Estimation and Watermarking of Motion Parameters in Model Based Image Coding

Min Chul Park
Systems Technology Division,
Korea Institute of Science and Technology,
P.O.BOX131 Cheongryang, Seoul 130-650, Korea
Tel. +82-2-958-6761, Fax.: +82-2-958-5609
e-mail : minchul@kist.re.kr

Abstract: In order to achieve an advanced human-computer interface system, it is necessary to analyze and synthesize facial motions just as they are in an interactive way, and to protect them from unwanted use and/or illegal use for their privacy, various uses in applications and the costs of obtaining motion parameters. To estimate facial motion, a method of using skin color distribution, luminance, and geometrical information of a face is employed. Digital watermarks are embedded into facial motion parameters and then these parameters are scrambled so that it cannot be understood. Experimental results show high accuracy and efficiency of the proposed estimation method and the usefulness of the proposed watermarking method.

1. Introduction

When most of the coding methods that have been based on waveform coding, the model-based coding method was described as a new coding scheme in which it described image content in a structural way [1]. In 3D model-based approaches, real facial motion in facial image sequences has been principally studied for the purpose of compressing video signals. 3D model-based image coding makes use of the 3D properties of the scene and of the 3D face model based on the analysis and synthesis techniques [1-4]. Some interesting applications, such as speech driven image animation with the MPEG-4 visual standard supports, and virtual space teleconferencing have been proposed [5-6]. In order to achieve an advanced human-computer interface system, it is necessary to analyze and synthesize facial motions just as they are in an interactive way and to protect them from unwanted use and/or illegal use [7-10]. To estimate facial motion, a method of using skin color distribution, luminance, and geometrical information of a face is employed. Digital watermarks are embedded into facial motion parameters and then these parameters are scrambled so that it cannot be understood. This paper suggests an efficient method to estimate facial lip motion as a representative facial motion and a watermarking method to prove copyright when estimated parameters, i.e., animation parameters, are used without any approval in advance.

2. Advanced Model Based Image Coding

In model based image coding, images were considered as having structural features such as contours and regions. These image features have been exploited to encode images at very low bit rates, while retaining enough visible structures in the reconstruction so as to maintain a certain level of quality. Structural model based image approach used 3D structural models of the scene. The coder analyzes input images and the decoder generates output images using the shared model [1-3]. Figure 1 shows a general description of mode based image coding scheme.

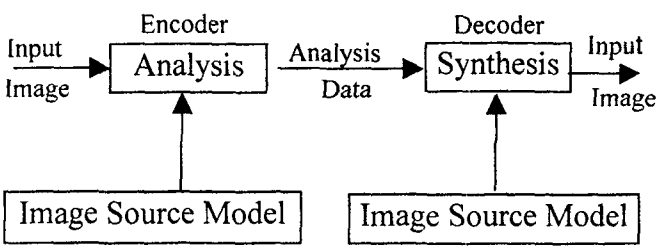


Figure 1: General description of Model Based Image Coding Scheme

There are 2D and 3D approaches in model based image coding schemes. In 3D approaches, real facial motion in face image sequence has been principally studied for the purpose of compressing video signals. 3D model based coding makes use of the 3D properties of the scene and a 3D face model based on the analysis and synthesis technique. However, this typical model based image coding produces synthetic images in 2D space at the receiver terminals. It would be more effective and interesting communication if the receiver terminals e.g., personal computer can produce synthetic images in 3D space while keeping low transmission bit rates. When facial motion in a (virtual) scene is synthesized or animated, real facial motion parameters are used to improve the reality and amusement value respectively. Meanwhile it is necessary to protect these images from unwanted use and copyright infringement because these digital data are easy to edit and copying keeping high quality. To solve these problems, the advanced scheme of model based image coding shown in figure 2 was proposed [4]. The typical scheme of model based image coding was extended by interactive control of a face object and digital watermarking.

3. Estimation of Facial Motion Parameters

Lip movement of a facial motion is estimated by extracting some feature points of the mouth shape. To extract some feature points of the mouth face region is roughly estimated firstly from a given image using the distribution of the hue

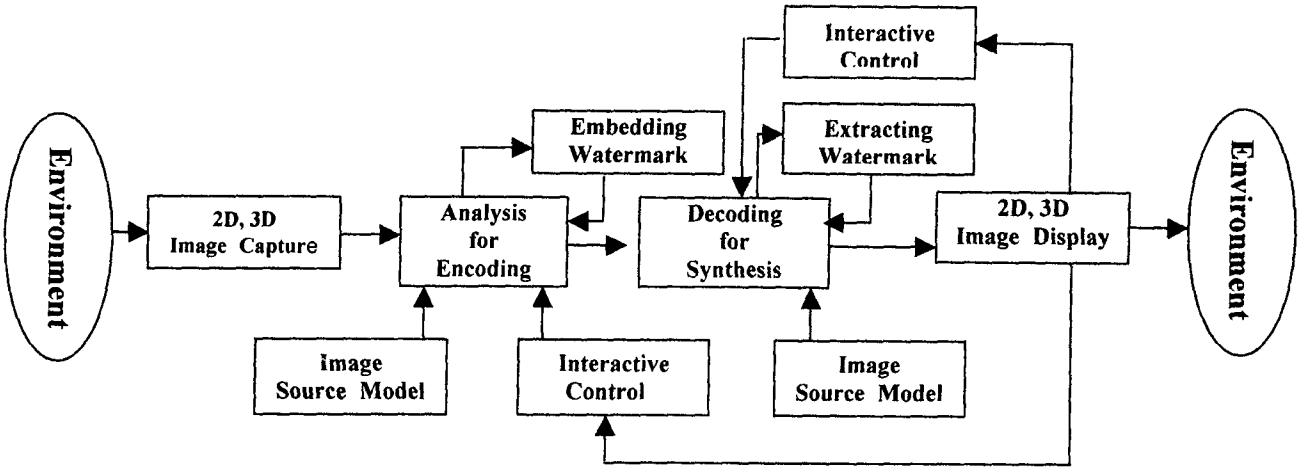
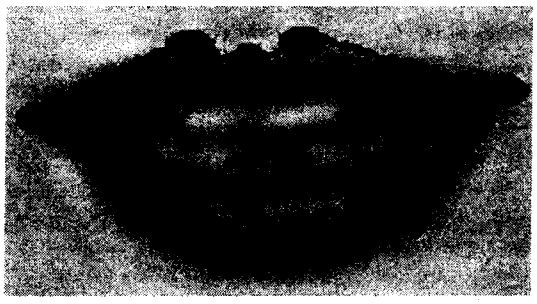


Figure 2: Description of the proposed advanced model based image coding scheme

components of human skin color and luminance. Then, detection of the mouth region is carried out to estimate some feature points, which are needed for automatic model fitting of model based image coding and MPEG-4 face animation, to make the mouth shape in a natural way. The most difficult process of the estimation of the mouth shape is to determine the outline of the region because the colors of the teeth and tongue are similar with those of reflected area of the mouth (i.e., almost white) and lip respectively. The decision value of the outline of the mouth usually depends on the input face images. In this experiment, this value is found out heuristically over the 299 frames.

4. Automatic Wireframe Model Fitting to the Mouth

Feature points, which are needed for the automatic model fitting, are estimated from the segmented mouth region. 8 mouth feature points, which are defined in MPEG-4 face animation, shown in figure 3 as (2,2), (2,3), (8,1), (8,2), (8,3), (8,4), (8,9), (8,10) are used in this experiment. The positions of these 8 feature points are estimated at each frame. Some other feature points of the mouth are generated based on the positions of the estimated 8 feature points.



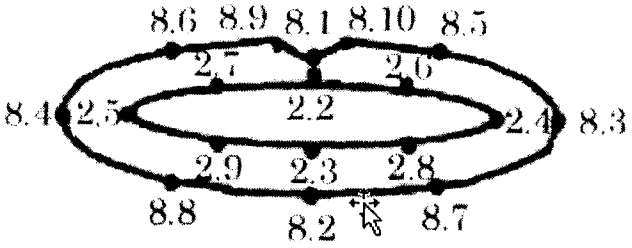


Figure 3: Mouth feature points employed for facial motion in the proposed method (left) and animation parameters in MPEG-4 (right).

5. Watermarking of Facial Motion Parameters

A watermark is embedded into feature points shown in figure 3. When 8 feature points are used as targets at each frame 16 different digit number (0 to 9) is available as watermark information because a float type feature point consists of x and y coordinates. Figure 4 shows an example of a coordinate and related bit positions where watermarks are embedded. Two different embedding methods are presented. One is embedding a digit number as a watermark by adding to a coordinate and the other is by subtracting a digit number as a watermark from coordinate. These algorithms are developed in mind detecting watermarks without watermarks in hand and have a revised and developed form of algorithm from [11]. An adding algorithm is described in the next program. (A) in Figure 4 show a float type number of a coordinate and (B) describes a watermarked float type number of the coordinate.

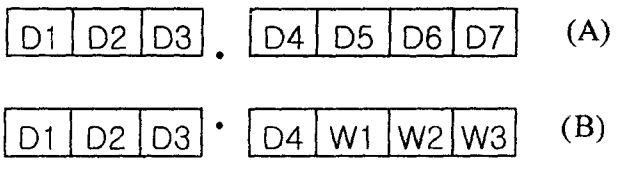


Figure 4: An example of embedding a digit number into a float type coordinate

Though there are several combinations of choosing bit positions D3 and D4 in (A), and W1 and W2 are chosen in the proposed method. D4 to D7 describe the digit numbers below a decimal point. W1 to W3 are the digit numbers that are related with watermarking. In the proposed method, D3 and D4 from an original number are used for embedding a digit number as watermarks are W1 and W2 describe the embedded digit numbers which are the changes from the original D5 and D6. More combinations can be made by selecting different places of numbers, as an example, D2 instead of D3 or D4 instead of D3. WM describes a digit number as a watermark. Watermarks are extracted by checking out W2 and W1 and using the adding equation " *sum* = D3+D4+WM". When watermarks WM are extracted WM are calculated by "WM=*sum*-D3-D4". A subtracting algorithm is described as the following. The detection of watermarks process is also carried out by checking out W2 and W1 and using the adding equation "*sum*=D4+D3-WM". When watermarks WM are extracted WM are calculated by "WM=D3+D4-*sum*".

{program: an adding algorithm }

```
int WM : float input, sum;
sum=D3+D4+WM;
if(sum1>0 && sum1<10) {
    W1=sum;
    W2=0;
}
else if(sum>=10 && sum<20) {
    if(sum==10) {
        W1=0;
W2=1;
      }
      else {
          W1=1;
          W2=sum-10;
}
else(sum>=20) {
    if(sum==20) {
        W1=0;
        W2=2;
    }
    else {
        W1=2;
        W2=sum-20;
    }
}
```

{program: a subtracting algorithm}

```
int WM; float input, sum;
sum=D3+D4-WM;
if(sum>=10) {
    if(sum==10) {
          W1=0;
          W2=1;
}
else {
          W1=1;
        W2=sum-10;
      }
else if(sum>=0 && sum<10) {
      W1=sum;
      W2=0;
}
else(sum<0) {
      W1=2;
      W2=sum+10;
}
```

6. Experimental Results

Facial feature points are used as targets for data hiding because the positions of facial feature points are moved by facial animation parameters. Facial parts group these feature points, and the mouthpart group among the several groups is selected in the experiment. As targets, mouth feature point in Akiyo sequence of 299 frames is used. Month feature points are estimated using the proposed segmentation algorithm at each frame. A watermark is embedded to 299 frames of 8 feature points. When 8 feature points are used as targets at each frame 16 different decimal numbers can be embedded as a watermark because one feature point consists of x and y coordinates. A long length watermark is also available as targets by using several or tens of frames. In this subsection, a short watermark that includes characters and numbers just like a password for computer-login is used. As a watermark "MCPARK7123" which is a combination of my name and student ID number is used. Character code "MCPARK" is changed into " 77 67 80 65 82 75" as numbers and "7123" is added as decimal numbers. The embedding algorithm is based on the algorithm mentioned in the previous subsection. Figure 5 shows the experimental results. Wireframe models on the right side image represent the estimated result and those on the left side represent watermarked wireframe models. The embedded watermark is completely detected at each frame. Mouth shapes in the left side are original, i.e., estimated mouth shapes, and ones in the rightside have watermarked shapes.

7. Conclusion

The proposed estimation method is tested for one of the standard image sequence known as Akiyo image sequence. The fitted wireframe models match well enough to describe the mouth shape as shown in figure 5. Experimental results show the usefulness and efficiency in its approach. Watermarks are embedded into the motion parameters invisibly and detected with the quality of good enough to prove the ownership and to achieve the goal of protection using the proposed watermarking method. As for further works, techniques using self error-correcting method to improve correlation ratio and the techniques embedding watermarks which look as freckles or wrinkles on the facial images are considered.

Figure 5: Original mouth shape and watermarked one are synthesized using wire frame model.

8. References

[1] K. Aizawa and T. S. Huang, "Model-Based Image Coding: Advanced Video CodingTechniques for Very Low Bit-Rate Applications," IEEE Proc., vol. 83, no. 2, pp. 259--271, Feb.1995.

[2] C. S. Choi, K. Aizawa, H. Harashima and T. Takebe, "Analysis and Synthesis of FacialImage Sequences in Model-Based Image Coding," IEEE Trans. CSVT, vol. 4, no. 3, pp. 257-275, June 1994.

[3] H. Li, A. Lundmark, and R. Forchheimer, "Image Sequence Coding at Very Low Bitrates: A Review," IEEE Trans. on Image Processing,} vol. 3, no 5, pp. 589--609, Sept. 1994.

[4] M. C. Park, N. P. Chandrasiri, T. Naemura, M. Kaneko. and H. Harashima, "Analysis and Synthesis of 3D Face mage Sequence Based on Model Based Image Coding Scheme with the Internet Browser," IEEE Intl. Conf. on mage Processing, vol. 4, pp.207-211., Oct.1999.

5] Jorn Ostermann, "Methodologies used for evaluation of ideo tools and algorithms in MPEG-4," Signal roc.:Image Communi., vol. 9, no.4, pp.343-365, May 1997.

6] Peter K. Doenges, Tolga K. Capin, F. Lavagetto, J.)stermann, Igor S. Pandzic, Eric D. Petajan, "MPEG-4," ignal Proc.:Image Communi., vol. 9, no.4, pp.433-463, lay 1997.

[7] A. H. Tewfik, M. Swanson, "Data Hiding for Multimedia Personalization, Interaction, and Protection," IEEE Signal Proc. Magazine,} pp.41-44., July 1997.

[8] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for Data Hiding," IMB Systems Journal, vol. 35, nos 3\&4, 1996.

[9] E. Praun, H. Hoppe, and A. Finkelstein, "Robust Mesh Watermarking," Proc. of ACM Siggraph 99 Conf., http://www.cs.princeton.edu/gfx/proj/meshwm/

[10] N. F. Johnson and S. Jajodia, "Exploring Steganography: Seeing the Unseen," Computer, pp. 26-34, 1998.

[11] M. C. Park, Y. K. Choi, T. Naemura, K. Aizawa and H. Harashima, "Data-hiding for Model Based Image Coding and MPEG-4 (Face Animation)," Intl. Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, pp.83-86, Sept. 1999.