AN OBJECT BASED WATERMARKING SOLUTION FOR MPEG4 VIDEO AUTHENTICATION

Dajun He, Qibin Sun and Qi Tian

Media Engineering, Labs for Information Technology
21 Heng Mui Keng Terrace, Singapore, 119613
Email: {djhe, qibin, tian}@lit.a-star.edu.sg

ABSTRACT

This paper presents an object based watermarking solution for MPEG4 video authentication. The watermark is embedded in Discrete Fourier Transform (DFT) domain before MPEG4 encoding. Groups of DFT coefficients in the low-middle frequency band are selected for watermarking. The coefficients in every group are divided into two sub-groups based on a pre-defined pattern, and the energy relationship between these two sub-groups is used to hide the watermark. The experimental results show that our algorithm is robust against MPEG4 compression as well as object based video manipulations such as rotation, scaling and inaccurate segmentation.

1. INTRODUCTION

The object based MPEG4 standard is proving to be very attractive for various applications in the areas of the Internet, video editing and wireless communication because of its object-based nature. It usually reduces the bit rate in transmission and simplifies video editing in multimedia applications. For example, during editing the objects can be scaled, translated and rotated to meet specific requirements. In video transmission, it may only transmit objects, not background, in real time to meet dynamic network bandwidth requirements. On the other hand, such flexibilities also pose new challenges to multimedia security (e.g., content integrity protection) because the video object (VO) can be easily accessed, modified or copied from one video sequence to another. Therefore an object-based watermarking solution to protect the integrity of the video becomes very important.

Boulgouris [1] and Lu [2] embed the watermark into the Discrete Cosine Transform (DCT) domain. In order to solve the synchronization issue in watermark extraction, Boulgouris separately embedded another special pattern into the spatial domain. Lu solved this problem by adjusting a video object to align its major eigenvector and minor eigenvector with the canonical orientation. Lie [3] took re-padding and re-quantization, very common in MPEG4, into consideration by embedding the watermark using the relationships among three groups of selected DCT coefficients in the middle frequencies. Considering that watermarked DCT blocks on the mask boundary are very sensitive to mask or segmentation error, P. Bas [4] and Piva [5] choose to embed a watermark in the spatial domain and wavelet domain respectively before objects are encoded. P. Bas also used eigenvalue information of the object to solve the synchronization problem caused by rotation.

A content-based watermarking system for content integrity protection is illustrated in Figure 1. Firstly, the raw video is segmented into foreground video (objects) and background video. The watermark is securely generated by combining the features extracted from both the objects and the background. The watermark is then embedded into foreground objects so that a secure link between object and background is created. At the receiver site, integrity between the object and background can be verified by correlating two sets of features: one is the extracted watermark from the object and the other is the extracted feature from both received object and background. If the correlation value is higher than a presetting threshold, we can claim that the video content is authentic. The selection of invariant features for video object and background will be described in another paper. These features should be characteristic of the specific video object and robust to normal video processing. In this paper, we focus on proposing a robust watermarking solution by assuming that the features have been extracted, and a specified long bit string as the watermark has also been created.

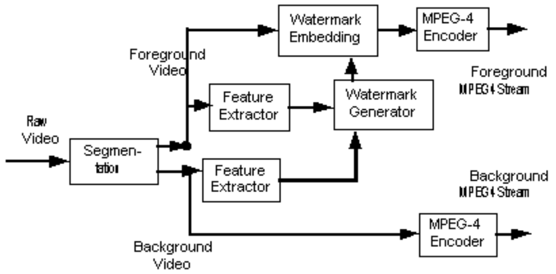


Figure 1. A brief watermarking system for MPEG4 video authentication

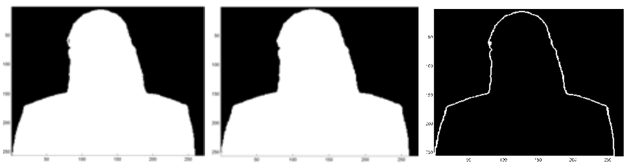
In this paper, we propose a new watermarking algorithm performed in the DFT domain. Experiments show that our algorithm is robust against scaling, rotation, translation, segmentation error as well as MPEG4 compression. The paper is organized as follows. Section 2 is an analytical description of our watermark embedding and extraction algorithm. Experimental results and conclusions are presented in Section 3 and 4 respectively.

2. PROPOSED WATERMARKING SCHEME

2.1. Problems and their countermeasures

A good watermarking system design should be application dependent. Since our application is for the purpose of content authentication, the basic requirement is that the proposed solution should be able to extract watermark messages without accessing original content. Some bit errors are allowed as long as an error correction coding scheme is still workable with a reasonable payload.

As mentioned in the introduction, in order not to constrain existing MPEG4 advantages on object access and manipulations, an object-based watermarking solution should be robust against rotation, scaling and translation. However, in addition to all robustness requirements listed above, we also take segmentation or object mask error in to account in this paper as required by real applications. Figure 2 illustrates the mask error in object manipulations.



(a) (b) (c)

Figure 2. Mask errors caused from object manipulations

Figure 2(a) is the original object mask. Figure 2(b) is the new object mask after rotating and scaling the original object and then rotating and scaling it back again (the mask has to follow the same operation for correct display). Figure 2(c) is their difference between Figure 2(a) and 2(b). In such a case, the watermark embedded in those blocks located at the mask boundary will be either destroyed or lost. This situation motivates us to explore embedding watermark in the DFT domain because the magnitudes of the DFT coefficients will not have significant changes caused by small segmentation or mask errors.

To develop an algorithm robust against scaling, we shall study the DFT coefficients before and after scaling. Figure 3 shows the energy distributions of DFT coefficients with different scales. We found that the energy of DFT coefficients shows great fluctuation in high frequencies; even some middle frequency information is lost if the scaling factor is around 0.5. Note that compression can also be considered as low-pass filtering. As illustrated in Figure 4, the whole DFT domain is partitioned into 4 areas: low frequency (black), low-middle frequency (textual), middle-high frequency (light gray) and high frequency (dark gray). In order to balance between robustness and quality, the low-middle frequency area is selected for watermarking in our algorithm.

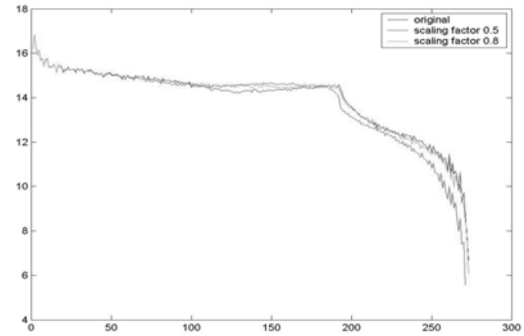


Figure 3. The Energy distributions of DFT coefficients in different scaling factor (0.5, 0.8, 1.0)

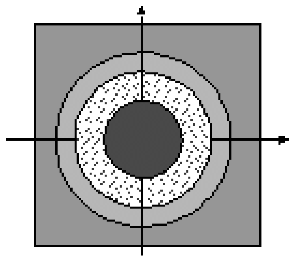


Figure 4. Four areas in DFT domain.

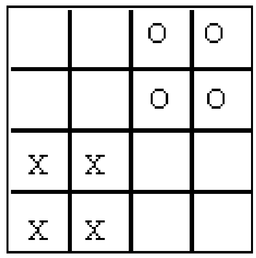


Figure 5. DFT coefficients grouping for watermarking

From Figure 3, we also found that the relationship between two DFT coefficients is not well preserved after object scaling, even in middle frequency band. However, their envelopes of energy distributions under different scales are still quite similar. Thus, we attempt to exploit the relationship between two groups of DFT coefficients to embed the watermark rather than two DFT coefficients, as shown in Figure 5.

We also use PCA (Principle Component Analysis) to solve the watermark synchronization problem, which is the same as Lu's method [2]: The eigenvector of the video object is calculated first. Then, the video object is adjusted to align its mask's maximum eigenvector with the X-axis in two-dimensional space before the watermark is embedded or retrieved. However, such embedding and extraction will involve many object rotation operations, which will also cause some variances in DFT coefficients. Furthermore, such embedding and extraction also cause mask errors which result in an inaccurate PCA calculation. In other words, the alignment parameters used for watermark embedding and extraction could be slightly different, although ideally they should be the same. We select the grouping pattern shown in Figure 5 for watermarking in order to alleviate such an incurred distortion. (In Figure 5, its bottom-left corner is towards to the origin of DFT domain). This pattern is not very sensitive to rotation errors. More detailed watermark embedding and extraction procedures are given in the next two subsections.

2.2. Watermark embedding

The embedding procedure is described as the follows.

1. Adjust the video object to align its mask's maximum eigenvector with the X-axis in two-dimensional space.
2. The size of the object is expanded into a pre-defined size by padding. The size is selected in a way such as 384x384, 256x256, 192x192, ... , based on the size of original object by taking its possible aligned size into account (e.g., the size of rotated object is larger than the original object size). In addition, the padding can use either zero padding or "mean value" padding.
3. For the expanded object image, we compute the DFT coefficients. Shift the DC coefficient to the center of DFT domain.
4. In the low-middle frequency area, we randomly select DFT coefficients group. As shown in Figure 5, eight DFT coefficients in each 4x4 group are classified into two sub-groups indicated as "o" and "x", we denote them as SG1 and SG2 respectively. The embedding procedure is as following:
   a. Compute the energy for SG1 and SG2.

$$E_{1,2} = \sum_{SG1,SG2} \| F(u,v) \| \qquad (1)$$

   $F(u,v)$ is DFT coefficient value at frequency (*u, v*).

   b. Calculate the adaptive threshold for each group

$$TH = \alpha * \left(\frac{f}{M/2} \right)^{\beta} \qquad (2)$$

   where $f = \sqrt{u^2 + v^2}$ stands for the frequency at point (*u, v*), *M* is the padded object image size (e.g., 384 or 256), and α, β are two constants estimated from statistical distribution of all DFT coefficients of this object image sequence. For instance, we could derive $\alpha = 30000, \beta = 1.8$ for the "Akiyo" test video. Note *TH* plays an very important role in balancing the watermarked video quality and watermark robustness.

   c. In the case that the watermark bit is "1":

   Step 1: If $E_1 - E_2 > 0$, then $MTH = (2/3)TH$ else $MTH = (4/3)TH$ (3)

   This threshold modification can further improve the watermark detection while maintaining the video quality.

   Step 2:

   If E_1 and E_2 satisfy equation (4), skip to step e). Otherwise, E_1 and E_2 are modified iteratively until equation (4) is satisfied.

$$E_1 - E_2 > MTH \qquad (4)$$

$$E_1 = E_1 + \Delta, \; E_2 = \max(E_2 - \Delta, 0) \qquad (5)$$

   where Δ is the step size.

   Step 3:

   Based on the new modified energy E_1 and E_2, modify the magnitude of each DFT coefficient individually in SG1 and SG2 while keeping their phases unchanged. The modification is related to each coefficient's original magnitude.

   d. In the case that the watermark bit is "0": exchange E_1 and E_2. The remaining procedure is the same as (c).

   e. For each watermarked coefficient, its symmetric coefficient with reference to the DC component of DFT should also be modified to ensure the watermarked video pixel has real value.
5. The watermarked image is acquired using IDFT. This image is rotated back to it original orientation. Then, a video object is extracted again using the method in MPEG4 standard[6]. This extracted video object can be used as the input of MPEG4 encoder.

2.3. Detection Scheme

Watermark detection is the reverse procedure of watermark embedding. The received video object is scaled back to its original resolution first. Then, we follow the described embedding procedure to calculate E_1 and E_2 for each sub-group are calculated. Finally, the watermark bit is extracted based on the following criterion:

if $E_1 - E_2 > 0$, watermark bit is '1'
else watermark bit is '0'

3. EXPERIMENTS

We have tested our algorithm on 2 MPEG4 object based video sequences: "Akiyo" and "Attack Akiyo". "Attack Akiyo" is created for the purpose of our authentication system performance evaluation (Figure 1). Figure 6 shows one original image and its watermarked version. For each object frame, we embedded 195 bits. We took 200 frames for testing. Figure 7 shows the results of watermarked video quality in terms of PSNR measure. Remember that in our embedding procedure, we need to estimate the orientation of the object first and then align it before embedding. After embedding, we also need to transform the watermarked object back to its original location. Hence, we present two PSNR results in Figure 7. The solid line measured the PSNR before and after watermarking with only watermarking distortion, while dotted line measured the PSNR between original object video and watermarked object video (not only watermarking distortion but also alignment distortion).


(a)


(b)

Figure 6. (a) Original frame (b) Watermarked frame

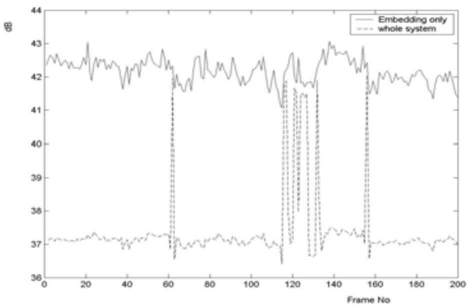


Figure 7. PSNR results of watermarked video

The robustness evaluation was performed on different scales, rotations and MPEG4 compressions. Figure 8 shows the robustness in terms of the ratio of correct extracted bits among a total of 195 watermark bits. We found that the correct ratios are all above 0.8. The actual information bits can be acquired correctly by using error correction coding.

4. CONCLUSIONS

In this paper, we have proposed a new watermarking solution for authentication of object-based video in the DFT domain. Before watermark embedding and retrieval, the eigenvector of the object mask is used to adjust the object's orientation. The watermark is embedded in a set of DFT coefficient groups that are randomly selected and located in the low-middle frequency area. Experimental results further demonstrated that the proposed solution is robust during MPEG4 compression and normal MPEG4 object manipulations such as scaling, rotation, as well as segmentation and mask errors.

REFERENCES

[1] Boulgouris N.V., Koravos F.D., Strintzis M.G., "Self-synchronizing watermark detection for MPEG-4 objects", *ICECS01*, Vol.3, Page(s): 1371 –1374, 2001

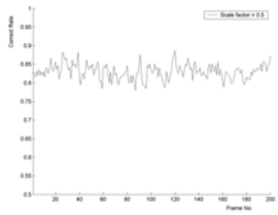
[2] Chun-Shien Lu; Hong-Yuan Mark Liao, "Video object-based watermarking: a rotation and flipping resilient scheme", *Processing of ICIP01,* Vol.2, Page(s): 483 –486, 2001

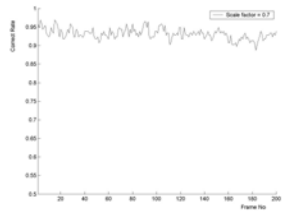
[3] Wen-Nung Lie; Guo-Shiang Lin; Ta-Chun Wang, "Digital watermarking for object-based compressed video", *ISCAS*01, Vol. 2, Page(s): 49 –52, 2001

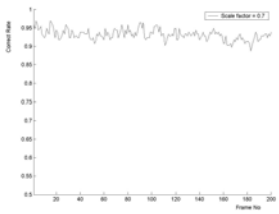
[4] P. Bas, and B. Macq, " A new Video-Object Watermarking Scheme Robust to Object Manipulation", *Proc. of ICIP01*, Tessaloniki, Greece, 2001

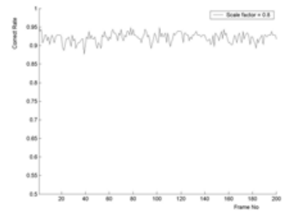
[5] Piva, R. Caldelli, A. De 2001Rosa, " A DWT-based object watermarking system for MPEG-4 video streams", *Proceeding of ICIP00,* Vancouver, Canada, September 10-13,2000.

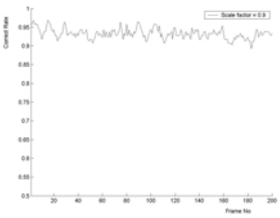
[6]. ISO/IEC 14496-2, "Information technology – Coding of audio-visual objects, Part 2"

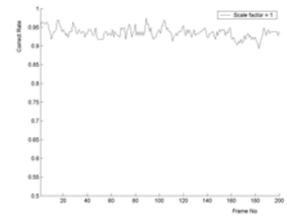

(a) Scaling 0.5

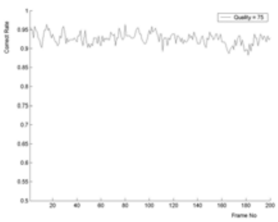

(b) Scaling 0.6


(c) Scaling 0.7


(d) Scaling 0.8


(e) Scaling 0.9


(f) No Scaling


(g) MPEG4 compression

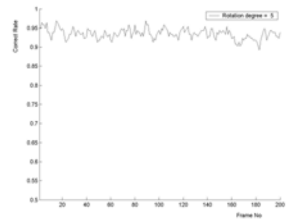

(h) Rotation: 5^o

Figure 8. Robustness tests.
Horizontal: Frame number 1-200 with the interval 20
Vertical: Ratio of correct extracted bits 0.5-1.0 with the interval 0.05