Semantic Image Analysis Using a Learning Approach and Spatial Context

G. Th. Papadopoulos[1,2], V. Mezaris[2] , S. Dasiopoulou[1,2] and I. Kompatsiaris[2]

[1] Informatics and Telematics Institute
Centre for Research and Technology Hellas

[2] Aristotle University
Of Thessaloniki

1st International Conference on Semantic and Digital Media Technologies
(SAMT 2006), **December 6-8, Athens, Greece**

Outline

- Introduction
  - problem definition
  - Knowledge-assisted analysis
- Proposed approach - overview
- Knowledge Infrastructure
- Knowledge-assisted analysis approach
  - Segmentation and feature extraction
  - Support vector machines
  - Spatial relation extraction
  - Genetic algorithm
- Experimental Results
- Conclusions

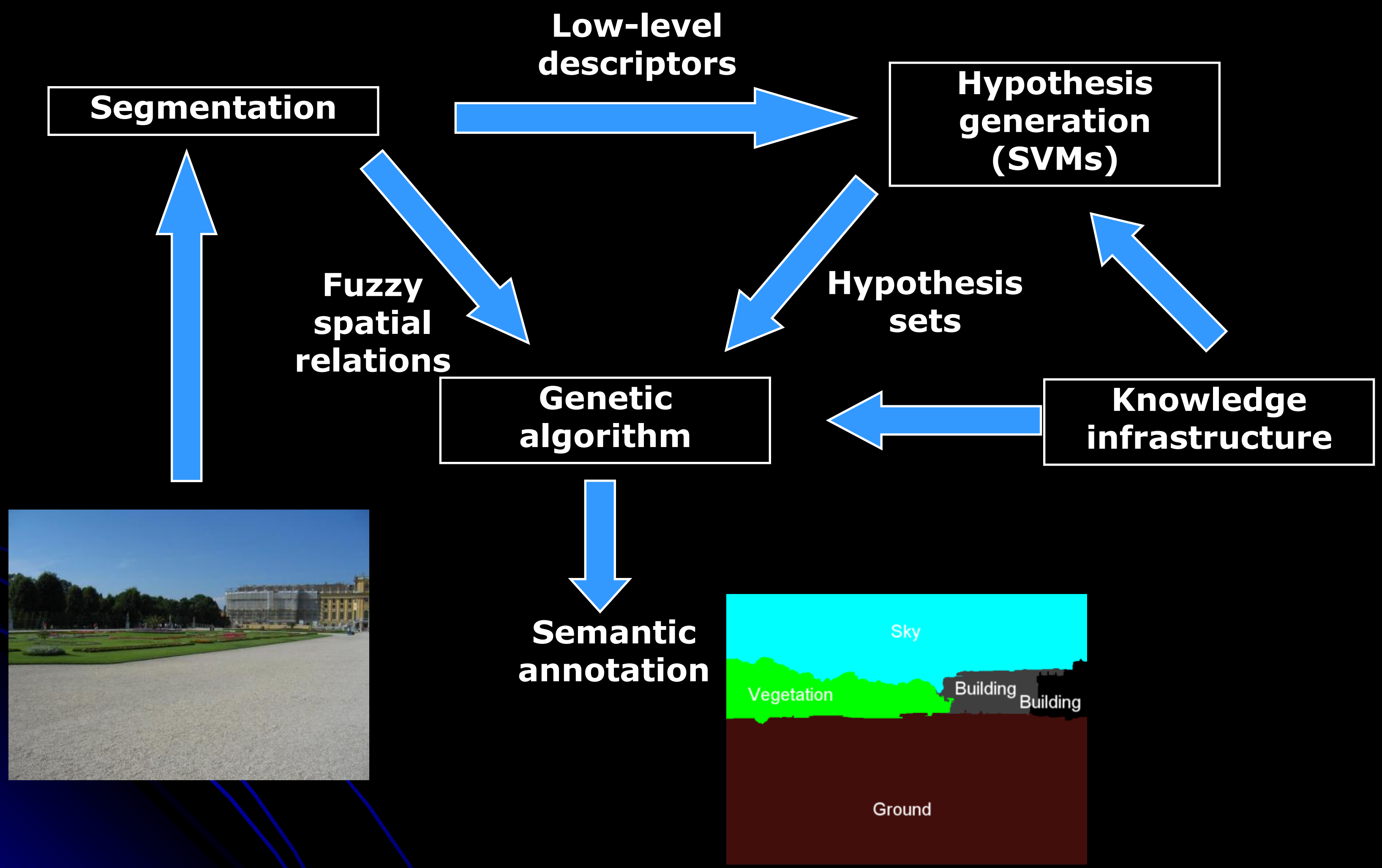
Introduction

- Enormous amount of digital images available on the internet and professional or personal collections, more digital content created at astonishing rates
- Image manipulation (search, retrieval, etc.)
  - Has become everyday practice
  - Is not yet efficient enough
    - Largely based on keywords
    - Often requires manual effort for annotation
  - Can greatly benefit from taking into account the semantics of image content
- Problem definition: understanding the image content
  - Knowledge-assisted image analysis techniques have emerged

Introduction

- Knowledge-assisted image analysis
  - Two main categories, depending on the adopted knowledge acquisition and representation process:
  - Implicit knowledge:
    - Machine learning methods (SVMs, GAs, HMMs, etc)
    - Robust for discovering complex relationships and interdependencies
    - Can handle problems of high-dimensionality
  - Explicit knowledge:
    - Model-based approaches (ontologies, rules, etc)
    - Use explicitly defined facts, models, relations and rules
    - Support “visual” inference in the specified context
  - Proposed approach combines implicit and explicit knowledge

Proposed approach - overview

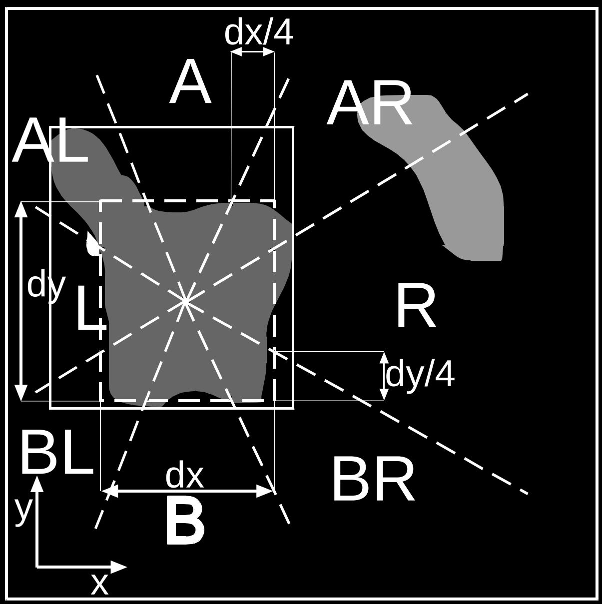


Knowledge Infrastructure

- Explicit knowledge represented in the form of a domain ontology
- Ontologies
  - Tools for defining explicit and machine-processable semantics
  - Support automated inference
- Proposed domain ontology includes
  - objects of interest (concepts)
  - spatial context

Knowledge Infrastructure

- Objects of interest
  - E.g. *Sky*, Water, Ground, etc.
  - Defined by expert
  - Matching objects with regions based on implicit knowledge (SVMs) acquired by training
- Spatial context
  - Based on 8 fuzzy spatial relations, defined by expert
  - Computed with the help of reduced region bounding box and cone-shaped areas
  - Learnt by training



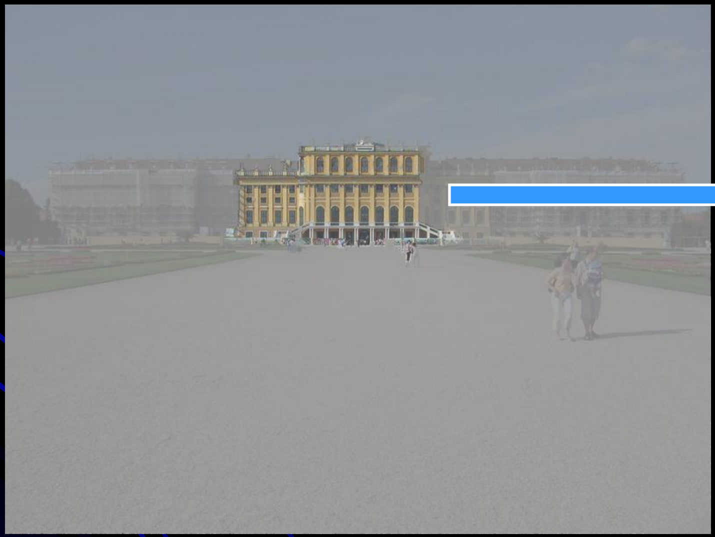
A above *AR above-right*
B below *AL above-left*
L left *BR below-right*
R right *BL below-left*

Segmentation and feature extraction

- Segmentation
  - Using extension of the Recursive Shortest Spanning Tree (RSST) algorithm
    - Fast
    - Produces accurate region boundaries
  - Forms regions, does not associate them with objects
- Descriptor extraction
  - MPEG-7 descriptors
    - Scalable Color
    - Edge Histogram
    - Region Shape
    - Homogeneous Texture
  - More descriptors could be employed to improve performance

Support vector machines

- Learning algorithm suitable for high-dimensional data
- Proposed approach
  - One SVM trained for every object, using low level descriptors
  - Each SVM estimates degree of confidence for region-object matching
  - Each region evaluated by all trained SVMs – hypothesis set created

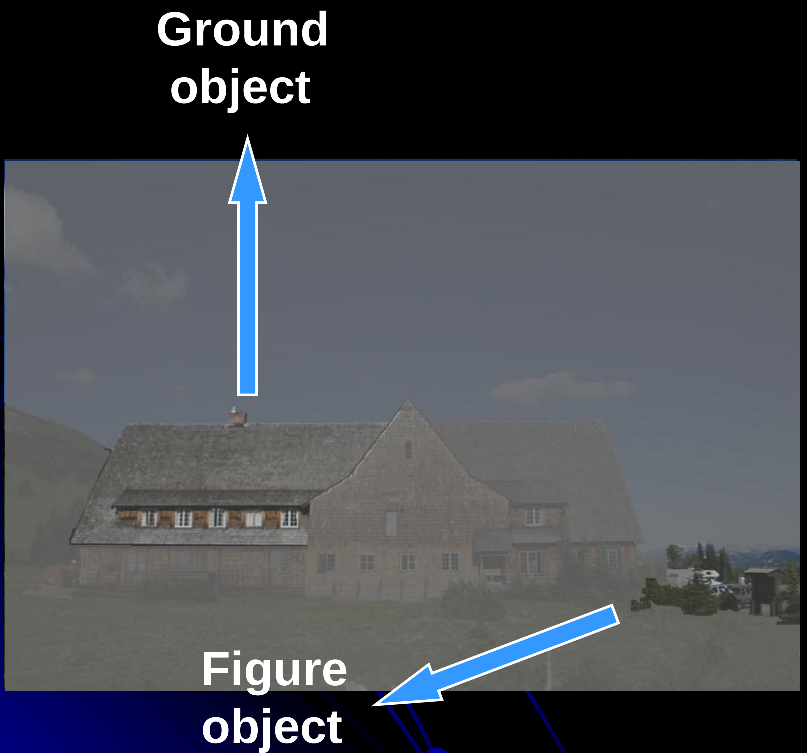


Segment's hypothesis set

Sky: 0.11
Water: 0.09
Building: 0.89
Rock: 0.51
Ground: 0.31
Vegetation: 0.35

Spatial relation extraction

- Using 8 fuzzy spatial relations defined in the ontology
- Proposed approach
  - All relations are evaluated for each pair of regions



Relative position of Figure object with respect to Ground object

Right: 0.39
Above-right: 0
Above: 0
Above-left: 0
Left: 0
Below-left: 0
Below: 0
Below-right: 0.61

Genetic Algorithm

- Appropriate for solving global optimization problems
- Proposed approach
  - Region-object association with respect to spatial context is defined as an optimization problem
    - Each chromosome represents a possible image interpretation
    - Each gene of a chromosome represents association of a region with a object
  - Fitness function takes into account
    - Hypothesis sets generated by SVMs, i.e. region-object matching based on visual descriptors
    - Spatial context, i.e consistency of spatial relations between regions in the image with relations stored in the ontology for the objects

Genetic Algorithm

- Fitness function
  - Parameter λ adjusts weight of spatial relations consistency
- Visual similarity based term calculated as normalized sum of degrees of confidence in hypothesis sets
- Two approaches for evaluating the consistency of spatial relations
  - Using a Euclidean distance-based function
  - Using a set of Triangular fuzzy membership functions

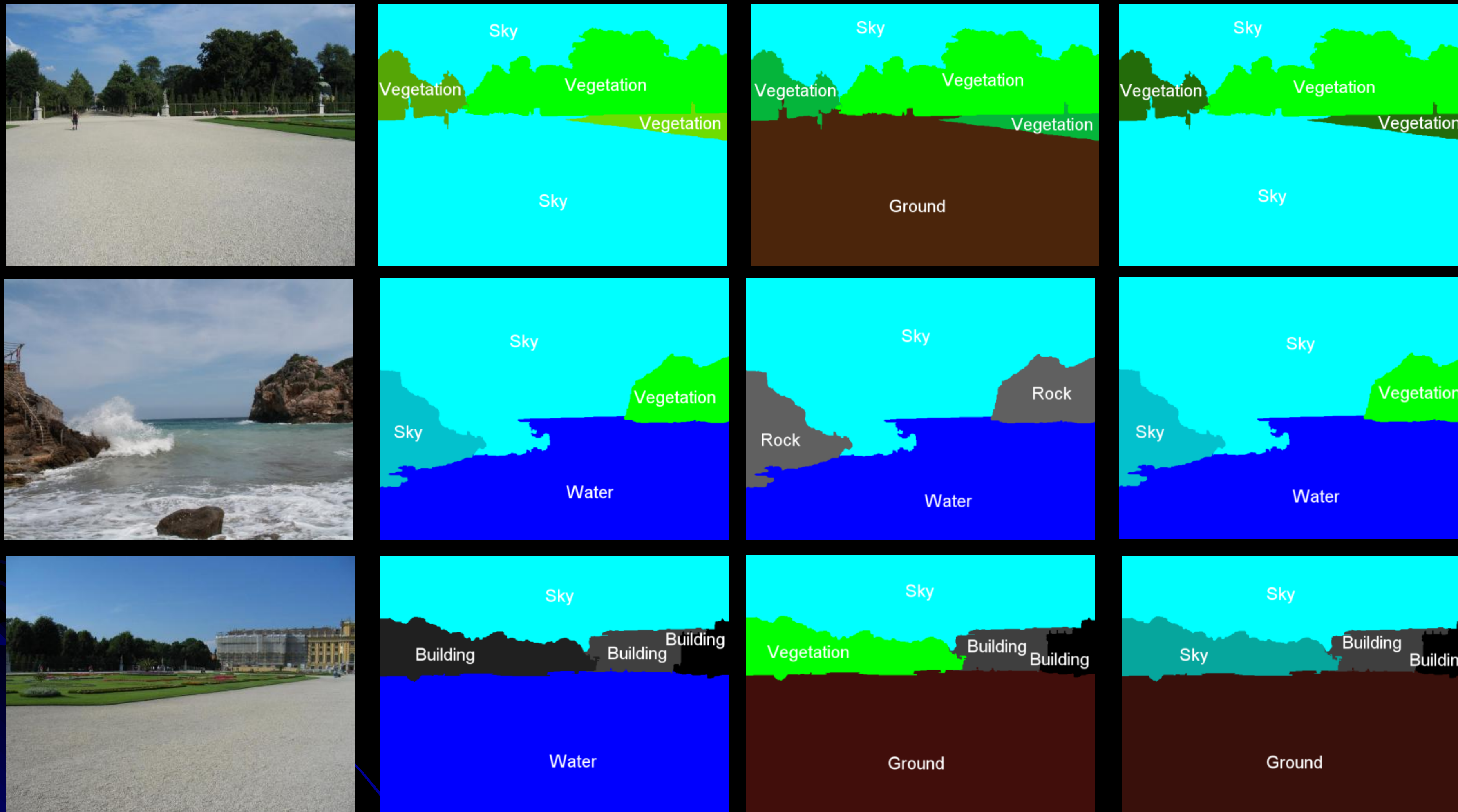
Genetic Algorithm

- Implementation
  - A population of 200 individual chromosomes is initialized with respect to the initial hypotheses
  - Selection, crossover, mutation used for evolution of population
  - GA termination condition
    - the diversity of the population becomes equal to/less than 0.001, or
    - the number of generations exceeds 50

Experimental Results

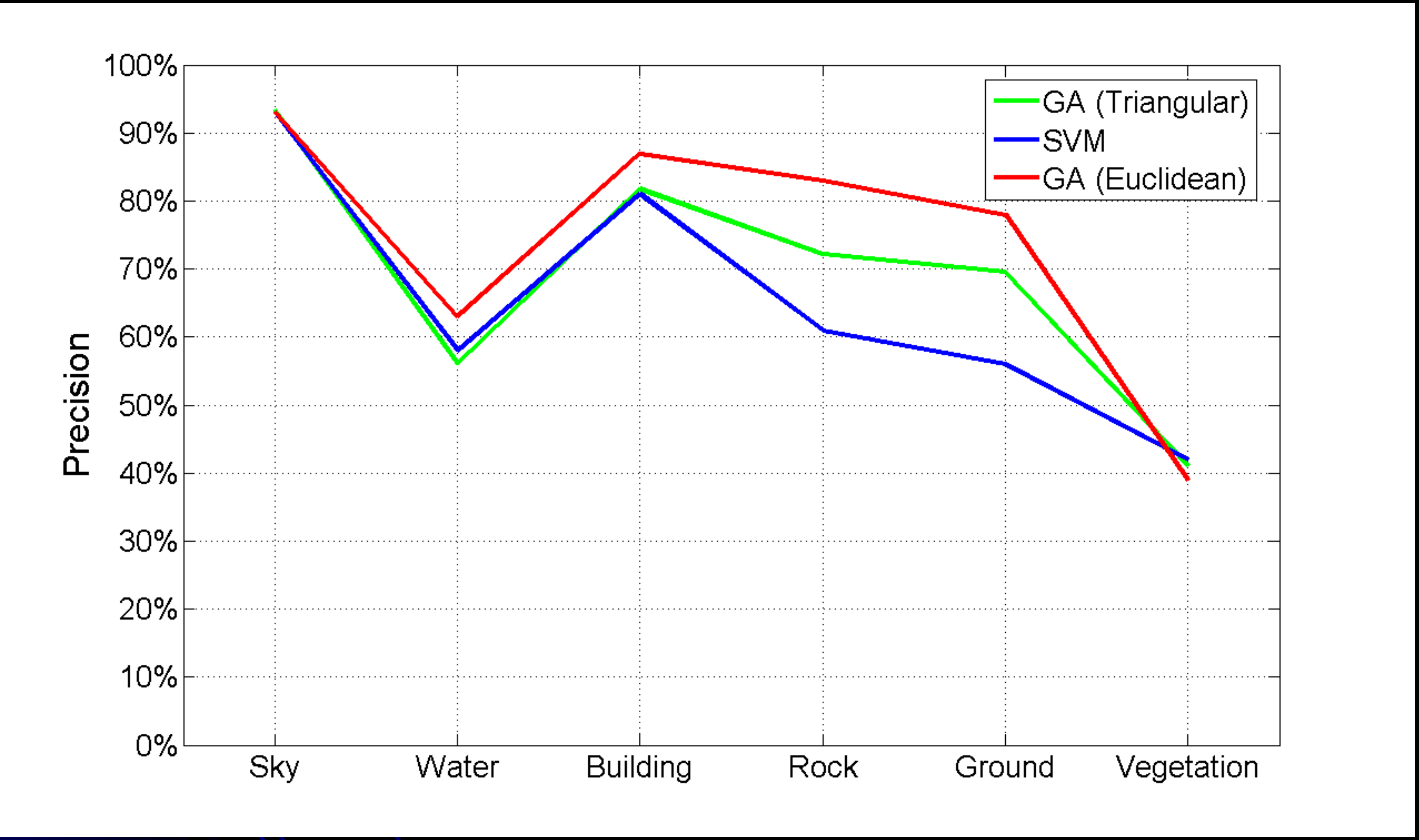
- Domain of experimentation: Outdoor Images
- 6 Supported concepts:
  - Sky
  - Water
  - Ground
  - Building
  - Vegetation
  - Rock
- 400 testing images

Experimental Results

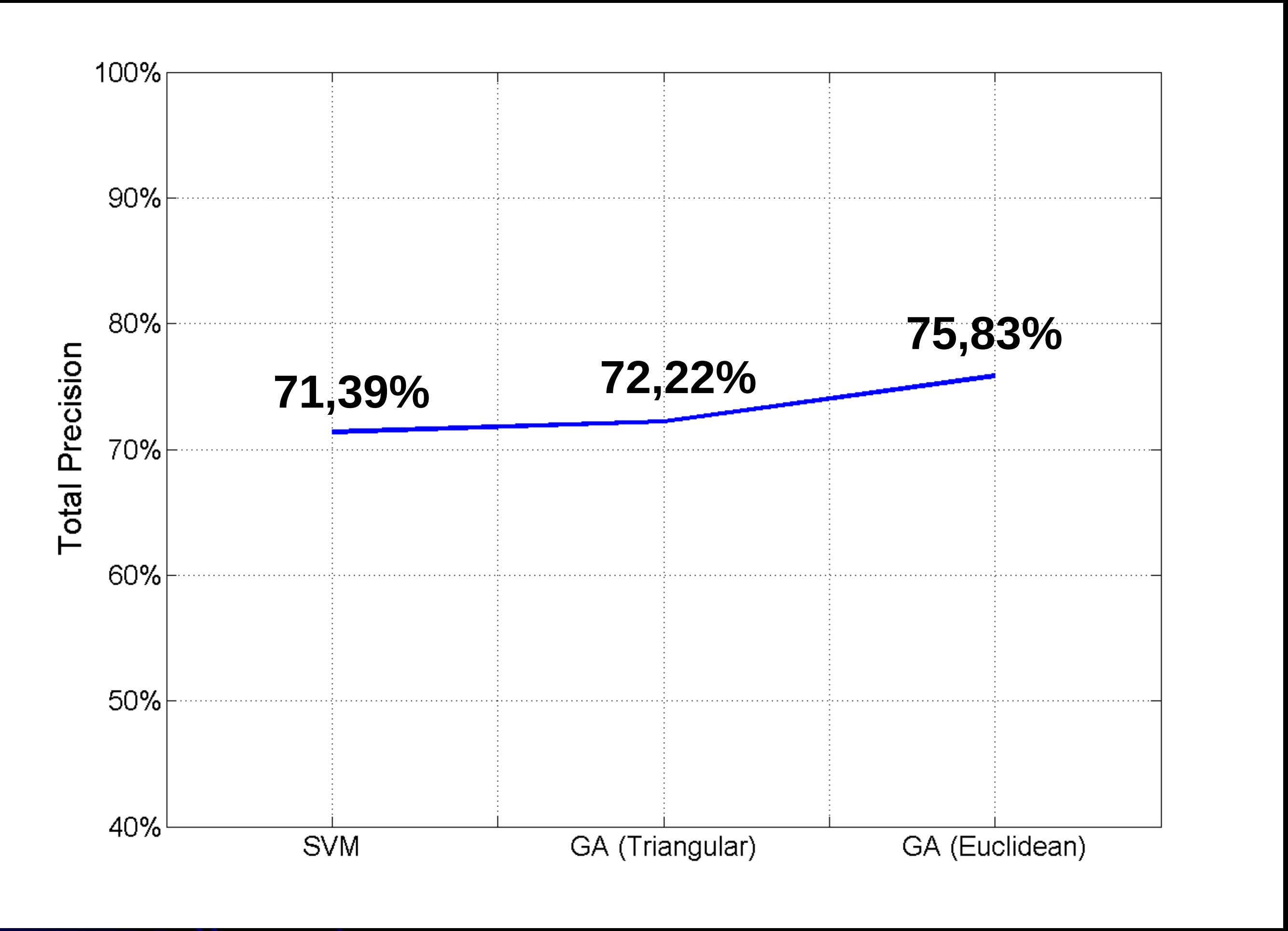


Original Image	Initial Hypotheses	GA (Euclidean) Interpretation	GA (Triangular) Interpretation

Experimental Results



Experimental Results



Conclusions & Future Work

- Conclusions
  - Combination of explicit and implicit knowledge efficient and effective
  - Formulation of semantic image analysis as an optimization problem produces promising results
  - Exploitation of domain-specific spatial-related contextual information improves object detection
  - Euclidean distance based evaluation of spatial constraints produces better results than using a set of triangular fuzzy membership functions
- Future work
  - Examine new approaches for evaluating spatial constraints
  - Extend the proposed framework to support more concepts and multiple domains, contextual information