykin. *Neural Networks: a Comprehensive Foundation*. Prentice Hall, 2008.
- [60] L. Hecker, L. Burns, L. Katz, J. Elkind, and K. Elkind. Benefits of assistive reading software for students with attention disorders. *Annals of Dyslexia*, 52(1):243–272, 2002.
- [61] B. Heisele, T. Serre, and T. Poggio. A component-based framework for face detection and identification. *International Journal of Computer Vision*, 74(2):167–181, 2007.
- [62] C. Hennessey, B. Nouredin, and P. Lawrence. A single camera eye-gaze tracking system with free head motion. In *Proceedings of the 2006 symposium on Eye tracking research & applications*, pages 87–94. ACM New York, NY, USA, 2006.
- [63] I. Horrocks and U. Sattler. A tableaux decision procedure for *SHOIQ*. In *Proc. 19th Int. Joint Conf. on Artificial Intelligence (IJCAI 05)*, 2005.
- [64] M.S. Hsieh, D.C. Tseng, and Y.H. Huang. Hiding digital watermarks using multiresolution wavelet transform. *IEEE Transactions on Industrial Electronics*, 48(5):875–882, 2001.
- [65] C.-T. Hsu and J.-L. Wu. Dct-based watermarking for video. *IEEE Transactions on Consumer Electronics*, 44:206–216, February 1998.
- [66] M.K. Hu. Visual pattern recognition by moment invariants. *Information Theory, IRE Transactions on*, 8(2):179–187, 1962.
- [67] C. Huang, H. Ai, Y. Li, and S. Lao. High-performance rotation invariant multiview face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 671–686, 2007.
- [68] KS Huang and MM Trivedi. Robust real-time detection, tracking, and pose estimation of faces in video streams. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 3, 2004.
- [69] Z. Huang. Clustering large data sets with mixed numeric and categorical values. In *The First Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 21–34, 1997.
- [70] S. Ioannou. *Αποτελεσματικές τεχνικές εξαγωγής χαρακτηριστικών και ανάλυσης εκφράσεων στην επικοινωνία ανθρώπου μηχανής*. PhD thesis, 2006.
- [71] S.V. Ioannou, A.T. Raouzaïou, V.A. Tzouvaras, T.P. Mailis, K.C. Karpouzis, and S.D. Kollias. Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, 18(4):423–435, 2005.
- [72] A.K. Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 1989.
- [73] CR Johnson, E. Hendriks, IJ Berezchnoy, E. Brevdo, SM Hughes, I. Daubechies, J. Li, E. Postma, and JZ Wang. Image processing for artist identification. *IEEE Signal Processing Magazine*, 25(4):37–48, 2008.

- [74] M. Jones and P. Viola. Fast multi-view face detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.
- [75] C. Karagiannidis, DG Sampson, and F. Cardinali. An architecture for web-based e-learning promoting reusable adaptive educational e-content. *Educational Technology & Society*, 5(4), 2002.
- [76] K. Karpouzis, G. Caridakis, L. Kessous, N. Amir, A. Raouzaïou, L. Malatesta, and S. Kollias. Modeling naturalistic affective states via facial, vocal, and bodily expressions recognition. *LECTURE NOTES IN COMPUTER SCIENCE*, 4451:91, 2007.
- [77] E. Kitzinger. *Byzantine Art in the Making: Main Lines of Stylistic Development in Mediterranean Art, 3Rd-7Th Century*. Harvard University Press, 1977.
- [78] E. Koch and J. Zhao. Towards robust and hidden image copyright labeling. In *Workshop on Nonlinear Signal and Image Processing*, pages 452–455.
- [79] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97:273–324, 1997.
- [80] I. Kokkinos. *Synergy between Image Segmentation and Object Recognition using Geometrical and Statistical Computer Vision Techniques*. PhD thesis, Electrical and Computer Engineering, National Technical University of Athens., 2006.
- [81] D. Kundur and D. Hatzinakos. A robust digital image watermarking method using wavelet-based fusion. In *Image Processing, 1997. Proceedings., International Conference on*, volume 1, 1997.
- [82] H. Leung, STS Wong, and HHS Ip. In the name of art. *IEEE Signal Processing Magazine*, 25(4):49–54, 2008.
- [83] A.S. Lillard. Young children’s conceptualization of pretense: action or mental representational state? *Child Development*, pages 372–386, 1993.
- [84] Claro Software Ltd. Claro software for dyslexia.
- [85] Sensory Software International Ltd. Sensory software international ltd.,
- [86] LTSA. Standard for learning technology-learning technology systems architecture (ltsa), 2003.
- [87] S.C. Marsella, D.V. Pynadath, and S.J. Read. Psychsim: Agent-based modeling of social interactions and influence. In *Proceedings of the International Conference on Cognitive Modeling*, pages 243–248, 2004.
- [88] J.C. Martin, G. Caridakis, L. Devillers, K. Karpouzis, and S. Abrilian. Manual annotation and automatic image processing of multimodal emotional behaviors: validating the annotation of tv interviews. *Personal and Ubiquitous Computing*, 13(1):69–76, 2009.

- [89] Y. Matsumoto, T. Ogasawara, and A. Zelinsky. Behavior recognition based on head pose and gaze direction measurement. In *2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings*, volume 3, 2000.
- [90] T. Meier and N. King. A flexible bayesian framework for image segmentation. In *Proceedings of the International Conference on Image Processing*, volume 3, pages 203–207, 1999.
- [91] E. Mendelson. *Introduction to Mathematical Logic*. Chapman & Hall/CRC, 1997.
- [92] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1):63–86, 2004.
- [93] NE Mitrakis, JB Theocharis, and V. Petridis. A multilayered neuro-fuzzy classifier with self-organizing properties. *Fuzzy Sets and Systems*, 159(23):3132–3159, 2008.
- [94] A. Mohan, C. Papageorgiou, T. Poggio, K. Commun, and R. City. Example-based object detection in images by components. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(4):349–361, 2001.
- [95] Ph. Mylonas, P. Tzouveli, and S. Kollias. An integrated approach towards intelligent educational content adaptation. *Emerging Artificial Intelligence Applications in Computer Engineering, Frontiers in Artificial Intelligence and Applications*, IOS Press, 160, October 2007.
- [96] Bernhard Nebel. *Reasoning and Revision in Hybrid Representation Systems*. Springer-Verlag, 1990.
- [97] N. Nikolaidis and I. Pitas. Robust image watermarking in the spatial domain. *Signal Processing*, 66(3):385–403, 1998.
- [98] N. Tsapatsoulis, Y. Avrithis, and S. Kollias. Facial image indexing in multimedia databases. *Pattern Analysis and Applications*, 2001.
- [99] D. Obolensky. *The Byzantine Commonwealth*. Weidenfeld and Nicolson, 1971.
- [100] Holy Monastery of St. Catherine at Mount Sinai. Web site.
- [101] N. Otsu. A threshold selection method from gray-level histogram. *IEEE Trans. Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
- [102] C. Papageorgiou and T. Poggio. A trainable system for object detection. *International Journal of Computer Vision*, 38(1):15–33, 2000.
- [103] A. Pelagotti, AD Mastio, AD Rosa, and A. Piva. Multispectral imaging of paintings. *IEEE Signal Processing Magazine*, 25(4):27–36, 2008.
- [104] K. Petridis, S. Bloehdorn, C. Saathoff, N. Simou, S. Dasiopoulou, V. Tzouvaras, S. Handschuh, Y. Avrithis, Y. Kompatsiaris, and S. Staab. Knowledge representation and semantic annotation of multimedia content. *IEE Proceedings-Vision, Image and Signal Processing*, 153(3):255–262, 2006.

- [105] M.T. Pham and T. Cham. Fast training and selection of haar features using statistics in boosting-based face detection. In *Proceedings of the International Conference on Computer Vision*, pages 1–7, 2007.
- [106] SPERO profiles. Spero profiles: (<http://www.image.ntua.gr/statspero/profiles>).
- [107] A. Protopapas and C. Skaloumbakas. Traditional and computer-based screening and diagnosis of reading disabilities in Greek. *Journal of learning disabilities*, 40(1):15, 2007.
- [108] R. Quillian. *Word Concepts: A Theory and Simulation of some basic capabilities*. Behavioral Science, 1967.
- [109] K. Raftopoulos, P. Tzouveli, K. Ntalianis, S. Kollias, D.Kalomoirakis, N. Fyssas, and G. Foukaneli. Saint classification in byzantine art. In *Proceedings of the 14th International Conference on Virtual Systems and MultiMedia (VSMM'08)*, October 2008.
- [110] A. Ribes, R. Pillay, F. Schmitt, and C. Lahanier. Studying that smile. *IEEE Signal Processing Magazine*, 25(4):14–26, 2008.
- [111] D.T Rice. *Art of the Byzantine Era*. Thames and Hudson London, 1963.
- [112] HA Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. In *1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96*, pages 203–208, 1996.
- [113] P. Salembier, T. Sikora, and BS Manjunath. *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley and Sons, Inc. New York, NY, USA, 2002.
- [114] A. Schmidt. Ontology-based user context management: the challenges of dynamics and imperfection. In *on the Move to Meaningful Internet Systems 2006: Coopis, DOA, GADA, and ODBASE*, pages 995–1011, 2006.
- [115] A. Schmidt and M. Schneider. Adaptive reading assistance for dyslexic students: Closing the loop. *Proc. of ABIS 07, Halle, Germany*, 2007.
- [116] M.S. Schmidt-Schauß and G. Smolka. Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26, 1991.
- [117] H. Schneiderman and T. Kanade. Object detection using the statistics of parts. *International Journal of Computer Vision*, 56(3):151–177, 2004.
- [118] JM Shapiro, D.S.R. Center, and NJ Princeton. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, 1993.
- [119] Sonant Software. Sonant software.
- [120] SPERO. Spero: A teleinformatics system for continuous collection, processing, diffusion of material for teacher training in special education.
- [121] SPERO statistics. Spero statistics: (<http://www.image.ntua.gr/statsspero>).

- [122] G. Stoilos, N. Simou, G. Stamou, and S. Kollias. Uncertainty and the semantic web. *IEEE INTELLIGENT SYSTEMS*, pages 84–87, 2006.
- [123] G. Stoilos, G. Stamou, and J.Z. Pan. Reasoning with the fuzzy description logic f-*SHIN*. Technical report, 2006.
- [124] Giorgos Stoilos. *Imperfect Knowledge and the Semantic Web: Expressive Fuzzy Description Logics*. PhD thesis, Department of Electrical & Computer Engineering, National & Technical University of Athens, Zographou, Greece, Feb 2008.
- [125] Giorgos Stoilos, Giorgos Stamou, Vassilis Tzouvaras, Jeff Z. Pan, and Ian Horrocks. Reasoning with very expressive fuzzy description logics. *Journal of Artificial Intelligence Research*, 30(5):273–320, 2007.
- [126] U. Straccia. Reasoning within fuzzy description logics. *Journal of Artificial Intelligence Research*, 14:137–166, 2001.
- [127] K.K. Sung and T. Poggio. Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):39–51, 1998.
- [128] M. D. Swanson, B. Zhu, B. Chau, and A. H. Tewfik. Object-based transparent video watermarking. In *Proceedings of the IEEE Workshop on Multimedia Signal Processing*, June 1997.
- [129] Cisco Systems. Blueprint for enterprise e-learning. White paper, June 2002.
- [130] Kurzweil Educational Systems. Kurzweil educational systems.
- [131] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1):116–132, 1985.
- [132] Dionysiou tou ek Phourna. *Hermeneia tis zografikis technis (Interpretation of the Byzantine art)*. B. Kirschbaum, 1909.
- [133] P. Tzouveli, Ph. Mylonas, and S. Kollias. An intelligent e-learning system based on learner profiling and learning resources adaptation. *Computers & Education*, 51(1):224–238, August 2008.
- [134] P. Tzouveli, A. Schmidt, M.Schneider, A. Symvonis, and S. Kollias. Adaptive reading assistance for the inclusion of students with dyslexia: The agent-dysl approach. In *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008)*, July 2008.
- [135] Paraskevi Tzouveli, Nikos Simou, Giorgios Stamou, and Stefanos Kollias. Semantic classification of byzantine icons. *IEEE Intelligent Systems*, 24(2):35–43, Mar./Apr. 2009.
- [136] R.G. Van Schyndel, A.Z. Tirkel, and C.F Osborne. A digital watermark. In *International Conference on Image Processing*, volume 2, pages 86–90, 1994.

- [137] P. Viola and M. Jones. Fast and robust classification using asymmetric adaboost and a detector cascade. *Advances in neural information processing systems*, 2:1311–1318, 2002.
- [138] P. Viola and M.J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
- [139] W3C. Web accesibility initiative.
- [140] E.M. Wadlington and P.L Wadlington. What educators really believe about dyslexia. *Reading Improvement*, 42(1):16–34, 2005.
- [141] M. Wallace. *Intelligent knowledge-based systems in uncertain environments*. PhD thesis, School of Electrical and Computer Engineering, National Technical University of Athens, 20050,0.
- [142] H. Wang and S.F Chang. A highly efficient system for automatic face region detection inmpeg video. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(4):615–628, 1997.
- [143] R. Wolfgang and E. Delp. A watermark for digital image. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages 211–214, September 1996.
- [144] J. Wood. Invariant pattern recognition: a review. *Pattern Recognition*, 29(1):1–17, 1996.
- [145] S. Wrobel, D. Wettschereck, E. Sommer, and W. Emde. Extensibility in data mining systems. *Arbeitspapiere-Gmd*, pages 177–192, 1996.
- [146] J. Wu, SC Brubaker, MD Mullin, and JM Rehg. Fast asymmetric learning for cascade face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):369–382, 2008.
- [147] M. Wu and B. Liu. Watermarking for image authentication. In *Proceedings of IEEE International Conference on Image Processing*, 1998.
- [148] Xiong Z. Wu X., Zhu W. and Zhang Y.-Q. Object-based multiresolution watermarking of images and video. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, May 2000.
- [149] M.H. Yang, D.J. Kriegman, and N. Ahuja. Detecting faces in images: A survey. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 24(1):34–58, January 2002.
- [150] M.H. Yang, DJ Kriegman, and N. Ahuja. Detecting faces in images: A survey. *IEEE Transactions on Pattern analysis and Machine intelligence*, 24(1):34–58, 2002.
- [151] L.A Zadeh. Fuzzy sets, information and control. *Vol*, 8(3):338–353, 1965.
- [152] W. Zeng and B. Liu. A statistical watermark detection technique without using originalimages for resolving rightful ownerships of digital images. *IEEE Transactions on Image Processing*, 8(11):1534–1548, 1999.

- [153] Z.Q. Zhang. Floatboost learning and statistical face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1112–1123, 2004.
- [154] Z.H. Zhou and X. Geng. Projection functions for eye detection. *Pattern Recognition*, 37(5):1049–1056, 2004.
- [155] W. Zhu, Z. Xiong, and Y.-Q. Zhang. Multiresolution watermarking for images and video. *IEEE Transactions on Circuits and Systems for Video Technology*, 9:545–550, June 1999.
- [156] X. Zhu and X. Wu. Class noise vs. attribute noise: A quantitative study. *Artificial Intelligence Review*, 22(3):177–210, 2004.

□

Παράρτημα Α

Σώμα ορολογίας για πρόσωπο Βυζαντινής μορφής

Classes

AboveEarsHair

AboveEarsHair \sqsubseteq Hair
AboveEarsHair $\sqsubseteq \neg$ Bald
AboveEarsHair $\sqsubseteq \neg$ BelowEarsHair

AcidBeard

AcidBeard \sqsubseteq Beard
AcidBeard $\sqsubseteq \neg$ CircleLikedBeard
AcidBeard $\sqsubseteq \neg$ CircleBeard
AcidBeard $\sqsubseteq \neg$ NoBeard
AcidBeard $\sqsubseteq \neg$ WideBeard
AcidBeard $\sqsubseteq \neg$ FluffyBeard

AndrewFace

AndrewFace \equiv OldManFace $\sqcap \exists$ hasComponent CurlyBeard \sqcap
 \exists hasComponent (BelowEarsHair \sqcap BlackHair)
AndrewFace \sqsubseteq OldManFace

AntoniosFace

AntoniosFace \equiv OldManFace $\sqcap \exists$ hasComponent CurlyBeard \sqcap
 \exists hasComponent (BelowEarsHair \sqcap CurlyHair \sqcap StraightHair
 \sqcap WhiteHair)
AntoniosFace \sqsubseteq OldManFace

Bald

Bald $\equiv \exists$ hasSegment (HighBeforeFleshSegment \sqcap VeryLowDarkSegment
 \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment \sqcap HighBeforeFleshSeg-
ment \sqcup MidBeforeFleshSegment)
Bald \sqsubseteq Hair
Bald $\sqsubseteq \neg$ StraightHair
Bald $\sqsubseteq \neg$ GrayHair

Bald \sqsubseteq \neg HairyHair
Bald \sqsubseteq \neg ThickHairOnTheRight
Bald \sqsubseteq \neg SparseHair
Bald \sqsubseteq \neg ThinHairOnTheRight
Bald \sqsubseteq \neg WhiteHair
Bald \sqsubseteq \neg AboveEarsHair
Bald \sqsubseteq \neg BraidOnTheRightShoulder
Bald \sqsubseteq \neg CurlyHair
Bald \sqsubseteq \neg BraidOnTheLeftShoulder
Bald \sqsubseteq \neg ThickHairOnTheLeft
Bald \sqsubseteq \neg ThinHairOnTheLeft
Bald \sqsubseteq \neg BlackHair
Bald \sqsubseteq \neg BelowEarsHair
Bald \sqsubseteq \neg BraidOnBothShoulders
Bald \sqsubseteq \neg WildHair

BassiliosFace

BassiliosFace \equiv VeryOldManFace \sqcap \exists hasComponent (AboveEarsHair \sqcap BlackHair
 \sqcap StraightHair) \sqcap \exists hasComponent (LongBeard \sqcap StraightBeard)
BassiliosFace \sqsubseteq MansFace

Beard

Beard \sqsubseteq AcidBeard \sqcup CircleBeard \sqcup CircleLikedBeard \sqcup FluffyBeard
 \sqcup NoBeard \sqcup WideBeard
Beard \sqsubseteq FaceComponent
Beard \sqsubseteq LongBeard \sqcup NoBeard \sqcup NotVeryLongBeard \sqcup ShortBeard
 \sqcup UpToKneesLongBeard \sqcup UpToWaistBeard
Beard \sqsubseteq BlackBeard \sqcup GrayBeard \sqcup NoBeard \sqcup WhiteBeard
Beard \sqsubseteq CurlyBeard \sqcup NoBeard \sqcup StraightBeard \sqcup WildBeard
Beard \sqsubseteq HairyBeard \sqcup NoBeard \sqcup SparseBeard
Beard \sqsubseteq HairyChin \sqcup NoBeard \sqcup SparseBeardChin
Beard \sqsubseteq \neg Hair
Beard \sqsubseteq \neg Cheek
Beard \sqsubseteq \neg Forehead
Beard \sqsubseteq \neg Moustache

BeforeFleshSegment

BeforeFleshSegment \sqsubseteq Segment

BelowEarsHair

BelowEarsHair \sqsubseteq Hair
BelowEarsHair \sqsubseteq \neg Bald
BelowEarsHair \sqsubseteq \neg AboveEarsHair

BlackBeard

BlackBeard \equiv HighDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
 \sqcap LowBeforeFleshSegment \sqcup MidBeforeFleshSegment

BlackBeard \sqsubseteq Beard

BlackBeard \sqsubseteq \neg NoBeard

BlackBeard \sqsubseteq \neg WhiteBeard

BlackHair

BlackHair \equiv HighDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
 \sqcap LowBeforeFleshSegment \sqcup MidBeforeFleshSegment

BlackHair \sqsubseteq Hair

BlackHair \sqsubseteq \neg Bald

BlackHair \sqsubseteq \neg WhiteHair

BlackMoustache

BlackMoustache \equiv HighDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
Segment

\sqcap LowBeforeFleshSegment \sqcup MidBeforeFleshSegment

BlackMoustache \sqsubseteq Moustache

BlackMoustache \sqsubseteq \neg NoMoustache

BlackMoustache \sqsubseteq \neg WhiteMoustache

BraidOnBothShoulders

BraidOnBothShoulders \equiv BraidOnTheLeftShoulder \sqcap BraidOnTheRightShoulder

BraidOnBothShoulders \sqsubseteq Hair

BraidOnBothShoulders \sqsubseteq \neg Bald

BraidOnTheLeftShoulder

BraidOnTheLeftShoulder \equiv BelowEarsHair \sqcap ThickHairOnTheLeft

BraidOnTheLeftShoulder \sqsubseteq Hair

BraidOnTheLeftShoulder \sqsubseteq \neg Bald

BraidOnTheRightShoulder

BraidOnTheRightShoulder \equiv BelowEarsHair \sqcap ThickHairOnTheRight

BraidOnTheRightShoulder \sqsubseteq Hair

BraidOnTheRightShoulder \sqsubseteq \neg Bald

Cheek

Cheek \sqsubseteq OldCheek \sqcup YoungCheek

Cheek \sqsubseteq FaceComponent

Cheek \sqsubseteq \neg Beard

Cheek \sqsubseteq \neg Hair

Cheek \sqsubseteq \neg Moustache

Cheek \sqsubseteq \neg Forehead

CircleBeard

CircleBeard \sqsubseteq Beard

CircleBeard \sqsubseteq \neg AcidBeard

CircleBeard \sqsubseteq \neg NoBeard

CircleBeard \sqsubseteq \neg WideBeard

CircleBeard \sqsubseteq \neg FluffyBeard

CircleLikedBeard

CircleLikedBeard \sqsubseteq Beard

CircleLikedBeard \sqsubseteq \neg AcidBeard

CircleLikedBeard \sqsubseteq \neg FluffyBeard

CircleLikedBeard \sqsubseteq \neg WideBeard

CircleLikedBeard \sqsubseteq \neg NoBeard

CurlyBeard

CurlyBeard \sqsubseteq Beard

CurlyBeard \sqsubseteq \neg NoBeard

CurlyBeard \sqsubseteq \neg StraightBeard

CurlyBeard \sqsubseteq \neg WildBeard

CurlyHair

CurlyHair \sqsubseteq Hair

CurlyHair \sqsubseteq \neg StraightHair

CurlyHair \sqsubseteq \neg Bald

CurlyHair \sqsubseteq \neg WildHair

CurlyMoustache

CurlyMoustache \sqsubseteq Moustache

CurlyMoustache \sqsubseteq \neg StraightMoustache

CurlyMoustache \sqsubseteq \neg NoMoustache

DarkSegment

DarkSegment \sqsubseteq Segment

DimitriosFace

DimitriosFace \equiv \exists hasComponent NoBeard \sqcap \exists hasComponent YoungCheek

\sqcap \exists hasComponent YoungForehead \sqcap

\exists hasComponent (BelowEarsHair \sqcap BlackHair \sqcap StraightHair)

DimitriosFace \sqsubseteq MansFace

EleniFace

EleniFace $\equiv \exists$ hasComponent NoBeard $\sqcap \exists$ hasComponent OldCheek
 $\sqcap \exists$ hasComponent YoungForehead $\sqcap \exists$ hasComponent (BelowEarsHair
 \sqcap BlackHair \sqcap StraightHair)
EleniFace \sqsubseteq WomansFace

EuthimiosFace

EuthimiosFace \equiv VeryOldManFace $\sqcap \exists$ hasComponent Bald \sqcap
 \exists hasComponent (LongBeard \sqcap StraightBeard)
EuthimiosFace \sqsubseteq MansFace

Face

Face $\equiv \forall$ hasComponent FaceComponent
Face \sqsubseteq Thing

FaceComponent

FaceComponent \sqsubseteq Beard \sqcup Cheek \sqcup Forehead \sqcup Hair \sqcup Moustache
FaceComponent \sqsubseteq Segment

Figure

Figure $\equiv \exists$ hasSegment Face
Figure \sqsubseteq Thing

FleshSegment

FleshSegment \sqsubseteq Segment

FluffyBeard

FluffyBeard \sqsubseteq Beard
FluffyBeard $\sqsubseteq \neg$ WideBeard
FluffyBeard $\sqsubseteq \neg$ NoBeard
FluffyBeard $\sqsubseteq \neg$ CircleLikedBeard
FluffyBeard $\sqsubseteq \neg$ CircleBeard
FluffyBeard $\sqsubseteq \neg$ AcidBeard

Forehead

Forehead \sqsubseteq OldForehead \sqcup YoungForehead
Forehead \sqsubseteq FaceComponent
Forehead $\sqsubseteq \neg$ Hair
Forehead $\sqsubseteq \neg$ Beard
Forehead $\sqsubseteq \neg$ Moustache
Forehead $\sqsubseteq \neg$ Cheek

GeorgiosFace

GeorgiosFace $\equiv \exists$ hasComponent NoBeard $\sqcap \exists$ hasComponent YoungCheek
 $\sqcap \exists$ hasComponent YoungForehead $\sqcap \exists$ hasComponent (BelowEarsHair
 \sqcap BlackHair \sqcap CurlyHair)
GeorgiosFace \sqsubseteq MansFace

GrayBeard

GrayBeard \equiv MidDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
 \sqcap HighBeforeFleshSegment \sqcup MidBeforeFleshSegment
GrayBeard \sqsubseteq Beard
GrayBeard $\sqsubseteq \neg$ NoBeard

GrayHair

GrayHair \equiv MidDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
 \sqcap HighBeforeFleshSegment \sqcup MidBeforeFleshSegment
GrayHair \sqsubseteq Hair
GrayHair $\sqsubseteq \neg$ Bald

GrayMoustache

GrayMoustache \equiv MidDarkSegment \sqcap VeryLowFleshSegment \sqcap VeryLowLightSegment
 \sqcap HighBeforeFleshSegment \sqcup MidBeforeFleshSegment
GrayMoustache \sqsubseteq Moustache
GrayMoustache $\sqsubseteq \neg$ NoMoustache

Hair

Hair \sqsubseteq Bald \sqcup ThickHairOnTheRight \sqcup ThinHairOnTheRight
Hair \sqsubseteq Bald \sqcup ThickHairOnTheLeft \sqcup ThinHairOnTheLeft
Hair \sqsubseteq AboveEarsHair \sqcup Bald \sqcup BelowEarsHair
Hair \sqsubseteq Bald \sqcup BlackHair \sqcup GrayHair \sqcup WhiteHair
Hair \sqsubseteq Bald \sqcup HairyHair \sqcup SparseHair
Hair \sqsubseteq FaceComponent
Hair $\sqsubseteq \neg$ Beard
Hair $\sqsubseteq \neg$ Forehead
Hair $\sqsubseteq \neg$ Cheek
Hair $\sqsubseteq \neg$ Moustache

HairyBeard

HairyBeard \sqsubseteq Beard
HairyBeard $\sqsubseteq \neg$ SparseBeard
HairyBeard $\sqsubseteq \neg$ NoBeard

HairyChin

HairyChin \sqsubseteq Beard

HairyChin $\sqsubseteq \neg$ SparseBeardChin

HairyChin $\sqsubseteq \neg$ NoBeard

HairyHair

HairyHair \sqsubseteq Hair

HairyHair $\sqsubseteq \neg$ SparseHair

HairyHair $\sqsubseteq \neg$ Bald

HighBeforeFleshSegment

HighBeforeFleshSegment \sqsubseteq BeforeFleshSegment

HighDarkSegment

HighDarkSegment \sqsubseteq DarkSegment

HighFleshSegment

HighFleshSegment \sqsubseteq FleshSegment

HighLightSegment

HighLightSegment \sqsubseteq LightSegment

Image

Image $\equiv \forall$ hasPart Segment

Image \sqsubseteq Thing

IoannisTheologosFace

IoannisTheologosFace \equiv VeryOldManFace $\sqcap \exists$ hasComponent (AboveEarsHair
 \sqcap StraightHair) $\sqcap \exists$ hasComponent (NotVeryLongBeard
 \sqcap WildBeard)

IoannisTheologosFace \sqsubseteq MansFace

JesusFace

JesusFace \equiv YoungMansFace $\sqcap \exists$ hasComponent (BraidOnTheLeftShoulder
 \sqcap HairyHair \sqcap StraightHair) $\sqcap \exists$ hasComponent (CircleLikedBeard
 \sqcap StraightBeard)

JesusFace \sqsubseteq MansFace

KaterinaFace

KaterinaFace \equiv VeryYoungFace \sqcap \exists hasComponent (StraightHair \sqcap BraidOnTheLeftShoulder
 \sqcup BraidOnTheRightShoulder)
KaterinaFace \sqsubseteq WomansFace

KonstantinosFace

KonstantinosFace \equiv \exists hasComponent OldCheek \sqcap \exists hasComponent OldForehead
 \sqcap \exists hasComponent (AboveEarsHair \sqcap BlackHair) \sqcap
 \exists hasComponent (AcidBeard \sqcap HairyChin \sqcap ShortBeard \sqcap StraightBeard)
KonstantinosFace \sqsubseteq MansFace

LightSegment

LightSegment \sqsubseteq Segment

LongBeard

LongBeard \sqsubseteq Beard
LongBeard \sqsubseteq \neg UpToWaistBeard
LongBeard \sqsubseteq \neg ShortBeard
LongBeard \sqsubseteq \neg UpToKneesLongBeard
LongBeard \sqsubseteq \neg NoBeard

LowBeforeFleshSegment

LowBeforeFleshSegment \sqsubseteq BeforeFleshSegment

LowDarkSegment

LowDarkSegment \sqsubseteq DarkSegment

LowFleshSegment

LowFleshSegment \sqsubseteq FleshSegment

LowLightSegment

LowLightSegment \sqsubseteq LightSegment

LukeFace

LukeFace \equiv \exists hasComponent OldCheek \sqcap \exists hasComponent OldForehead \sqcap
 \exists hasComponent (AboveEarsHair \sqcap GrayHair \sqcap StraightHair)
 \sqcap \exists hasComponent (AcidBeard \sqcap ShortBeard \sqcap SparseBeard \sqcap SparseBeard-
Chin \sqcap StraightBeard)
LukeFace \sqsubseteq MansFace

MansFace

MansFace \sqsubseteq Face

MatureManFace

MatureManFace $\equiv \exists$ hasComponent BlackHair $\sqcap \exists$ hasComponent OldCheek \sqcap
 \exists hasComponent OldForehead $\sqcap \exists$ hasComponent (HairyChin \sqcap ShortBeard)
MatureManFace \sqsubseteq MansFace

MidBeforeFleshSegment

MidBeforeFleshSegment \sqsubseteq BeforeFleshSegment

MidDarkSegment

MidDarkSegment \sqsubseteq DarkSegment

MidFleshSegment

MidFleshSegment \sqsubseteq FleshSegment

MidLightSegment

MidLightSegment \sqsubseteq LightSegment

MinasFace

MinasFace $\equiv \exists$ hasComponent OldCheek $\sqcap \exists$ hasComponent OldForehead \sqcap
 \exists hasComponent (BelowEarsHair \sqcap BlackHair \sqcap CurlyHair)
 $\sqcap \exists$ hasComponent (CircleBeard \sqcap HairyChin \sqcap ShortBeard)
MinasFace \sqsubseteq MansFace

Moustache

Moustache \sqsubseteq FaceComponent
Moustache \sqsubseteq CurlyMoustache \sqcup NoMoustache \sqcup StraightMoustache
Moustache \sqsubseteq BlackMoustache \sqcup NoMoustache \sqcup WhiteMoustache
Moustache $\sqsubseteq \neg$ Hair
Moustache $\sqsubseteq \neg$ Cheek
Moustache $\sqsubseteq \neg$ Forehead
Moustache $\sqsubseteq \neg$ Beard

NikolaosFace

NikolaosFace $\equiv \exists$ hasComponent OldCheek $\sqcap \exists$ hasComponent OldForehead \sqcap
 \exists hasComponent (AboveEarsHair \sqcap BlackHair \sqcap StraightHair)
 $\sqcap \exists$ hasComponent (CircleBeard \sqcap HairyChin \sqcap ShortBeard)
NikolaosFace \sqsubseteq MansFace

NoBeard

NoBeard \sqsubseteq Beard

- NoBeard $\sqsubseteq \neg$ GrayBeard
- NoBeard $\sqsubseteq \neg$ BlackBeard
- NoBeard $\sqsubseteq \neg$ CurlyBeard
- NoBeard $\sqsubseteq \neg$ SparseBeardChin
- NoBeard $\sqsubseteq \neg$ SparseBeard
- NoBeard $\sqsubseteq \neg$ CircleBeard
- NoBeard $\sqsubseteq \neg$ HairyChin
- NoBeard $\sqsubseteq \neg$ UpToWaistBeard
- NoBeard $\sqsubseteq \neg$ UpToKneesLongBeard
- NoBeard $\sqsubseteq \neg$ StraightBeard
- NoBeard $\sqsubseteq \neg$ WideBeard
- NoBeard $\sqsubseteq \neg$ HairyBeard
- NoBeard $\sqsubseteq \neg$ LongBeard
- NoBeard $\sqsubseteq \neg$ WhiteBeard
- NoBeard $\sqsubseteq \neg$ NotVeryLongBeard
- NoBeard $\sqsubseteq \neg$ FluffyBeard
- NoBeard $\sqsubseteq \neg$ ShortBeard
- NoBeard $\sqsubseteq \neg$ CircleLikedBeard
- NoBeard $\sqsubseteq \neg$ WildBeard
- NoBeard $\sqsubseteq \neg$ AcidBeard

NoMoustache

NoMoustache \equiv HighBeforeFleshSegment \sqcap VeryLowDarkSegment \sqcap VeryLowFleshSegment

\sqcap VeryLowLightSegment \sqcap HighBeforeFleshSegment \sqcup MidBeforeFleshSegment

- NoMoustache \sqsubseteq Moustache
- NoMoustache $\sqsubseteq \neg$ StraightMoustache
- NoMoustache $\sqsubseteq \neg$ WhiteMoustache
- NoMoustache $\sqsubseteq \neg$ CurlyMoustache
- NoMoustache $\sqsubseteq \neg$ BlackMoustache
- NoMoustache $\sqsubseteq \neg$ GrayMoustache

NotVeryLongBeard

NotVeryLongBeard \sqsubseteq Beard

- NotVeryLongBeard $\sqsubseteq \neg$ NoBeard
- NotVeryLongBeard $\sqsubseteq \neg$ UpToWaistBeard
- NotVeryLongBeard $\sqsubseteq \neg$ UpToKneesLongBeard

OldCheek

OldCheek \equiv HighDarkSegment $\sqcap \exists$ isAboveOf (HighFleshSegment \sqcup HighLightSegment \sqcup MidLightSegment)

- OldCheek \sqsubseteq Cheek
- OldCheek $\sqsubseteq \neg$ YoungCheek

OldForehead

OldForehead \sqsubseteq Forehead
OldForehead $\sqsubseteq \neg$ YoungForehead

OldManFace

OldManFace $\equiv \exists$ hasComponent OldCheek $\sqcap \exists$ hasComponent OldForehead
 $\sqcap \exists$ hasComponent (HairyChin \sqcap NotVeryLongBeard)
OldManFace \sqsubseteq MansFace

PelagiaFace

PelagiaFace \sqsubseteq WomansFace

ProkopiosFace

ProkopiosFace \equiv VeryYoungFace $\sqcap \exists$ hasComponent (BelowEar-
sHair \sqcap StraightHair)
ProkopiosFace \sqsubseteq MansFace

Segment

Segment \sqsubseteq Thing

ShortBeard

ShortBeard \sqsubseteq Beard
ShortBeard $\sqsubseteq \neg$ UpToKneesLongBeard
ShortBeard $\sqsubseteq \neg$ NoBeard
ShortBeard $\sqsubseteq \neg$ UpToWaistBeard
ShortBeard $\sqsubseteq \neg$ LongBeard

SimonFace

SimonFace $\equiv \exists$ hasComponent OldCheek $\sqcap \exists$ hasComponent OldForehead
 $\sqcap \exists$ hasComponent (AboveEarsHair \sqcap BlackHair \sqcap StraightHair)
 $\sqcap \exists$ hasComponent (CircleBeard \sqcap HairyChin \sqcap ShortBeard)
SimonFace \sqsubseteq MansFace

SparseBeard

SparseBeard \sqsubseteq Beard
SparseBeard $\sqsubseteq \neg$ HairyBeard
SparseBeard $\sqsubseteq \neg$ NoBeard

SparseBeardChin

SparseBeardChin \sqsubseteq Beard
SparseBeardChin $\sqsubseteq \neg$ HairyChin
SparseBeardChin $\sqsubseteq \neg$ NoBeard

SparseHair

SparseHair \sqsubseteq Hair
SparseHair $\sqsubseteq \neg$ HairyHair
SparseHair $\sqsubseteq \neg$ Bald

StefanosFace

StefanosFace $\equiv \exists$ hasComponent NoBeard $\sqcap \exists$ hasComponent YoungForehead
 $\sqcap \exists$ hasComponent (AboveEarsHair \sqcap BlackHair
 \sqcap StraightHair)
StefanosFace \sqsubseteq MansFace

StraightBeard

StraightBeard \sqsubseteq Beard
StraightBeard $\sqsubseteq \neg$ CurlyBeard
StraightBeard $\sqsubseteq \neg$ NoBeard
StraightBeard $\sqsubseteq \neg$ WildBeard

StraightHair

StraightHair \sqsubseteq Hair
StraightHair $\sqsubseteq \neg$ CurlyHair
StraightHair $\sqsubseteq \neg$ Bald
StraightHair $\sqsubseteq \neg$ WildHair

StraightMoustache

StraightMoustache \sqsubseteq Moustache
StraightMoustache $\sqsubseteq \neg$ CurlyMoustache
StraightMoustache $\sqsubseteq \neg$ NoMoustache

TheodorosFace

TheodorosFace \equiv OldManFace $\sqcap \exists$ hasComponent CurlyBeard \sqcap
 \exists hasComponent (BelowEarsHair \sqcap BlackHair \sqcap CurlyHair)
TheodorosFace \sqsubseteq OldManFace

ThickHairOnTheLeft

ThickHairOnTheLeft \sqsubseteq Hair
ThickHairOnTheLeft $\sqsubseteq \neg$ Bald
ThickHairOnTheLeft $\sqsubseteq \neg$ ThinHairOnTheLeft

ThickHairOnTheRight

ThickHairOnTheRight \sqsubseteq Hair
ThickHairOnTheRight $\sqsubseteq \neg$ Bald
ThickHairOnTheRight $\sqsubseteq \neg$ ThinHairOnTheRight

ThinHairOnTheLeft

ThinHairOnTheLeft \sqsubseteq Hair
ThinHairOnTheLeft \sqsubseteq \neg Bald
ThinHairOnTheLeft \sqsubseteq \neg ThickHairOnTheLeft

ThinHairOnTheRight

ThinHairOnTheRight \sqsubseteq Hair
ThinHairOnTheRight \sqsubseteq \neg Bald
ThinHairOnTheRight \sqsubseteq \neg ThickHairOnTheRight

Thing

ThomasFace

ThomasFace \equiv \exists hasComponent (AboveEarsHair \sqcap StraightHair) \sqcap
 \exists hasComponent (NotVeryLongBeard \sqcap WideBeard)
ThomasFace \sqsubseteq MansFace

UpToKneesLongBeard

UpToKneesLongBeard \sqsubseteq Beard
UpToKneesLongBeard \sqsubseteq \neg ShortBeard
UpToKneesLongBeard \sqsubseteq \neg NoBeard
UpToKneesLongBeard \sqsubseteq \neg LongBeard
UpToKneesLongBeard \sqsubseteq \neg NotVeryLongBeard

UpToWaistBeard

UpToWaistBeard \sqsubseteq Beard
UpToWaistBeard \sqsubseteq \neg NotVeryLongBeard
UpToWaistBeard \sqsubseteq \neg ShortBeard
UpToWaistBeard \sqsubseteq \neg NoBeard
UpToWaistBeard \sqsubseteq \neg LongBeard

VeryLowBeforeFleshSegment

VeryLowBeforeFleshSegment \sqsubseteq BeforeFleshSegment

VeryLowDarkSegment

VeryLowDarkSegment \sqsubseteq DarkSegment

VeryLowFleshSegment

VeryLowFleshSegment \sqsubseteq FleshSegment

VeryLowLightSegment

VeryLowLightSegment \sqsubseteq LightSegment

VeryOldManFace

VeryOldManFace $\equiv \exists$ hasComponent OldCheek $\sqcap \exists$ hasComponent OldForehead
 $\sqcap \exists$ hasComponent WhiteHair \sqcap
 \exists hasComponent (HairyBeard \sqcap HairyChin)
VeryOldManFace \sqsubseteq MansFace

VeryYoungFace

VeryYoungFace $\equiv \exists$ hasComponent BlackHair $\sqcap \exists$ hasComponent NoBeard
 $\sqcap \exists$ hasComponent YoungForehead $\sqcap \exists$ hasPart YoungCheek
VeryYoungFace $\equiv \exists$ hasComponent BlackHair $\sqcap \exists$ hasComponent NoBeard
 $\sqcap \exists$ hasComponent YoungForehead
VeryYoungFace \sqsubseteq Face

VirginMaryFace

VirginMaryFace $\equiv \exists$ hasComponent NoBeard $\sqcap \exists$ hasComponent OldCheek
 $\sqcap \exists$ hasComponent OldForehead $\sqcap \exists$ hasComponent (BlackHair \sqcap StraightHair)
VirginMaryFace \sqsubseteq WomansFace

WhiteBeard

WhiteBeard \equiv HighFleshSegment \sqcap HighLightSegment \sqcap VeryLowBeforeFleshSegment
 \sqcap VeryLowDarkSegment
WhiteBeard \sqsubseteq Beard
WhiteBeard $\sqsubseteq \neg$ NoBeard
WhiteBeard $\sqsubseteq \neg$ BlackBeard

WhiteHair

WhiteHair \equiv HighFleshSegment \sqcap HighLightSegment \sqcap VeryLowBeforeFleshSegment
 \sqcap VeryLowDarkSegment
WhiteHair \sqsubseteq Hair
WhiteHair $\sqsubseteq \neg$ BlackHair
WhiteHair $\sqsubseteq \neg$ Bald

WhiteMoustache

WhiteMoustache \equiv HighFleshSegment \sqcap HighLightSegment \sqcap VeryLowBeforeFleshSegment
 \sqcap VeryLowDarkSegment
WhiteMoustache \sqsubseteq Moustache
WhiteMoustache $\sqsubseteq \neg$ NoMoustache
WhiteMoustache $\sqsubseteq \neg$ BlackMoustache

WideBeard

WideBeard \sqsubseteq Beard

WideBeard \sqsubseteq \neg FluffyBeard

WideBeard \sqsubseteq \neg CircleLikedBeard

WideBeard \sqsubseteq \neg CircleBeard

WideBeard \sqsubseteq \neg AcidBeard

WideBeard \sqsubseteq \neg NoBeard

WildBeard

WildBeard \sqsubseteq Beard

WildBeard \sqsubseteq \neg CurlyBeard

WildBeard \sqsubseteq \neg NoBeard

WildBeard \sqsubseteq \neg StraightBeard

WildHair

WildHair \sqsubseteq Hair

WildHair \sqsubseteq \neg CurlyHair

WildHair \sqsubseteq \neg Bald

WildHair \sqsubseteq \neg StraightHair

WomansFace

WomansFace \sqsubseteq Face

YoungCheek

YoungCheek \equiv HighDarkSegment \sqcap \exists isAboveOf (HighBeforeFleshSegment
 \sqcup LowDarkSegment \sqcup MidBeforeFleshSegment)

YoungCheek \sqsubseteq Cheek

YoungCheek \sqsubseteq \neg OldCheek

YoungForehead

YoungForehead \sqsubseteq Forehead

YoungForehead \sqsubseteq \neg OldForehead

YoungMansFace

YoungMansFace \equiv \exists hasComponent BlackHair \sqcap \exists hasComponent YoungCheek
 \sqcap \exists hasComponent YoungForehead \sqcap \exists hasComponent (BlackBeard
 \sqcap ShortBeard \sqcap SparseBeard \sqcap SparseBeardChin)

YoungMansFace \sqsubseteq MansFace

Object properties

hasComponent

\sqsubseteq hasPart
 $\top \sqsubseteq \forall$ hasComponent FaceComponent

hasPart

hasSegment

\sqsubseteq hasPart

isAboveOf

isAboveOf \equiv isBelowOf⁻

isBelowOf

isAboveOf \equiv isBelowOf⁻

isLeftOf

isLeftOf \equiv isRightOf⁻

isRightOf

isLeftOf \equiv isRightOf⁻

Data properties

Individuals

org.coode.owl.latex.LatexWriter@2eb0c175

Παράρτημα Β

Ερωτηματολόγια αρχικής εξαγωγής προφίλ χρηστών



Ερωτήσεις για τη σχολική μονάδα

Περιγραφικά στοιχεία της σχολικής μονάδας

Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

1. Περιγραφή της σχολικής μονάδας		
Επωνυμία		
Διεύθυνση		
Πόλη, Χώρα		
Τηλέφωνο		
Τ.Κ.		
Φαξ		
Email		
2. Γενικές πληροφορίες για τη σχολική μονάδα		
	ΣΥΝΟΛΟ	ΑΜΕΑ
Νήπια		
Μαθητές στο δημοτικό		
Μαθητές στο γυμνάσιο		
Εκπαιδευτικοί		
Εκπαιδευτικοί Ειδικής Εκπαίδευσης		
Αριθμός μαθητών που καταρτίζονται στις τεχνολογίες της πληροφορικής		
Αριθμός Η/Υ στο σχολείο		
Αριθμός Η/Υ στο σχολείο που χρησιμοποιούνται από τους μαθητές στη μαθησιακή διαδικασία		

3. Η περιοχή λειτουργίας του σχολείου είναι:			
Αγροτική			
Σε μια μικρή πόλη (<15.000)			
Σε μια πόλη μέσου μεγέθους (<100.000)			
Σε μια μεγάλη πόλη (>100.000)			
4. Ειδικές εκπαιδευτικές ανάγκες (στο σύνολο του μαθητικού πληθυσμού)			
	Αγόρια	Κορίτσια	
Ήπιες ως μέτριες μαθησιακές δυσκολίες			
Σοβαρές / πολύπλοκες μαθησιακές δυσκολίες			
Κινητικά προβλήματα			
Αυτισμός/ δυσκολίες στην επικοινωνία			
Προβλήματα ακοής (κωφοί-βαρήκοοι)			
Προβλήματα όρασης (τυφλά-αμβλύωπα)			
Συναισθηματικές διαταραχές / διαταραχές συμπεριφοράς			
Πολιτισμικές και γλωσσικές δυσκολίες			
Άλλες δυσκολίες			
5. Έχει επιδοτηθεί το σχολείο σας με ενδο-υπηρεσιακή τα τελευταία τρία χρόνια, σε κάποια από τις ακόλουθες κατηγορίες;			
	Με πόρους από την Ευρωπαϊκή Ένωση	Με εθνικούς πόρους	Με άλλους πόρους
Κατάρτιση προσωπικού στις ΤΠΕ			
Κατάρτιση προσωπικού στις ΤΠΕ για ΑΜΕΑ			
ΤΠΕ υλικό (hardware) και λογισμικό			
ΤΠΕ υλικό (hardware) και λογισμικό για ΑΜΕΑ			
Κατάρτιση των νεοδιορισμένων εκπαιδευτικών για ΑΜΕΑ			



Ερωτήσεις για τον εκπαιδευτικό

Α. Παιδαγωγική χρήση των Τεχνολογιών Πληροφορικής

Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

A.1.Αλλαγές στις καθημερινές δραστηριότητες του διδακτικού προσωπικού που οφείλονται στη χρήση των Τεχνολογιών της Πληροφορικής και Επικοινωνίας A.1.1. Η ευρεία χρήση των ΤΠΕ ως "εργαλείων" στο σχολείο θα συντελέσει σε αλλαγές απόψεων σχετικά με την πρακτική στη διδασκαλία. Παρακαλούμε δηλώστε πού επικεντρώνεται αυτή η αλλαγή.		
	Πολύ συχνά	Σπάνια
Αλλαγές στη στάση του διδακτικού προσωπικού είναι απαραίτητες		
Θα πρέπει να σχεδιαστούν και να υλοποιηθούν νέες διδακτικές δραστηριότητες για τους μαθητές		
Δεν θα επηρεάσει άλλες διδακτικές δραστηριότητες		
Θα αυξήσει την ομαδική εργασία και την συνεργασία των εκπαιδευτικών		
Θα αποτελέσει προσωπική πρόκληση για το διδακτικό προσωπικό η απόκτηση νέων γνώσεων και δεξιοτήτων που θα απαιτούνται		
Σχόλια ή άλλες κατηγορίες		

Α.2.Αλλαγές στη διδακτική-μαθησιακή διαδικασία που οφείλονται στη χρήση των Τεχνολογιών της Πληροφορικής και Επικοινωνίας		
Α.2.1. Η ευρεία χρήση των ΤΠΕ ως "εργαλείων" στα σχολεία θα οδηγήσει τους εκπαιδευτικούς και μαθητές, στη σχολική τάξη, σε διαφορετικές λειτουργίες. Παρακαλούμε διευκρινίστε ποιες :		
	Πολύ συχνά	Σπάνια
Η μάθηση με τη χρήση των ΤΠΕ κινητοποιεί τους μαθητές		
Η μάθηση έχει μεγάλη διαδραστική ικανότητα αν σε αυτήν χρησιμοποιούνται οι ΤΠΕ		
Οι μαθητές θα έχουν μεγαλύτερη ικανότητα κατανόησης του διδακτικού υλικού		
Σχόλια ή άλλες κατηγορίες		
Α.3.Αλλαγές σχετικές με τη διαφορετικότητα των μαθητών και την συνεκπαίδευσή τους που επιτυγχάνονται με τη χρήση των τεχνολογιών.		
	Πολύ συχνά	Σπάνια
Οι τεχνολογίες πληροφορικής θα επιτρέψουν την προσαρμογή των μαθητών σε διαφορετικούς ρυθμούς εργασίας και την ένταξη των μαθητών με ειδικές ανάγκες		
Οι τεχνολογίες της πληροφορικής μπορούν να χρησιμοποιηθούν για ενίσχυση στη μάθηση εξοικονομώντας τον χρόνο των εκπαιδευτικών		
Οι τεχνολογίες της πληροφορικής θα επιτρέψουν τους περισσότερους ικανούς μαθητές να διευρύνουν τον ορίζοντα της μάθησής τους		
Οι τεχνολογίες της πληροφορικής μπορούν να χρησιμοποιηθούν για ενίσχυση της επικοινωνίας μεταξύ των μαθητών		
Η χρήση των τεχνολογιών της πληροφορικής δεν παρέχει σημαντικά πλεονεκτήματα στους μαθητές με ειδικές εκπαιδευτικές ανάγκες		
Σχόλια ή άλλες κατηγορίες		

Α.3.2. Ποια είναι η γνώμη σας για το Διαδίκτυο (Internet) ως εργαλείο υποστήριξης στο διδακτικό σας πρόγραμμα; Πιστεύετε ότι το Διαδίκτυο είναι:	
Πολύ χρήσιμο	
Χρήσιμο	
Δεν είναι χρήσιμο	

Α.3.3. Πιστεύετε ότι το Διαδίκτυο θα οδηγήσει σε σημαντικές αλλαγές τον τρόπο διδασκαλίας σας; Πιστεύετε ότι:	
Είναι ήδη η περίπτωση αυτή;	
Θα αποτελέσει πιθανόν την περίπτωση αυτή στα επόμενα 3 χρόνια;	
Θα αποτελέσει πιθανόν την περίπτωση αυτή σε περισσότερα από 3 χρόνια;	
Δεν θα είναι πιθανή ποτέ η περίπτωση αυτή;	



Ερωτήσεις για τον εκπαιδευτικό

B1. Επαγγελματική εξέλιξη

Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

B.1.1.Είναι η θέση σας :		
	Με πλήρες ωράριο	Με μερική απασχόληση
Μόνιμη		
Προσωρινή		

B.1.2.Είστε αποκλειστικά εκπαιδευτικός ειδικής αγωγής ή απλώς εργάζεστε με μαθητές με ειδικές εκπαιδευτικές ανάγκες (Να αφεθεί κενό αν δεν ανήκετε σε κάποια από τις δύο κατηγορίες)	
Σε ένα Ειδικό Σχολείο;	
Σε μια τάξη εκπαιδευτικής ένταξης στο γενικό σχολείο;	
Σε παροχή μαθησιακής υποστήριξης σε ένα γενικό σχολείο;	

B.1.3.Προσόντα/ κατάρτιση σε Ειδικές Εκπαιδευτικές Ανάγκες;	
Πτυχίο	
Βεβαίωση μετεκπαίδευσης /Δίπλωμα Εξειδίκευσης σε υποστηρικτική διδασκαλία (ή ισότιμο άλλο)	
Μεταπτυχιακό Δίπλωμα (Masters/Διδακτορικό)	
Άλλα πτυχία σε προγράμματα για Ειδικές Εκπαιδευτικές Ανάγκες (Σημειώστε διάρκεια σε χρόνια και μήνες)	

B.1.4.Διδακτική εμπειρία		
Κάτω από 5 χρόνια		
5 με 10 χρόνια		
11 με 20 χρόνια		
Πάνω από 20 χρόνια		
B.1.5.Η ομάδα ηλικίας σας		
Κάτω από 20		
21 με 30		
31 με 40		
41 με 50		
Πάνω από 50		
B.1.6.Γένος		
Άρρην		
Θήλυ		
B.1.7.Η ηλικία των μαθητών σας		
	ΣΥΝΟΛΟ	ΑΜΕΑ*
Νήπια		
Μαθητές στο δημοτικό		
Μαθητές στο γυμνάσιο		
B.1.8.Η περιοχή του σχολείου σας είναι:		
Αγροτική		
Σε μια μικρή πόλη (<15.000)		
Σε μια πόλη μέσου μεγέθους (<100.000)		
Σε μια μεγάλη πόλη (>100.000)		



Ερωτήσεις για τον εκπαιδευτικό

B2. Γνώση της Τεχνολογίας

Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

B.2.1.		
	Ναι	Όχι
Έχετε Η/Υ στο σπίτι;		
Έχετε σύνδεση με το Διαδίκτυο (Internet) στο σπίτι;		
B.2.2. Πόσο συχνά χρησιμοποιείτε τη σύνδεσή σας με το Διαδίκτυο στο σπίτι;		
Κάθε μέρα		
Μια φορά την εβδομάδα		
Σπάνια		
Ποτέ		
B.2.3. Επιδoteίται ο Η/Υ και/ή η σύνδεσή σας με το Διαδίκτυο :		
Από το Κράτος		
Από τη Βιομηχανία		
Από το σχολείο		
Από άλλους οργανισμούς		
Δεν επιδοτείται		
Δεν γνωρίζω		

B.2.4.Για ποιο από τα παρακάτω χρησιμοποιήσατε τον Η/Υ τουλάχιστον μια φορά στον μήνα που πέρασε ;	
Πρόσβαση στο Διαδίκτυο	<input type="checkbox"/>
Λήψη/αποστολή ηλεκτρονικών μηνυμάτων	<input type="checkbox"/>
Παιχνίδια	<input type="checkbox"/>
Επεξεργασία κειμένου	<input type="checkbox"/>
Εκπαιδευτικό λογισμικό	<input type="checkbox"/>
Άλλο, παρακαλώ διευκρινίστε:	
B.2.5.Ποια από τις παρακάτω εργασίες εκτελέσατε τουλάχιστον μια φορά, χωρίς καμία βοήθεια;	
Εγκατάσταση λογισμικού	<input type="checkbox"/>
Εγκατάσταση ενός εκτυπωτή	<input type="checkbox"/>
Δημιουργία αντιγράφων ασφαλείας (backup)	<input type="checkbox"/>
Εγκατάσταση εφαρμογών (e.g. MS Office)	<input type="checkbox"/>
Αντιμετώπιση προβλημάτων σε σχέση με το λογισμικό	<input type="checkbox"/>
Αντιμετώπιση προβλημάτων σε σχέση με τον εκτυπωτή ή τη σύνδεση με το Διαδίκτυο	<input type="checkbox"/>
Κανένα από τα παραπάνω	<input type="checkbox"/>
Άλλες, παρακαλώ διευκρινίστε:	



Ερωτήσεις για τον εκπαιδευτικό

B3. Χρήση των τεχνολογιών στη διδασκαλία

Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

B.3.1.Υπάρχουν Η/Υ στο περιβάλλον της εργασίας σας;		
Ναι		
Όχι		
B.3.2.Πόσο συχνά χρησιμοποιήσατε Η/Υ την τελευταία εβδομάδα στο σχολείο;		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Κάθε μέρα		
3-4 μέρες		
1-2 μέρες		
Καθόλου		
B.3.3.Έχετε πρόσβαση στο Διαδίκτυο ή σε εκπαιδευτικό λογισμικό στο εργασιακό σας περιβάλλον;		
	Ναι	Όχι
Στη γενική εκπαίδευση		
Σε ειδικές εκπαιδευτικές ανάγκες		
B.3.4.Πόσες ώρες την εβδομάδα, κατά μέσο όρο, χρησιμοποιείτε το Διαδίκτυο ή εκπαιδευτικό λογισμικό με τους μαθητές σας;		
Στη γενική εκπαίδευση		
Σε ειδικές εκπαιδευτικές ανάγκες		

B.3.5.Χρησιμοποιείτε το Διαδίκτυο για έρευνα και αναζήτηση πληροφοριών που αφορούν σε ανάγκες των μαθητών ΜΕΑ;		
	Ναι	Όχι
Στη γενική εκπαίδευση		
Σε ειδικές εκπαιδευτικές ανάγκες		

B.3.6.Χρησιμοποιείτε το Διαδίκτυο από το σχολείο για να αναζητήσετε πρόσθετες πηγές εκπαιδευτικού υλικού		
	Ναι	Όχι
Στη γενική εκπαίδευση		
Σε ειδικές εκπαιδευτικές ανάγκες		

B.3.7.Σε ποιο χώρο γενικά αναζητούν οι μαθητές σας πρόσβαση στο Διαδίκτυο για εκπαιδευτικό σκοπούς;		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Στην τάξη		
Στο εργαστήριο Πληροφορικής		
Στη σχολική βιβλιοθήκη		
Στο σπίτι		
Δεν υπάρχει πρόσβαση		
Άλλο		

B.3.8.Χρησιμοποιείτε το Διαδίκτυο για να συνδεθείτε με άλλα σχολεία;		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Ναι, με σχολεία στην περιοχή μου		
Ναι, με σχολεία στη χώρα μου		
Ναι, με σχολεία στην Ευρωπαϊκή Ένωση		
Ναι, με σχολεία εκτός της Ευρωπαϊκής Ένωσης		
Όχι		

B.3.9.Αποψη που αφορά στην ποσότητα του διαθέσιμου εκπαιδευτικού λογισμικού που χρησιμοποιείτε στην τάξη για τους μαθητές		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Αρκετά ανεπαρκής		
Ανεπαρκής		
Επαρκής		
Περισσότερο από επαρκής		

B.3.10.Αποψη που αφορά στην τεχνολογική ποιότητα του εκπαιδευτικού λογισμικού που διατίθεται στην τάξη σας		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Τελείως απαρχαιωμένο		
Επί το πλείστον απαρχαιωμένο		
Επί το πλείστον κατάλληλο		
Πολύ κατάλληλο ή καλύτερο		

B.3.11.Αποψη που αφορά στη χρηστικότητα, σύμφωνα με τους όρους της διδακτικής, του εκπαιδευτικού λογισμικού που διατίθεται στην τάξη		
	Στη γενική εκπαίδευση	Σε ειδικές εκπαιδευτικές ανάγκες
Τελείως άχρηστο		
Επί το πλείστον άχρηστο		
Επί το πλείστον χρήσιμο		
Περισσότερο από χρήσιμο		



Ερωτήσεις για τον εκπαιδευτικό

Γ. Πρόσθετες πληροφορίες

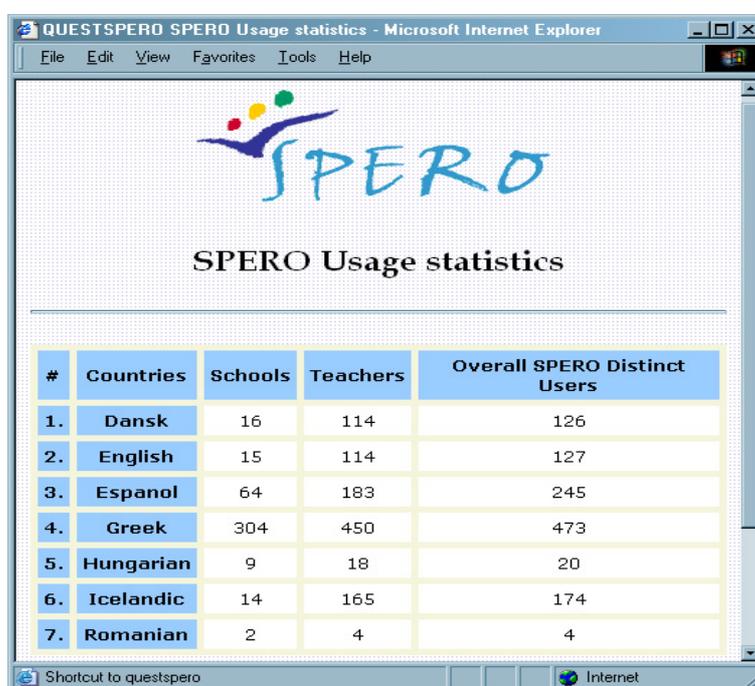
Παρακαλούμε να συμπληρώσετε τα ακόλουθα πεδία:

Εάν έχετε πρόσθετα σχόλια ή πληροφορίες που σχετίζονται με το περιεχόμενο των ερωτήσεων, παρακαλούμε χρησιμοποιήστε το παρακάτω διάστημα.

Παράρτημα Γ

Δείγματα της στατιστικής ανάλυσης των απαντήσεων των χρηστών

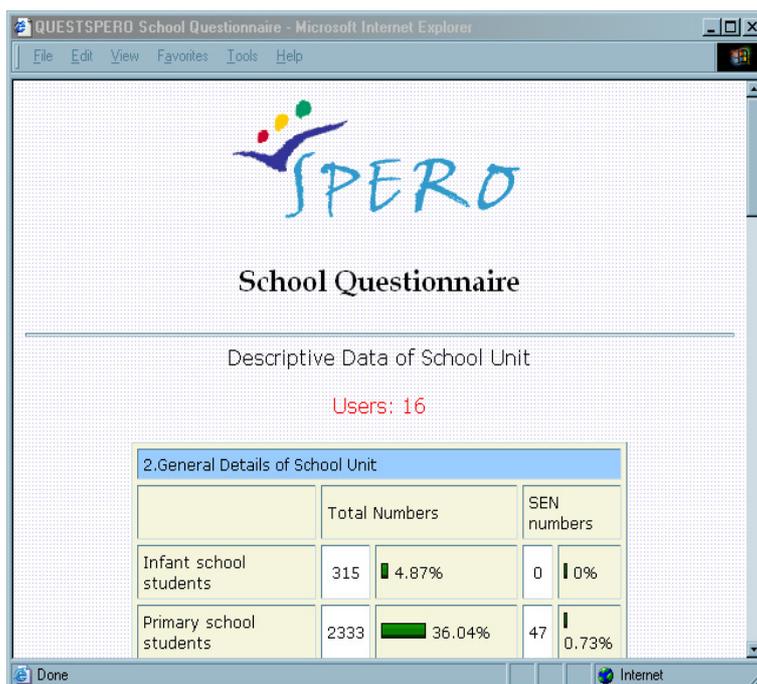
Η στατιστική ανάλυση της συνολικής συμμετοχής όλων των χωρών καθώς και η ανάλυση της συνολικής συμμετοχής των σχολείων και των χρηστών φαίνονται στα ακόλουθα σχήματα:



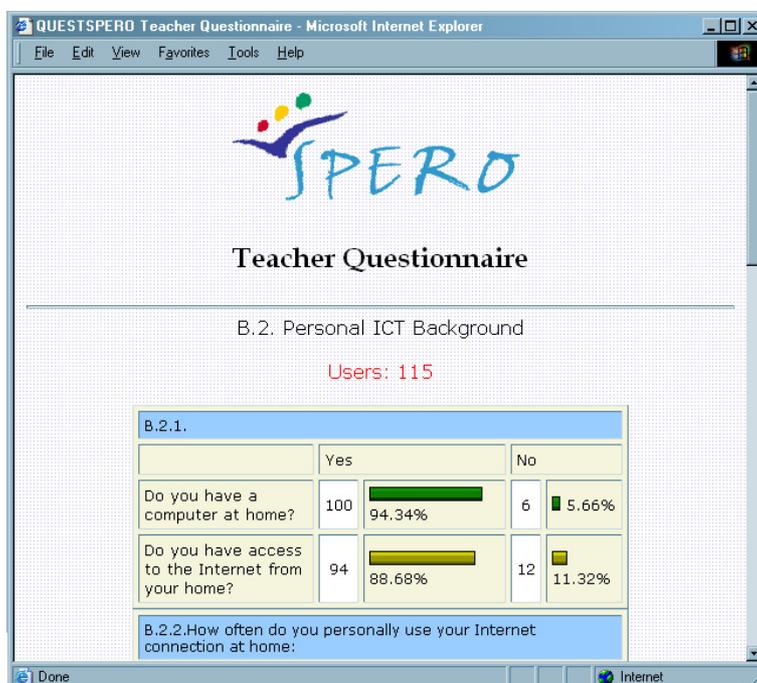
#	Countries	Schools	Teachers	Overall SPERO Distinct Users
1.	Dansk	16	114	126
2.	English	15	114	127
3.	Espanol	64	183	245
4.	Greek	304	450	473
5.	Hungarian	9	18	20
6.	Icelandic	14	165	174
7.	Romanian	2	4	4

Σχήμα 1: Στατιστική ανάλυση συνολικής συμμετοχής

Στη συνέχεια παρουσιάζονται οι στατιστικές αναλύσεις ως προς τις απαντήσεις των εκπαιδευομένων στα ερωτηματολόγια που τους αφορούν. Στις λεζάντες των σχημάτων αναφέρονται οι ερωτήσεις (Παράρτημα Β) που τέθηκαν στους εκπαιδευομένους και εμφανίζονται με ποσοστά οι απαντήσεις που αυτοί έδωσαν ανά χώρα προέλευσης.

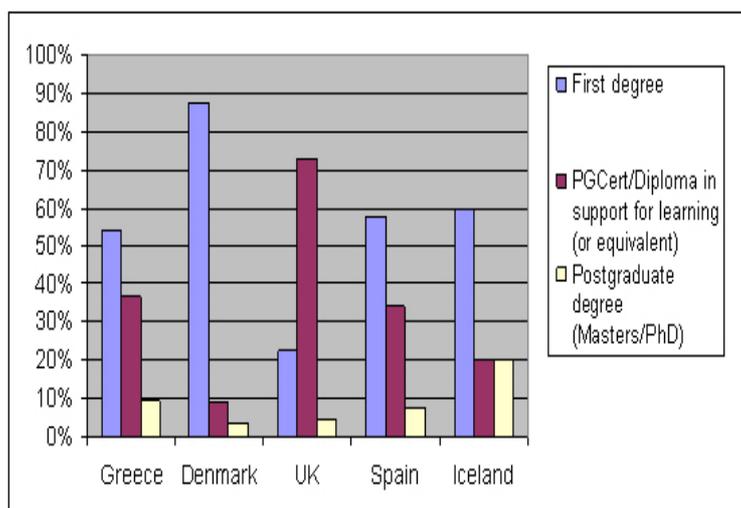


Σχήμα 2: Στατιστική ανάλυση συνολικής συμμετοχής σχολείων

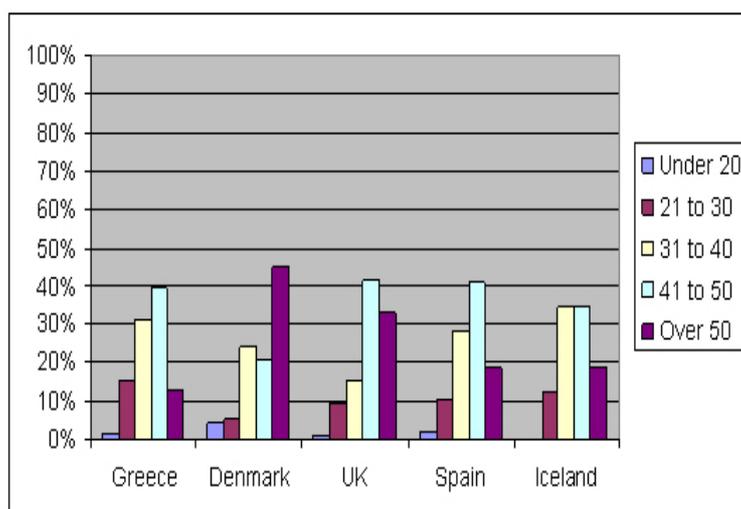


Σχήμα 3: Στατιστική ανάλυση συνολικής συμμετοχής χρηστών

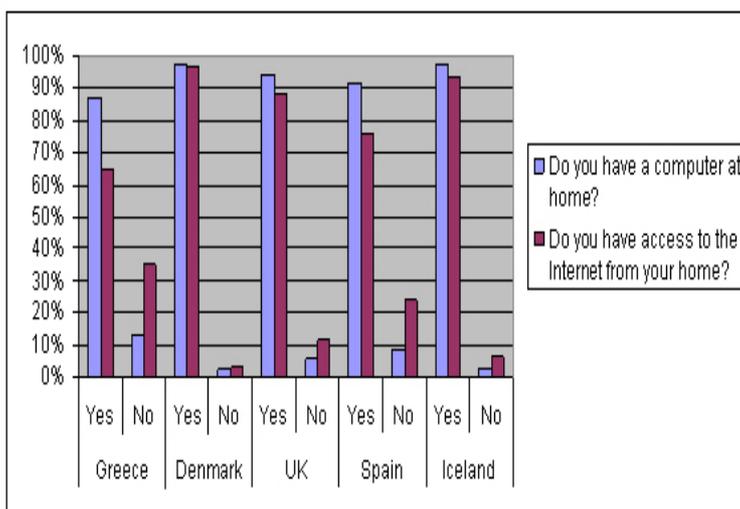
Σχήμα 4: Ερώτηση B.1.1.Είναι η θέση σας



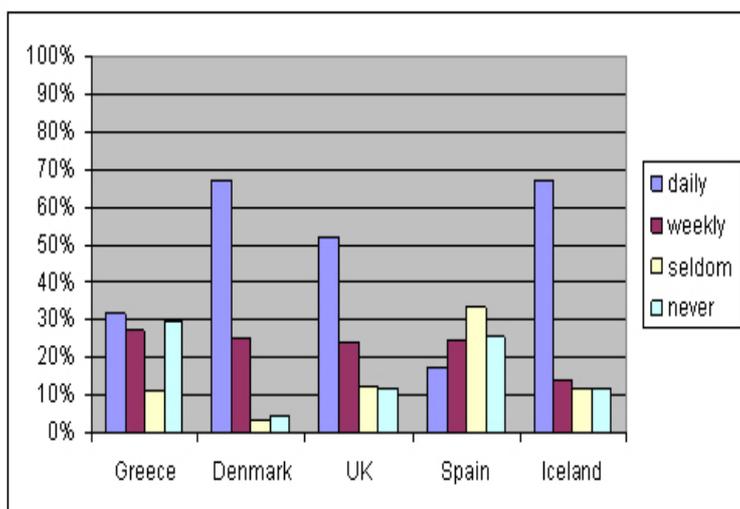
Σχήμα 5: Ερώτηση B.1.3. Προσόντα/ κατάρτιση σε Ειδικές Εκπαιδευτικές Ανάγκες



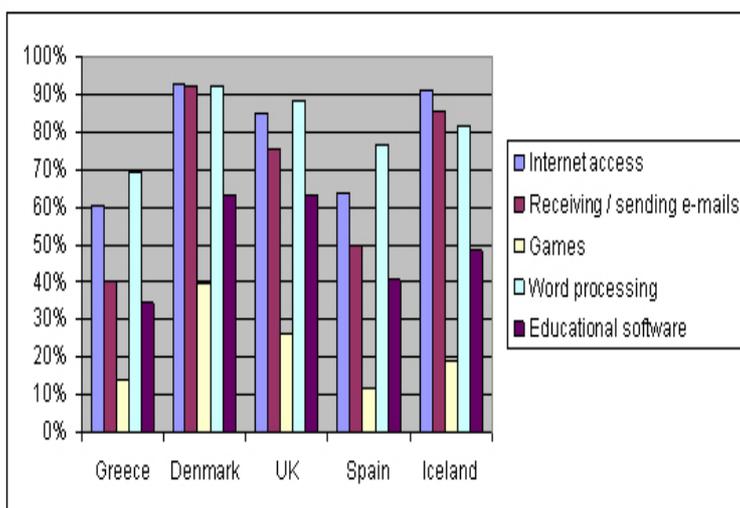
Σχήμα 6: Ερώτηση B.1.5. Η ομάδα ηλικίας σας



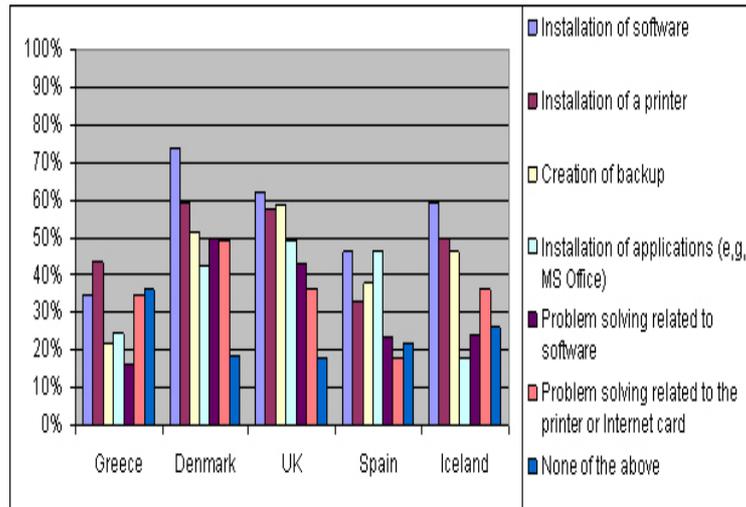
Σχήμα 7: Ερώτηση B.2.1. Έχετε Η/Υ στο σπίτι



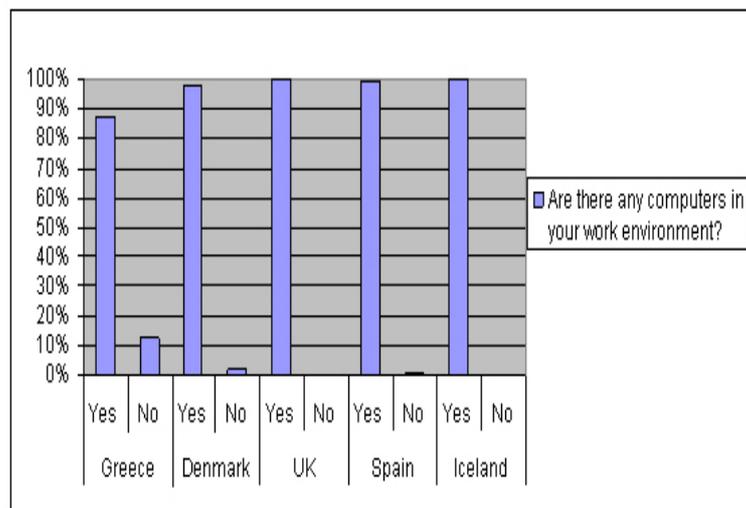
Σχήμα 8: Ερώτηση B.2.1. Έχετε σύνδεση με το Διαδίκτυο (Internet) στο σπίτι



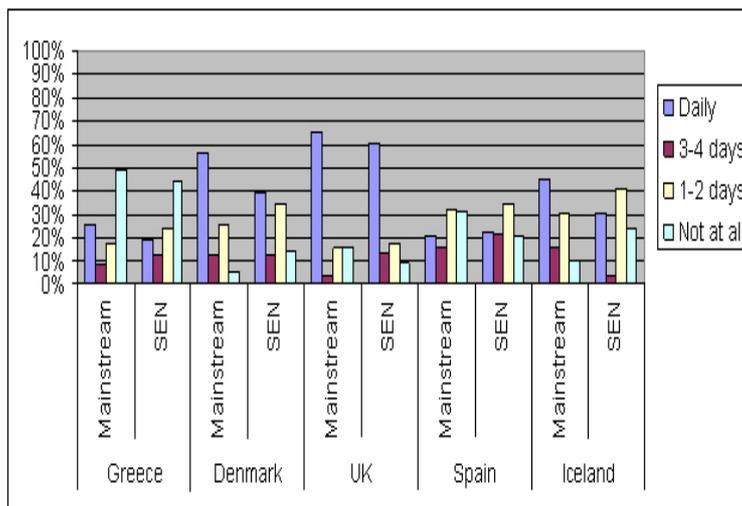
Σχήμα 9: Ερώτηση B.2.4. Για ποιο από τα παρακάτω χρησιμοποιήσατε τον Η/Υ τουλάχιστον μια φορά στον μήνα που πέρασε



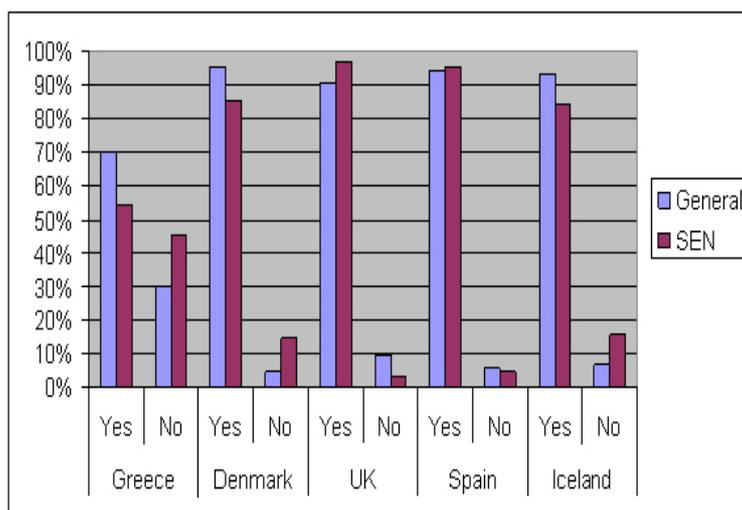
Σχήμα 10: Ερώτηση Β.2.5. Ποια από τις παρακάτω εργασίες εκτελέσατε τουλάχιστον μια φορά, χωρίς καμία βοήθεια



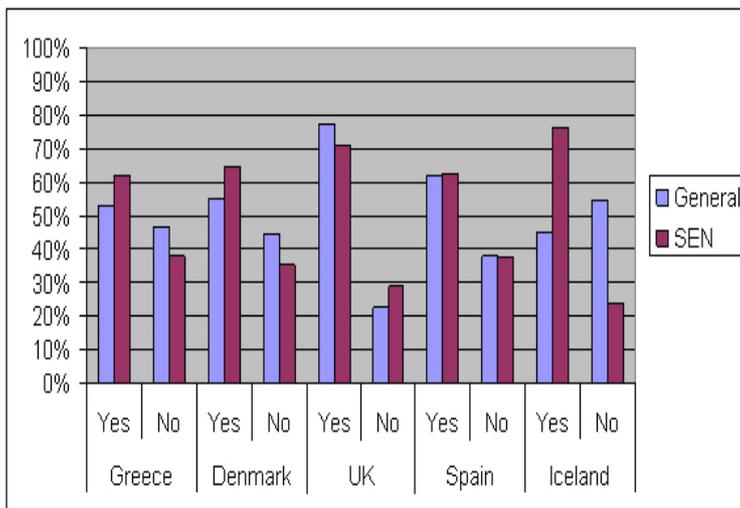
Σχήμα 11: Ερώτηση Β.3.1. Υπάρχουν Η/Υ στο περιβάλλον της εργασίας σας



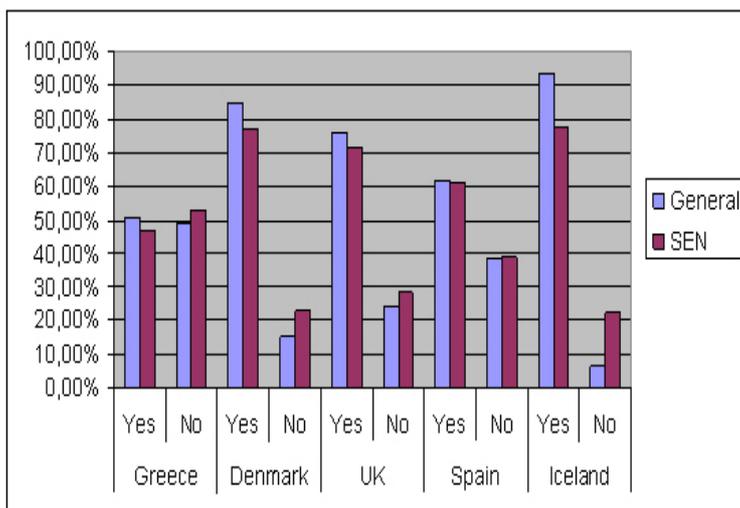
Σχήμα 12: Ερώτηση Β.3.2. Πόσο συχνά χρησιμοποιήσατε Η/Υ την τελευταία εβδομάδα στο σχολείο



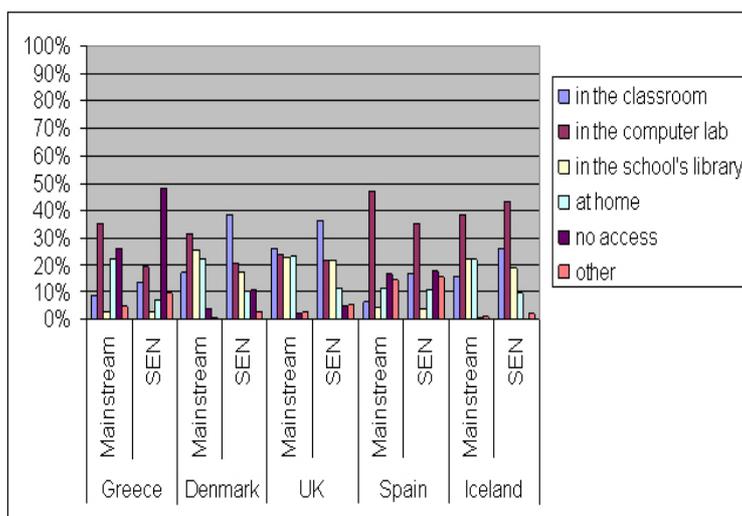
Σχήμα 13: Ερώτηση Β.3.3. Έχετε πρόσβαση στο Διαδίκτυο ή σε εκπαιδευτικό λογισμικό στο εργασιακό σας περιβάλλον



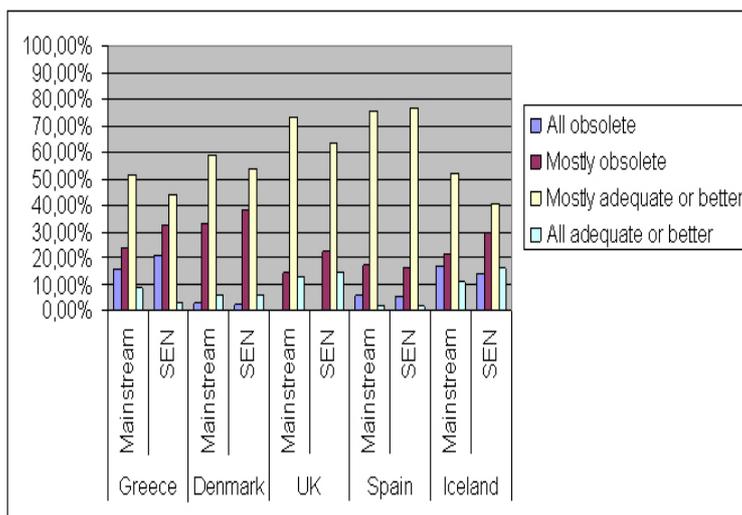
Σχήμα 14: Ερώτηση Β.3.5.Χρησιμοποιείτε το Διαδίκτυο για έρευνα και αναζήτηση πληροφοριών που αφορούν σε ανάγκες των μαθητών ΜΕΑ



Σχήμα 15: Ερώτηση Β.3.6.Χρησιμοποιείτε το Διαδίκτυο από το σχολείο για να αναζητήσετε πρόσθετες πηγές εκπαιδευτικού υλικού



Σχήμα 16: Ερώτηση Β.3.7. Σε ποιο χώρο γενικά αναζητούν οι μαθητές σας πρόσβαση στο Διαδίκτυο για εκπαιδευτικούς σκοπούς



Σχήμα 17: Ερώτηση Β.3.10. Αποψη που αφορά στην τεχνολογική ποιότητα του εκπαιδευτικού λογισμικού που διατίθεται στην τάξη σας