VIDEO CODING SYSTEM BASED ON MODELS OF OBJECTS

Radim Javůrek

Department of Radioelectronics, Faculty of Electrical Engineering,
Czech Technical University in Prague
Technická 2, 166 27, Prague 6, Czech Republic
Phone: +420 22435 2113, Fax: +420 22435 5829, E-mail: javurer@fel.cvut.cz

Abstract

In this paper methods for model based coding of video sequences are presented. Furthermore composed system for model based image coding is described. This system has been composed in Matlab programming environment. Finally experimental results are shown.

1. Introduction

In model based coding method knowledge of scene is exploited to high compression efficiency achievement. For shape and colour description of all objects 3-D models are used. These models are coded and transmitted only on beginning of transmitted sequence. Scene description is done by means of movement and deformation description of scene objects. For this description only several parameters are needed therefore very low bit rates can be achieved.

One of the first systems for facial expression parameterisation described Ekman a Friesen [3] in seventies. However when texture mapping on model was introduced in eighties model based coding became popular research topic [1,2,6].

In MPEG-4 standard [5] procedure of parameters coding and decoding for facial animation of head and shoulders scenes is included. Head and shoulders scene is scene where only head and upper part of men's body in front of usually static background is included. A face model is described with set of Facial Definition Parameters (FDPs). Movement in the face is determined by set of animation parameters named Facial Animation Parameters (FAPs). Parameters are based on minimal face movement and are closed to the muscle movement. Technique for parameters coding is not defined in standard but there is defined syntax and semantic of resulting bit stream.

2. Model based coding

System for model based coding can be described in this way: Sequence from video camera is analysed in coder and in each frame individual objects are identified. On each object 3-D model is adapted. Each frame is thus 2-D projection of 3-D scene objects. Next animation parameters are obtained. These parameters describe 3-D movement and deformation of individual objects. Parameters are given by difference in positions of corresponding characteristic points in two frames. Hence animation parameters estimation is connected with characteristic points estimation. There are several basic approaches for animation parameters estimation: manual animation, video tracking, optical tracking, visual text to speech and automatic speech synchronization. That way set of parameters is obtained which together with 3-D model describes appearance of objects. These parameters are coded and transmitted. Set of parameters is small and then final bit rate is low, typically less than 1 kbit/s. Models of the objects are transmitted only at the beginning of the sequence provided they was not saved in decoder during previous sequence. In decoder parameters are decoded and utilized to adjustment of the model to the new shape and position. Original frame is approximated by displaying of all 3-D models in new positions and if needed they are deformed. Situation is displayed on figure 1.

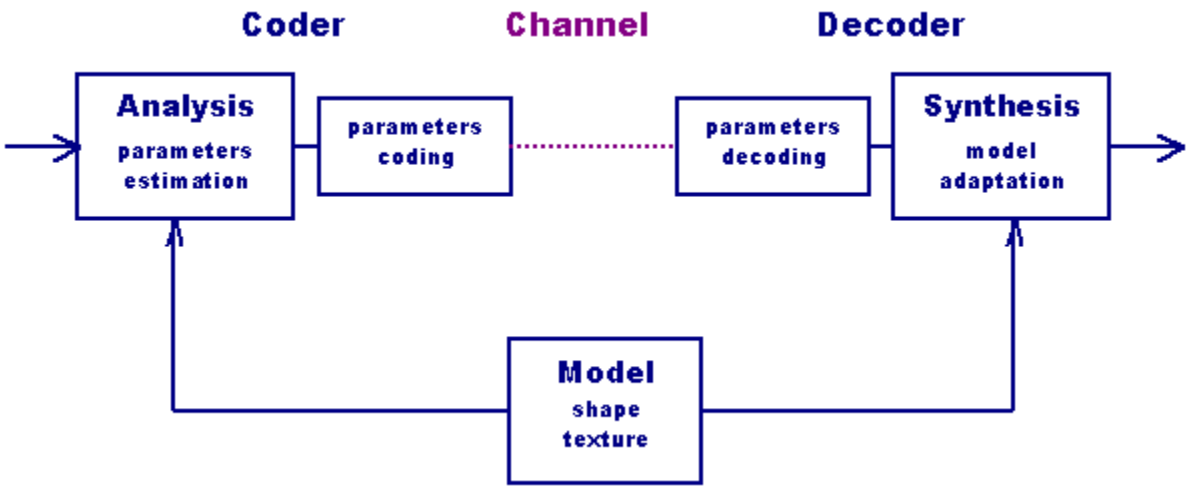

Fig. 1. *System for image coding based on models*

3. System description

Composed system for coding corresponds to the above mentioned scheme System is designed for head and shoulders scenes coding where only one moving object – men's head with upper part of body is presumed. Other objects form image background. Thus only one model of head is used in the system. This model consists of 113 vertexes and 68 triangles and is based on Candide model created by M. Rydfalk in 1987 [4]. For realistic looking results texture is mapped on the model. This texture is extracted from one frame of face in neutral state.

Methods used in individual parts of the system were selected with reference to final system complexity and its plasticity. For operation with model methods based on characteristic points are exploited. System was implemented in Matlab programming environment.

As an algorithm for face detection manual selection of small set of characteristic points from entire set of characteristic points defined in MPEG-4 were selected. Recently 13 characteristic points are used. This simple method is not usable in automatic systems for coding, but is very illuminating and as was mentioned above it is used as a referential method for other methods evaluation. Animation parameters are obtained from position difference of corresponding characteristic points. For coding of model and deformation and animation parameters coding scheme exploiting DCT, parameters prediction and variable code word length coding is used.

Procedure of parameters decoding in decoder is inverse to the coding procedure. For matching of the model to given face obtained characteristic points are exploited and model is matched by global motions (rotation, translation and scaling) of the model and local movement of individual vertexes of the model. Model animation is maintained by animation parameters obtained in the coder or by extracting animation parameters from ASCII text file. Inverse texture mapping algorithm applied on each triangle is used for texture mapping on 3-D model.

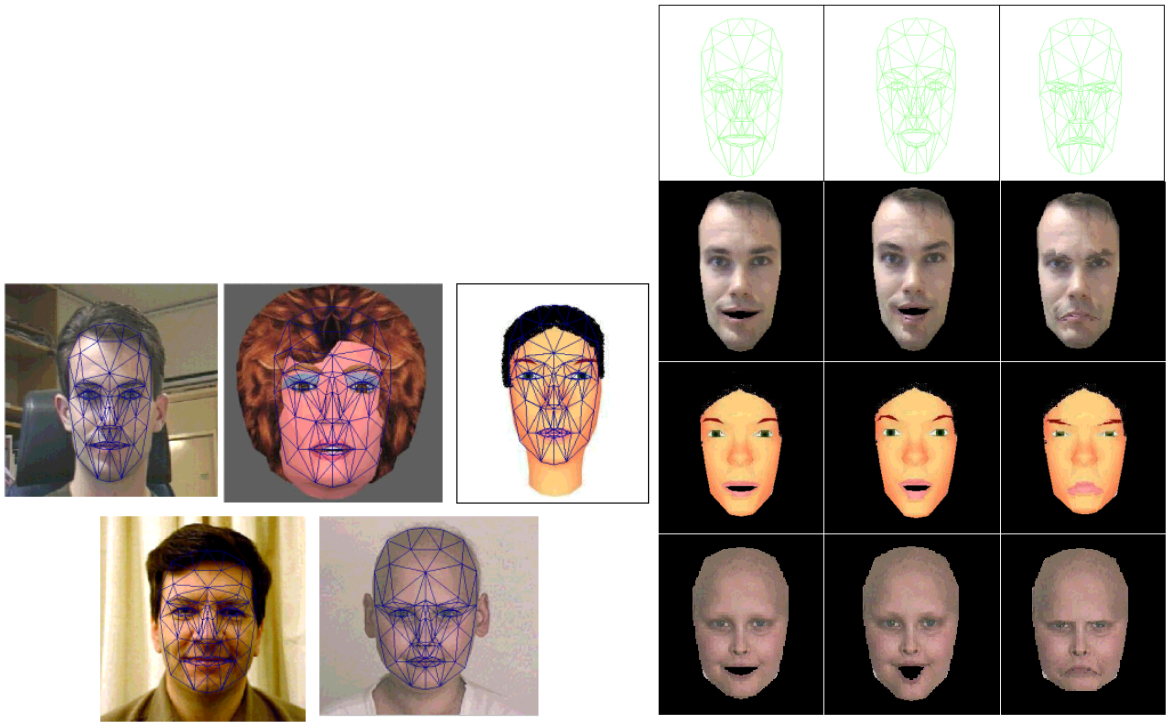

Fig. 2. *Model adaptation (left) and model animation (right)*

4. Conclusion

In this paper first an overview of model based coding methods is presented, then composed system for model based coding based on MPEG-4 standard is presented. Composed system is complete implementation of MPEG-4 facial animation codec. A new technique for generating visual speech from input text for Czech language is proposed. Video coding based on the models of object is efficient compression method and is suitable for real-time communication such as videoconferencing, games and e-commerce.

5. Acknowledgement

This work has been supported by the grant No. 102/02/0133 of the Grant Agency of Czech Republic. A part of this research work has been partially supported by the research program No. MSM 212300014.

References

[1] *Pandzic, I. S., Forchheimer, R., MPEG-4 Facial Animation: The Standard, Implementation and Applications, J. Wiley & sons, July 2002, 328 pages, ISBN: 0-470-84465-5*

[2] *Tekalp, A.M., Ostermann, J., Face and 2-D Mesh Animation in MPEG-4, Image Communication Journal, Tutorial Issue on MPEG-4 Standard, Elsevier, 2000*

[3] *Ekman, P., Friesen, W.V., Facial Action Coding System, Consulting Psychologists Press, Inc., Palo Alto, 1978*

[4] *Rydfalk, M., Candide: a parameterized face, Report No. LiTH-ISY-I-866, Dept of Electrical Engineering, Linköping University, Sweden, 1987*

[5] *MPEG Working Group on Visual, Coding of Audio-Visual Objects: Final Draft of International, Standard, Part 2 (Visual), ISO 14496-2, Doc. No. N2502 of JTC1/SC29/WG11 Atlantic City meeting, October 1998*

[6] *Eisert, P., Girod, B., Facial Expression Analysis for Model-Based Coding of Video Sequence, Picture Coding Symposium, Berlin, Sept. 1997, p. 33 – 38*