VIDEO OBJECT WATERMARKING USING HU MOMENTS

Paraskevi Tzouveli, Klimis Ntalianis and Stefanos Kollias

School of Electrical and Computer Engineering
National Technical University of Athens
Iroon Polytexneiou 9, 15780 Zografou, Greece
phone: +302107724352 fax: +302107722492
email: tpar@image.ntua.gr
web: http://www.image.ntua.gr

ABSTRACT

With the rapid growth of images database and computer networks broadcasting, the legal issues of multimedia copyright protection have become very important. In this paper, a method for video object watermarking using GVF snake and HU moments is proposed, providing copyright protection of the semantic content. To achieve this goal, the contour of the objects of an image is initialised by the user and then using a GVF snake the real position of the video objects contours are retrieved. Next, an invariant watermark is designed using invariant Hu moments of each video objects. The proposed algorithm is tested against various attacks such as JPEG lossy compression, blurring, filtering and cropping. Experimental results on real life images indicate the efficiency and robustness of the proposed scheme.

1. INTRODUCTION

Nowadays, the increasing importance of digital media, however, brings also new challenges as it is now straightforward to duplicate and even manipulate multimedia content. There is a strong need for security services in order to keep the distribution of digital multimedia work both profitable for the document owner and reliable for the customer.

Watermarking technology plays an important role in securing as it places an invisible mark in the multimedia data which is used to identify the legal owner. Many watermarking schemes [1]- [5] have been proposed for ownership protection, providing significant resistance to image processing attacks. The design of resistant to geometrical distortions watermarking systems remains an open and challenging problem, leading many researching activities to the proposal of watermarking schemes resilient to geometric distortions [2],[6], [7] of images or video sequences.

However most watermarking systems deal with multimedia files as binary large objects, without taking into consideration regions of semantic information. These regions may need better protection or can be the only regions that need protection, depending on the specific application.

Towards this direction, in this paper we propose a video object watermarking method which is resilience to geometric distortions. Having extracted the semantic video objects of an image or video sequence, Hu moments of each video object are estimated and an invariant function is incorporated for watermarking. The automatic localization of objects of interest in an image or video sequence is a challenging task. Objects of interest are presented in many techniques with active contours [8]-[10]. Snakes, or active contours, are used extensively in computer vision and image processing applications, particularly to locate object boundaries. Problems associated with initialization and poor convergence to concave boundaries, however, have limited their utility. In the proposed method, we use the GVF snake [10] in order to extract the contour of the semantic video object. The GVF snake develops a new external force for active contours, the GVF force, largely solving both problems.

Hu [11] derived from the regular moments, seven moment invariants, which are rotation, scaling and translation invariant. In [12] Hu's moment invariant functions are used for watermarking. The watermark is embedded by modifying the moment values of the image. In their implementation, they had to do an exhaustive search to determine the embedding strength.

In this paper, we propose a method for video object watermarking based on HU moments and applied to the video object that the GVF snake has extracted. The embedding proposed method is consisted of two sub-modules: the video object detection sub-module and the watermark insertion sub-module. The video object detection is based on GVF snake taking as input a initial contour close to the object that the user wants to watermark and extracts the real contour of the video object. The watermark insertion sub-module computes the Hu moments of the enclosed area of the contour that the GVF snake has extracted. Then, modifies the Hu moments of the video object using a predefined function, which determines the moment invariant values in a predetermined space of values, the watermarked video object is provided.

So, the embedding method provides the image with the watermarked video object. As the moment invariants continue to exist after geometric distortions of a video object, the extraction method, initially, takes the contour of the semantic object which can be considered as the candidate watermarked area. Then the values of Hu moments of the candidate semantic video object are computed and a watermark extraction module characterises the area as watermarked if it is. Finally, in this paper the experimental results indicate the efficiency and robustness of the proposed method.

2. ACTIVE CONTOURS AND GRADIENT VECTOR FLOW

The automatic localization of objects of interest in an image or video sequence is a challenging task. Objects of interest are presented in many techniques with active contours [3]-[5]. In the proposed method, we use the GVF snake [5] in order to extract the objects of interest from a video sequence (video object) and accelerate the proposed video object tracking procedure.

The basic idea in active contour models is to evolve a curve, subject to constraints from a given image, in order to detect objects in it. Starting with this curve within the image domain and moving it under the influence of internal and external forces derived from image data, we can acquire the boundaries of the objects of interest.

A new external force for active contours, called Gradient Vector Flow, has been proposed in [5], trying to tackle problems that are associated, with initialization and poor convergence, to boundary concavities. The GVF snake begins with the calculation of a field of forces, called the GVF forces, over the image domain. The GVF forces are used to drive the snake, modelled as a physical object having a resistance to both stretching and bending, toward the boundaries of the object. The GVF forces are calculated by applying generalized diffusion equations to both components of the gradient of an image edge map. Because the GVF forces are derived from a diffusion operation, they tend to extend very far away from the object so that snakes can find objects that are quite far away from the snake's initial position. This same diffusion creates forces which can pull active contours into concave regions. Experimental result is depicted in Figure 1 where the contour of semantic object (Akiyo) is depicted with red colour.

Figure 1. Semantic Object Extraction using GVF snake

3. WATERMARKING AND MOMENTS

Moments and functions of moments have been utilized as pattern features in a number of applications [13]-[16]. These features can provide global information about the image. Geometric moments and the corresponding moment invariants are reviewed in this section for completeness. Before presenting moment invariants, one needs to give some intuition on why they are invariant to general affine transform.

It is known that the effect of an affine transform on images is to translate, scale, and rotate the image. Then, the required invariants have to be invariant to translation, scaling, and orientation. *Geometric moments* (regular moments) are non-negative integers, which can be computed by:

$$m_{pq} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_p y_q f(x,y)dxdy$$

where m_{pq} is the (*p*+*q*)th order moment of any real continuous image function *f*(*x*,*y*) having a bounded support and a finite nonzero integral. In a digital implementation, these integrals are approximated by summations over small, discrete areas: $m_{pq} = \sum_x \sum_y x_p y_q f(x,y)$.

Corresponding *central moment* $\mu_{pq}^{(f)}$ of order (*p*+*q*) of the image *f*(*x*,*y*) are defined analogously as

$$\mu_{pq} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\bar{x})^p (y-\bar{y})^q f(x,y)dxdy$$

where the coordinates $\bar{x} = \frac{m_{1,0}}{m_{0,0}}, \bar{y} = \frac{m_{0,1}}{m_{0,0}}$ denote the centroids of *f*(*x*,*y*). The central moments of the image are invariant to translation as they are origin-independent.

Scaling invariance can be achieved by normalizing the moments of the scaled image by the scaled energy of the original. For this propose, further normalization for the effects of scaling can be computed by the definition of the formula: $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}{}^{\gamma}}$ where γ is the normalization factor $\gamma = \frac{p+q}{2} + 1$.

Traditionally, moment invariants are computed based on information provided by both the shape boundary and its interior region [11]. Hu [11] first introduced the mathematical foundation for two-dimensional moment invariants in 1962, based on methods of algebraic invariants and demonstrated their applications to shape recognition [11]. Using nonlinear combinations of geometric moments, a set of invariant moments, which have the desirable properties of being invariant under image translation, scaling, and rotation, is provided by the Hu method.

$$\phi_1 = n_{20} + n_{02}$$
$$\phi_2 = (n_{20} - n_{20})^2 + 4n_{11}^2$$
$$\phi_3 = (n_{30} - 3n_{12})^2 + (n_{03} - 3n_{21})^2$$
$$\phi_4 = (n_{30} - n_{12})^2 + (n_{03} + n_{21})^2$$
$$\phi_5 = (3n_{30} - 3n_{12})(n_{30} + n_{12}) \cdot \left[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2 \right] + (3n_{21} - n_{03})(n_{21} + n_{03}) \cdot \left[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2 \right]$$
$$\phi_6 = (n_{20} - n_{02}) \cdot \left[(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2 \right] + 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03})$$
$$\phi_7 = (3n_{21} - n_{03})(n_{30} + n_{12}) \cdot \left[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2 \right] + (3n_{12} - n_{03})(n_{21} + n_{03}) \cdot \left[3(n_{30} + n_{12})^2 - (n_{21} + n_{30})^2 \right]$$

Hu defines seven of these shape descriptor values, computed from central moments through order three, independent to object translation, scale and orientation. Translation invariance is achieved by computing moments that are normalised with respect to the centre of gravity so that the centre of mass of the distribution is at the origin (central moments). Size invariant moments are derived from algebraic invariants but these can be shown to be the result of simple size normalization. From the second and third order values of the normalized central moments, a set of seven invariant moments can be computed which are rotation independent. From the normalized central moments, a set of seven values, Hu moments, can be calculated by the following equations:

4. WATERMARK EMBEDDING METHOD

An overview of the proposed system's watermark embedding module is depicted in Figure 2. The embedding proposed method is consisted of two sub-modules: the video object detection sub-module and the watermark insertion sub-module.

Initially, the semantic video object *O* is extracted by GVF snakes detection modules. The watermark insertion sub module applies a mapping function Π to the extracted semantic video object *O*, adding noise and producing a modified semantic video object, which is called $\Pi(\mathrm{O})$. In our case study $\Pi(\mathrm{O}) = \log(\mathrm{O})$.

Now let us consider that $\Phi = [\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7]^T$ is the invariant Hu moments of the original semantic video object O. Let also $\Phi^* = [\phi_1^*, \phi_2^*, \phi_3^*, \phi_4^*, \phi_5^*, \phi_6^*, \phi_7^*]^T$ be the invariant Hu moments of the watermarked semantic video object $\tilde{\mathrm{O}}$.

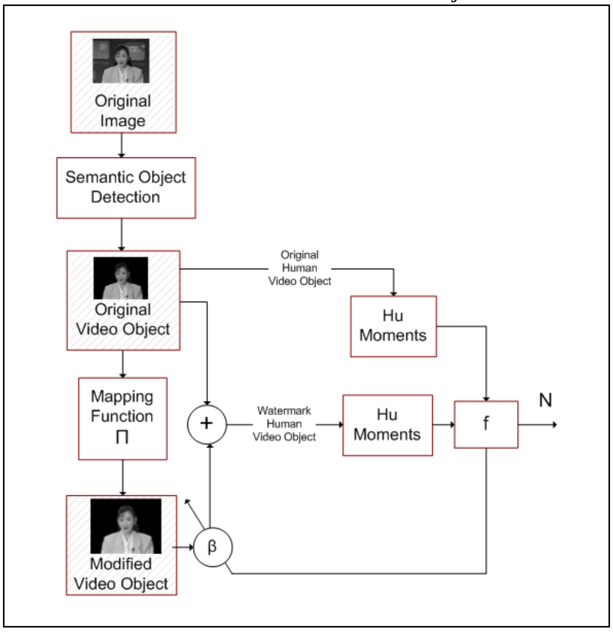

Figure 2. Block diagram of embedding proposed method

We can choose a function f, which can be any linear or non-linear combination of the invariant moments. In our case study, the function f is expressed as a sum value of the weighted average differences between Hu moments of the original semantic video object, ϕ, and the watermarked semantic video object, ϕ^*:

$$\mathrm{f}(\Phi^*, \Phi) = \sum_{i=1}^{7} w_i \left(\frac{\phi_i^* - \phi_i}{\phi_i} \right) \qquad (1)$$

where weights w_i take the values $w_1 = 1.5, w_2 = 1.25, w_3 = 1, w_4, w_5 = 0.75, w_6 = 0.50$ and $w_7 = 0.50$. These values have been set after several experimental tests, since the first and second Hu moments are the most robust compared to the other moments [11], so the values of the weighted factors w_1, w_2 are higher than the other weighted factors. The output of function f is called factor N (Figure 2).

The watermarked semantic video object $\tilde{O}$ can be computed by adding a perturbation ΔO to the original semantic video object O: $\tilde{O} = O + \Delta O$. The perturbation ΔO is produced by the following multiplication: $\Delta O = \beta \cdot \Pi(O)$. The weighted factor β is controlled by feedback in order to ensure that $f(\Phi^*, \Phi) \approx 20\%$.

5. WATERMARK EXTRACTION METHOD

The block diagram of the watermark extraction method is depicted in Figure 3. . The proposed extraction method is consisted of two sub-modules: the detection of the candidate video object sub-module and the watermark extraction sub-module.

Initially, a candidate image or video sequence (that may contain watermarked video objects) is received as an input to the extracted method. We assume that the candidate image contains a watermarked video object that has been geometric attacked. The moment invariants continue to exist after geometric distortions of a video object.

The detection sub-module of the extraction method takes as input the candidate image or video sequence and the user of the proposed system draws the initial contours of the video objects in which he/she characterises as candidate object. These initial contours are the inputs to the GVF snake which finally extracts video object of the candidate image or video sequence which can be considered as the candidate watermarked area.

In order to make it clear, let us assume that a malicious user has cropped the watermarked semantic video object (Akiyo) and has placed it in a different content. Let us assume that this image is the candidate image for the extraction method, as is depicted in Figure 3. The user of the proposed system draws a general contour around to Akiyo's video object and the GVF snake extracts exactly the Akiyo's video object.

Having received the video object that the detection sub-module has extracted as possible watermarked video object, the watermark extraction sub-module of the detection module estimates the Hu moments of this object.

In parallel, the watermark extraction sub-module of the detection procedure receives in its input, the seven values of the Hu moments of the respective watermarked semantic video object. Now, the value of factor *N* can be found, counting the mean value of weighted average differences between Hu moments of the watermarked semantic video object, ϕ^*, and of the candidate region, ϕ'.

Authentication of the received semantic video object can be achieved by checking the validity of the equation $\mathrm{N} \le \varepsilon$ where ε is the margin of acceptable error between the two video objects.

Then, the detection procedure returns either 1, meaning that the candidate video object is the watermarked video object which has passed the proposed watermarking embedding procedure, or 0, meaning that the candidate video object is not watermarked.

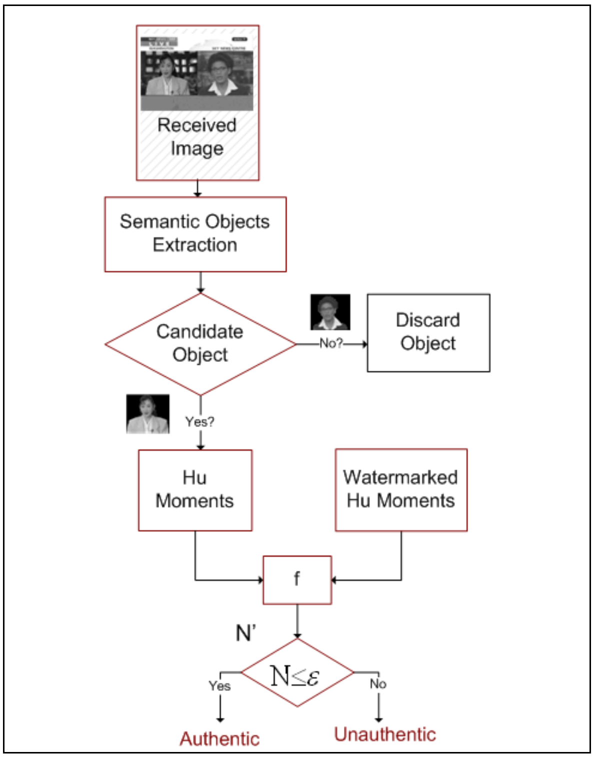

Figure 3. Block Diagram of the Watermark Detection Method

6. EXPERIMENTAL RESULTS

The GVF snake can sufficiently extract the semantic video object of an image which has passed the proposed watermarking embedding procedure and has received attacks i.e. Gaussian noise, jpeg compression, by adjusting threshold T to 40%.

Table I contains the absolute log values of the invariant Hu moments for semantic video object (Akiyo) for the original object and after several different attacks. Examination of Table I reveals the robustness of the invariants. It is noticeable that the variations between the same invariants of the watermarked semantic video object and the watermarked semantic video object under different attacks are insignificant.

In order authenticate a watermarked semantic video object that an image may include, we measure the value of the function f which is computed as the mean value of weighted average differences between Hu moments of the watermarked semantic video object and the watermarked semantic video object under different attacks. So, we assume that the watermarked semantic video object detection is achieved for $\varepsilon \leq 0.01$.

Table II depicts the robustness of the proposed technique under different attacks as well as noise addition, and image distortions. The proposed watermarking scheme serves as a 1-bit watermarking system as it answers the yes/no question of authenticity. It is obvious that the watermarked video object can be detected after most attacks. Also in this case threshold ε is set equal to 0.01. This means that the attacked video object can be different from the watermarked object only by 1%. Robustness of the proposed scheme to geometric manipulations is guaranteed. This is due to the fact the used moments are designed to be invariant to geometric attacks.

TABLE I. HU MOMENTS INVARIANCE FOR AKIYO IMAGE

Attack	Hu Moments Invariance						
	ϕ_1^*	ϕ_2^*	ϕ_3^*	ϕ_4^*	ϕ_5^*	ϕ_6^*	ϕ_7^*
Original Image	6.63	17.54	20.36	24.05	46.58	32.90	46.64
Watermark Image	6.81	17.94	20.88	24.57	47.59	33.77	47.70
Gaussian Filtering	6.79	18.16	21.00	24.14	46.92	33.43	47.23
Median Filtering	6.78	17.91	20.88	24.59	47.63	33.78	47.73
JPEG Q=50%	6.82	17.94	20.88	24.57	47.58	33.77	47.69
JPEG Q=10%	6.80	17.96	20.87	24.55	47.54	33.75	47.69
Rotation 5^o	6.80	18.20	20.95	24.34	47.13	33.59	47.69
Rotation -1^o	6.78	17.94	20.88	24.57	47.59	33.77	47.70
Scaling 50%	6.78	17.95	20.89	24.58	47.60	33.78	47.71
Scaling 110%	6.76	17.90	20.88	24.59	47.62	33.77	47.72
Cropping 10%	6.43	16.95	23.86	22.95	46.62	31.80	47.03
Flipping	6.78	17.94	20.88	24.57	47.59	33.77	47.81
False Alarm	1.35	6.91	4.27	7.94	14.28	11.89	14.53

It can be noticed in Table II that if the watermarked video object is cropped, the detector cannot easily authenticate the content. The only difference between the original and the watermarked video objects is in their contrast, which is expected since moments and image histograms are tightly related. This is the price paid for achieving such robust watermarking. It should be emphasized that the decoder does not need to receive the original video object for detecting the watermark. The decoder only needs to receive the values of the moments of the watermarked video object. This is considered also as an advantage of the proposed system, while both the encoder and decoder have low complexity. Finally it should be noted that during detection just the error and the margin of error are computed.

7. CONCLUSIONS

In order to provide a copyright protection we propose a semiautomatic method for video object watermarking using GVF snake and HU moments. For this propose, a global contour of the objects of an image is initialised by the user and then using a GVF snake the real position of the video objects contours are retrieved. Next, an invariant watermark is designed using invariant Hu moments of each extracted video objects. As the moment invariants continue to exist after geometric distortions of a video object, the extraction method, initially, takes the contour of the semantic object which can be considered as the candidate watermarked area. Then the values of Hu moments of the candidate semantic video object are computed and using the proposed watermark extraction method the extracted video object is characterized as watermarked or not. Experimental results indicate the robustness of the proposed method.

TABLE II. RESULTS OF DETECTING THE WATERMARKED AKIYO VIDEO OBJECT AFTER SEVERAL ATTACKS

Attack	$f(\Phi^*, \Phi) = \frac{1}{7}\sum_{i=1}^{7} w_i \left(\frac{\phi_i^* - \phi_i}{\phi_i} \right)$	**Watermark Detection**
Gaussian Filtering	0.004023	Pass
Median Filtering	0.004023	Pass
JPEG Q=50%	0.001229	Pass
JPEG Q=10%	0.004691	Pass
Rotation 5^o	0.044546	Pass
Rotation -1^o	0.000012	Pass
Scaling 50%	0.002085	Pass
Scaling 110%	0.003790	Pass
Cropping 10%	0.40962	Fail
Flipping	0.002162	Pass

REFERENCES

[1] M. Kutter, S. K. Bhattacharjee and T. Ebrahimi, “Towards second generation watermarking schemes” in *Proc. IEEE Int. Conf. Image Processing* 1999, pp. 320–323.

[2] J. J. K. O’Ruanaidh and T. Pun, “Rotation, scale, and translation invariant spread spectrum digital image watermarking,” *Signal Processing*, vol. 66, pp. 303–317, 1998.

[3] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, “Rotation, scale, and translation resilientwatermarking for images,” *IEEE Trans. Image Processing*, vol. 10, pp. 767–782, May 2001.

[4] H. S. Kim, Y. Baek, and H. K. Lee, “Rotation-, scale-, and translation invariant image watermark using higher order spectra,” *Opt. Eng.*, vol. 42, no. 2, pp. 340–349, Feb. 2003.

[5] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Attacks on copyright watermarking systems,” in *Proc. 2nd Int.Workshop Information Hiding*, 1998, pp. 218–238.

[6] S. Pereira and T. Pun, “Robust template matching for affine resistant image watermarks” *IEEE Trans. Image Processing*, vol. 9, pp. 1123–1129, July 2000.

[7] M. Kutter, “Watermarking resisting to translation, rotation, and scaling,” in *Proc. SPIE Multimedia Systems Applications*, 1998, pp. 423–431.

[8] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active Contour models,” Int’l J. Computer Vision, vol. 1, pp. 312-333, 1988.

[9] H. S. Ip and S. Dinggang, “An Affine-Invariant Active Contour Model (AI-Snake) for Model-Based Segmentation,” Image and Vision Computing, 16(2), pp. 135-146, 1998.

[10] C. Xu and J.L.Prince, “Gradient Vector Flow: A New External Force for Snakes” IEEE Proceedings Conference on Computer Vision and Pattern Recognition, pp 66-71, 1997

[11] M. K. Hu, “Visual pattern recognition by moment invariants,” *IEEE Trans. Inform. Theory*, vol. 8, pp. 179–187, 1962.

[12] M. Alghoniemy and A. H. Tewfik, “Image watermarking by moment invariant,” in *Proc. IEEE Int. Conf. Image Processing*, 2000, pp. 73–76.

[13] B. A. Dudani, K. J. Breeding, and R. B. McGhee, “Aircraft identification by moment invariants,” *IEEE Trans. Comput.*, vol. C-26, pp. 39–46, 1977.

[14] Y. S. Abu-Mostafa and D. Psaltis, “Image normalization by complex moments,” *IEEE Trans. Pattern. Anal. Machine Intell.*, vol. PAMI-7, pp. 46–55, 1985.

[15] J. Wood, “Invariant pattern recognition: a review,” *Pattern Recognit.*, vol. 29, no. 1, pp. 1–17, 1996.

[16] A. P. Reeves, R. J. Prokop, S. E. Andrews, and F. P. Kuhl, “Three dimensional shape analysis using moment and Fourier descriptors,” *IEEE Trans. Pattern. Anal. Machine Intell.*, vol. 10, pp. 937–943, 1988.