CONFRONTING THE SYNCHRONIZATION PROBLEM OF SEMANTIC REGION UNDER GEOMETRIC ATTACKS

Paraskevi Tzouveli, Klimis Ntalianis and Stefanos Kollias

National Technical University of Athens
Electrical and Computer Engineering Department
Iroon Polytexneiou 9, 15780, Athens, Greece
email: (tpar,kntal)@image.ntua.gr

ABSTRACT

In this paper, an affine invariant watermarking scheme, robust to geometric attacks, is proposed and applied to face regions. Initially, face regions are unsupervisedly extracted from an initial image and a normalization procedure, invariant to geometric attacks is applied on each of these regions using a set of specific moment criteria. A spread spectrum-based DS-CDMA watermarking scheme is then used in order to provide a multi bits watermark. The multi bits watermark embedding and detection procedures are then applicable to each normalized face region. Finally, performance of the proposed face regions watermarking scheme is tested under various distortions, providing efficient and robust watermark retrieval.

1. INTRODUCTION

With information exchange and broadcasting, copyright protection of digital media has become a significant scientific field. For this purpose, several watermarking techniques have been presented in the literature trying to confront the problem of copyright protection. Many of the watermarking techniques are not resistant enough to geometric attacks. Such attacks as rotation, scaling, translation, shearing can severely affect a watermarked image, destroying the synchronization of the watermark bit stream during detection. Without synchronization the watermark cannot be extracted. Since the recognition of the need for watermarking schemes resilient to geometric distortions, many researchers [1]-[5] try to design watermarking techniques resistant to rotation, scaling, translation and shearing of images.

Most of them are based on the invariant property of the Fourier transform. Particularly, in [2], a method based on the invariant properties of Fourier–Mellin transform (FMT) was proposed to deal with geometric attacks. However, this method was effective in theory, but difficult to implement. In [6], a template was embedded in the DFT domain of the image, which was used to estimate the affine geometric attacks. Initially, the estimated distortion was corrected and afterwards watermark detection was performed. In [5], a watermarking scheme was proposed using moment based image normalization with a standard size and orientation. Thus, it is suitable for public watermarking where the original image is not available. The approach was used to embed a 1-bit watermark. Another watermarking scheme [3] aims at providing scale and translation invariance by utilizing an image normalization technique.

In most of the aforementioned techniques the watermark is a random sequence of bits and can only be detected by employing detection theory. The watermark is retrieved by subtracting the original from the candidate image and choosing an experimental threshold value to determine when the cross-correlation coefficient denotes a watermarked image or not. However most of the proposed techniques are frame-based and thus semantic regions, such as human faces, are not considered. Here it should be stressed that in several applications face regions are addressed as independent video objects and thus should be independently protected from the rest of the content.

Towards this direction and in order to automatically extract face areas, modern methods use skin colour characteristics. In [7], face detection is achieved, using a skin colour model based on the chrominance components of the YCrCb colorspace and a suitable face area shape model. The work presented in [8] proposes an adaptive 2D Gaussian model for skin colour distribution, whose parameters are re-estimated based on the current image. The mask area obtained from skin colour detection is processed using morphological tools.

In the proposed scheme, a face region watermarking scheme is presented, as face region is considered as a very important semantic feature. Towards this direction, initially detection of faces is performed based on a two-step process: detection of human skin regions and then extraction of face regions from the detected human regions. Afterwards a watermarking technique addressing the face regions of the image is presented, which alleviates the problem of geometric distortions. To achieve this goal, a normalization procedure, invariant to affine transform attacks is applied on each of the extracted face regions using a set of specific moment criteria. Afterwards, a spread spectrum-based DS-CDMA watermarking scheme is used in order to provide a

1-4244-0367-7/06/$20.00 ©2006 IEEE

multi bits watermark. Finally, a multi bits watermark embedding procedure is applied in each normalized face region. Performance of the proposed watermarking scheme is tested under various distortions, providing the desirable synchronization property in all cases.

2. UNSUPERVISED FACE DETECTION

It was shown in [7] that skin-tone colors are limited to a small area of the Cr-Cb chrominance plane of the YCrCb colourspace. Then all pixels of an image can be checked whether they belong to the skin tone color area or not by using a Bayesian formula. In the proposed face watermarking scheme, face detection is automatically performed using the algorithm of [8] According to this algorithm skin-tone colors distribution is approximated using a two-dimensional Gaussian density function. The adapted Gaussian model combined with a minimum risk threshold, estimated using the maximum likelihood criterion on the training set, is applied to the input image (Fig. 1a.) producing a binary image mask (Fig. 1b.), which guides the face watermarking procedure while in Fig. 1c. the detected face region is presented.

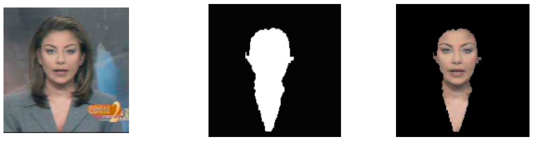

Figure 1. (a) Original Image, (b) Face Mask, (c) Face Region

Afterwards morphological operations (opening and closing) are applied to spatially filter the obtained image masks, while the morphological distance transform and size distribution techniques are used to isolate the disconnected areas and provide separate skin segments. Shape features are also employed to discard skin segments that possess irregular shapes. Finally remaining segments are bounded by rectangles and pixel verification is performed within each rectangle according to the adopted algorithm.

3. IMAGE NORMALIZATION AND TRANSFORM PARAMETERS DEFINITION

Having extracted the face regions from the initial image, a normalization procedure, invariant to affine transform attacks is applied on each of these regions. The proposed watermarking scheme is based on the idea of using a normalized face region both for watermark embedding and detection. By normalizing the extracted face region, the invariance to any affine distortions is achieved, providing integrity of the watermark. This invariance is constructed using the central moments of the face region of the image as function parameters.

Let *f(x,y)* denote a face region of an image and *g(x,y)* denote the affine transformation of *f(x,y)* obtained with affine matrix $A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$ and translation $d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ such that $g(x,y)=f(x_a,y_a)$ where:

$$\begin{bmatrix} x_a \\ y_a \end{bmatrix} = A \cdot \begin{bmatrix} x \\ y \end{bmatrix} - d \qquad \text{(Eq. 1)}$$

If $d = 0$, the moments $m'_{p,q}$ of *g(x,y)* are described as [10]:

$$m'_{p,q} = \sum_{i=0}^{p}\sum_{j=0}^{q}\binom{p}{i}\cdot\binom{q}{j}\cdot\alpha_{11}^{i}\cdot\alpha_{12}^{p-i}\cdot\alpha_{21}^{j}\cdot\alpha_{22}^{q-j}\cdot m_{i+j,p+q-i-j} \text{(Eq. 2)}$$

while the central moments $\mu'_{p,q}$ of *g(x,y)* are described as:

$$\mu'_{p,q} = \sum_{i=0}^{p}\sum_{j=0}^{q}\binom{p}{i}\cdot\binom{q}{j}\cdot\alpha_{11}^{i}\cdot\alpha_{12}^{p-i}\cdot\alpha_{21}^{j}\cdot\alpha_{22}^{q-j}\cdot \mu_{i+j,p+q-i-j} \text{(Eq.3)}$$

In order to achieve an affine invariant transform which will eliminate the four parameters ($\alpha_{11},\alpha_{12},\alpha_{21},\alpha_{22}$), four moments should be normalized, providing specific moment criteria. For this purpose, the following description of the affine transformation matrix is adopted [9]:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \cdot \begin{bmatrix} \alpha & 0 \\ 0 & \delta \end{bmatrix} \cdot \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} = A_R \cdot A_S \cdot A_X$$

where A_R is the rotation invariant matrix, A_S is the scaling in both x and y directions matrix and A_X is the shearing in x direction matrix.

Firstly, in order to eliminate the translation of the affine attack, we set ($\alpha_{11},\alpha_{12},\alpha_{21},\alpha_{22}$)=(1,0,0,1) and $(d1, d2) = (m_{10}/m_{00}, m_{01}/m_{00})$ where m_{00}, m_{01}, m_{10} are the moments of *f(x,y)*. This procedure results in the creation of the centered face region $f_1(x,y)$.

Then, a shearing transform is applied to $f_1(x,y)$ in the x direction with the shearing matrix A_X:

$$f_2(x,y) = A_x[f_1(x,y)]$$

Based on the previous equation and by setting $\mu_{30}^{(2)} = 0$, we result to the following equation:

$$\mu_{30}^{(2)} = \beta^3\mu_{03}^{(1)} + 3\beta^2\mu_{12}^{(1)} + 3\beta\mu_{21}^{(1)} + \mu_{30}^{(1)} = 0 \text{ (Eq. 4)}$$

from which the parameter β can be estimated.

Afterwards, scale normalization of $f_2(x,y)$ in x and y directions is performed, using the matrix A_S so that the resulting image, denoted by $f_3(x,y) = A_s[f_2(x,y)]$,

achieves $\mu_{05}^{(3)} > 0$ and $\mu_{50}^{(3)} > 0$, forcing the normalized image to a standard size.

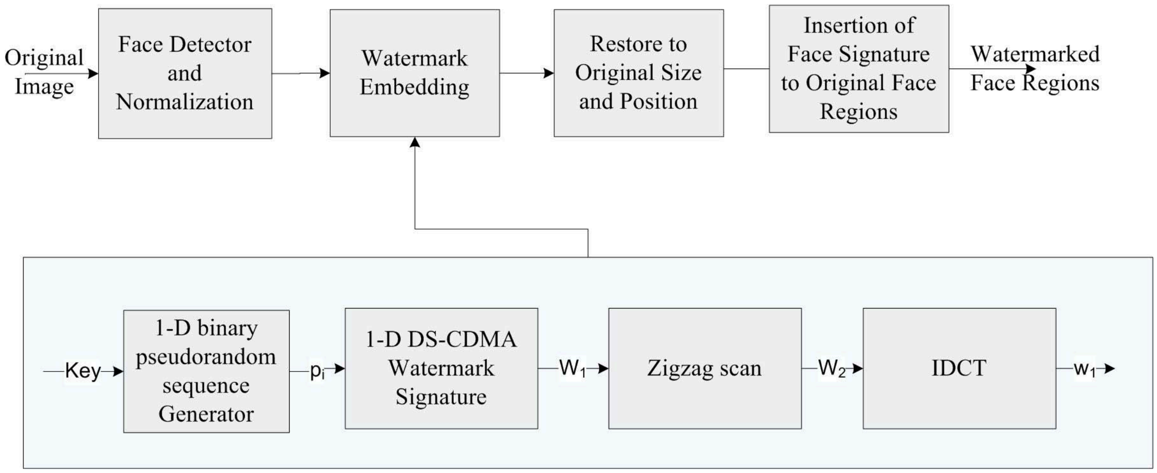
Original Image
Face Detector and Normalization
Watermark Embedding
Restore to Original Size and Position
Insertion of Face Signature to Original Face Regions
Watermarked Face Regions
Key
1-D binary pseudorandom sequence Generator
pi
1-D DS-CDMA Watermark Signature
W1
Zigzag scan
W2
IDCT
w1

Figure 2a: Block Diagram of Watermark Embedding Method

Finally, $f_3(x,y)$ is rotated using A_R, so that the resulting image, denoted by $f_4(x,y) = A_R[f_3(x,y)]$, is full normalized to affine transformations. Computation of angle φ is performed by requiring $\mu_{03}^{(3)} + \mu_{21}^{(3)} < 0$ and according to the following equation:

$$\phi = \arctan((\mu_{30}^{(3)} + \mu_{12}^{(3)}) / (\mu_{03}^{(3)} + \mu_{21}^{(3)})) \text{ (Eq. 5)}$$

It is important to note that each step in the normalization procedure is readily invertible. This will allow us to convert the normalized image back to its original size and orientation once the watermark is inserted.

4. WATERMARK EMBEDDING AND EXTRACTION

The proposed face region normalization-based watermark embedding method is illustrated in Fig. 2 while the watermark extraction method in Fig 3. Using the normalized face region, described above, the spread spectrum-based DS-CDMA watermarking scheme [11] is adopted, which is well known for its robustness to common signal processing attacks.

The steps that the watermark embedding method follows are demonstrated in Fig. 2. Firstly, the normalization procedure is applied to obtain a normalized face region and a normalized mask of the face region. Then, a 2D watermark with the same size as the normalized face region is created. Afterwards, the watermark signature is generated using the mask of the face region. Next, the inverse of the normalization procedure is applied to the watermark signature, so that the signature has the same size as the face region and then the face signature is added to the original face region, producing the final watermarked face area.

In order to create the aforementioned 2D watermark, M 1-D binary pseudo-random sequences pi, $i = 1,\ldots,M$ are generated using a private key as seed, where M is the number of bits in the watermark message ($m_i = \{0 \text{ or } 1\}$, $i = 1,\ldots,M$). Each of these sequences has zero mean and their values come from a binary alphabet {-1,1}. Then, a 1-D DS-CDMA watermark signature is created that modulates the watermark message: $W_1 = \sum_{i=1}^{M} (2m_i - 1) p_i$. Then, the 1-D signature is converted into a 2-D signature, W_2, using zigzag scan. Finally, the watermark sequence w_1 is produced by applying the IDCT to W_2: $\mathrm{w}_1 = \mathrm{IDCT}(\mathrm{W}_2)$.

Here, it should be noted that in this procedure we choose to transform the watermark signature to fit the face region instead of embedding the watermark into the normalized face region. This has the advantage that it avoids any distortion which might otherwise have incurred to the original face region. Another remark is that the masking step (i.e., discarding the part of the watermark signature outside the support of the normalized face region) is for the ease of implementation. It will not weaken the correlation property of the watermark signature, because the normalized face region is simply zero outside its support.

Now, the watermark extraction is performed according to the following procedure:

a) The face detection procedure is applied in order to provide face regions of the input image,
b) The normalization procedure is applied so as to obtain the normalized face regions,
c) the watermark patterns are re-generated using the same key and following the same procedure as during watermark embedding,
d) DCT is applied to the normalized face region (step a),
e) DCT coefficients are converted into a 1-D vector, denoted as c_w, through inverse zigzag scan and

f) the watermark message is decoded bit-by-bit using a correlation detector: $m_i' = \begin{cases} 1, & corr(cw, pi) > 0 \\ 0, & otherwise \end{cases}$

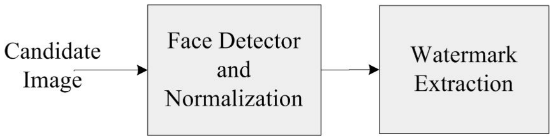

Figure 3: Watermark Extraction Method

5. EXPERIMENTAL RESULTS

In this section, the performance of the proposed watermarking scheme is tested under various attacks. A set of test images has been used for this purpose that contain human faces. In the following experimental results, the face region of an image (Fig. 1c) has been successfully extracted using the proposed face detector module. Then, the watermark embedding method is applied to this region, providing the final watermarked face region.

Afterwards, the watermarked face region undergoes a variety of geometric distortions and common signal processing attacks. In each case, the watermark extraction method is activated in order to recover the embedded watermark from each one of the distorted face regions.

Table 1 depicts the test results that the proposed method achieves. In these experiments the bit-error rate (BER) of the decoded watermark is used as validation measure. In our case the BER is defined as the ratio between the number of incorrectly decoded bits and the total number of embedding bits.

Table I: BER values after several attacks

Attack		BER
Scaling	20%	0
	50%	0.0005
Rotation	-10	0
	10^o	0
Shearing (both direction)	0%, 5%	0
	5%, 5%	0.0003
Cropping	10%	0.004
Flipping		0
Median Filtering (3x3)		0.02
Gaussian Filtering (3x3)		0
JPEG compression	Q=20	0.001
	Q=10	0.0007

As we can see in Table 1, the BERs for all geometric attacks are very low. In addition, the proposed method works very well under filtering attacks or jpeg compression. On the other hand, the most common problem that moments-based watermarking methods face is when an overcropping attack is applied. In case of our method, however, this could be a problem only if a malicious user crops large parts of the face region. But having cropped parts of this region, the semantic content is destroyed, something that even the malicious user does not want.

6. CONCLUSION

In this paper, we propose a way to protect face regions of an image using an affine invariant watermarking scheme. Firstly, face regions are detected in the initial image. Afterwards, a normalization procedure, invariant to affine transform attacks, is applied on the extracted regions using a set of specific moment criteria. A spread spectrum-based DS-CDMA watermarking scheme is used in order to provide a multi bits watermark. Finally, a multi bits watermark embedding procedure is applied, to provide the watermarked face regions. This approach can be used as a public watermarking scheme, providing robustness to general affine geometric attacks.

7. REFERENCES

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

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

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

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

[5] M. Alghoniemy and A. H. Tewfik, “Geometric distortion correction through image normalization,” *Proceedings ICME 2000,* vol.3, pp. 1291-1294, 2000

[6] P. Bas, J-M Chassery, B. Macq, “Geometrically Invariant Watermarking Using Feature Points”, *IEEE Trans. on Image Processing*, vol 11, No 9, pp.1014-1028, 2002

[7] H. Wang, S-F. Chang “A highly efficient system for automatic region detection in MPEG video”, IEEE Trans. Circuits and System for Video Technology, vol. 7, pp. 615-628, 1997.

[8] N. Tsapatsoulis, Y Avrithis, S. Kollias “Facial Image Indexing in Multimedia Databases” in Pattern Analysis & Applications, vol. 4, pp. 93-107, 2001.

[9] I. Rothe, H. Susse, and K. Voss, “The method of normalization to determine invariants”, *IEEE Trans. on Pattern Analysis and Machine Intelligent*., vol. 18, no. 4, Apr. 1996.

[10] P. Dong, J. Brankov, N. Galatsanos, Y. Yang, F. Davoine, “Digital Watermarking Robust to Geometric Distortions”, *IEEE Trans. Image Processing*, vol.14, no.12, 2140-2150, Dec 2005

[11] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.