# Automatic Generation of 3D wireframes of human persons for video coding

Markus Kampmann
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Appelstraße 9A, 30167 Hannover, F.R.Germany
Phone: ++49–511–7625314, email: kampmann@tnt.uni–hannover.de
WWW: http://www.tnt.uni–hannover.de/~kampmann

***ABSTRACT:*** *In this contribution, an algorithm for the automatic generation of 3D wireframes of human persons in head–and–shoulder videophone sequences is proposed. As a first part of this algorithm, ellipsoidal volume models of human persons are generated by evaluation of the silhouettes and the location of the faces of the persons in the scene. As a second part of this algorithm, the volume models are approximated by 3D wireframes with incorporated pre–defined 3D face models. Applying the proposed algorithm to typical videophone sequences, an approximation of the 3D shape of human persons is possible.*

## 1. Introduction

For coding of videophone sequences at very low bit rates, model–based coding using 3D models has been proposed[1][2]. In this coding scheme, the real 3D objects in the scene are described by 3D model objects. A 3D model object is defined by motion, shape and texture parameters. The shape of a 3D model object is represented by a 3D wireframe. At the transmitter, the parameters of each model object are estimated. These parameters are coded and transmitted to the receiver. Using computer–graphic methods, model images are synthesized from the transmitted parameters at the receiver. In typical videophone sequences, head and shoulder of human persons appear in the scene. This prior knowledge can be used to enhance the model objects. However, only parts of the objects are modelled in most approaches. In [3], a pre–defined 3D face model is used for modelling the face, in [4], the head is modelled by a 3D head model. In [5], a complete 3D wireframe for head and shoulders of a person is generated using the silhouette and assuming ellipsoidal shape for the human body. This wireframe includes a pre–defined 3D face model. This algorithm allows only a rough approximation of the 3D shape of a human person. It is restricted to the simple face model CANDIDE [6] (Fig. 1) and to only one person in the scene.

In this contribution, an algorithm for the automatic generation of 3D wireframes of head and shoulders of human persons in videophone sequences is proposed which allows an improved approximation of the 3D shape of a human person. Therefore, the silhouettes and the location of the face of the persons in the scene are evaluated. The algorithm is not restricted to the face model CANDIDE, but allows the incorporation of any face model (Fig. 2) as well as other models of parts of the body. Furthermore, the generation of 3D wireframes for more than one person in the scene is possible.
The new algorithm consists of two processing steps. In the first step, a volume model of a human person is generated. In the second step, the volume model is approximated by a 3D wireframe.

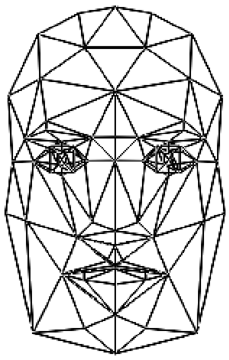

Fig. 1: Face model CANDIDE [6].

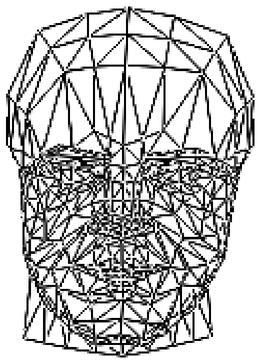

Fig. 2: Face model based on [8].

## 2. Generation of a volume model

The volume model of the human person is generated by evaluating the silhouette and the location of the face of the person assuming the anatomy of an average person. The silhouette is estimated automatically applying the algorithm in [5]. The location of the face is automatically estimated by adaptation of a face model to the person's face in the scene[9][10]. Applying the location of the person's face, the silhouette is divided into head, neck and upper part of the body (Fig. 3). Then, the head, the neck, the abdomen and the back are modelled by different ellipsoidal shapes assuming the anatomy of an average person. If two persons appear in the scene and these persons share a common silhouette, this silhouette is subdivided in the middle between the two located faces and an individual volume model is generated for each person.

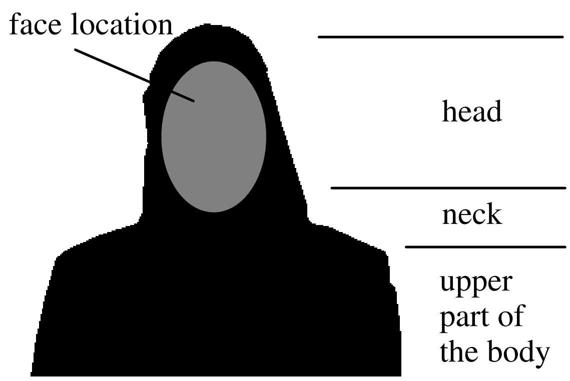


Fig. 3: Subdivision of a person's silhouette into head, neck and upper part of the body.

## 3. Generation of a 3D wireframe

After generation of the volume model, the volume model is approximated by a 3D wireframe. The generation of the 3D wireframe consists of three processing steps. In the first step, a pre–defined 3D face model is placed onto the surface of the volume model at the estimated location of the face. Then, triangles at the border of the volume model are generated which approximate the contour of the silhouette of the person. In the third step, a local growing triangulation algorithm is used[7] to connect the triangles of the face model and the generated triangles at the border of the volume model.

## 4. Experimental results

The described algorithm has been applied to the test sequence *Anis_Markus* and the MPEG4 test sequence *Akiyo* with a spatial resolution corresponding to CIF and a frame rate of 10Hz (Fig. 4). Fig. 5 shows the generated wireframes with the algorithm in [5] and with the proposed algorithm for the testsequence *Akiyo*. Whereas the algorithm in [5] uses the simple face model CANDIDE, the proposed algorithm allows the use of more complex face models (Fig.2). Compared with [5], the proposed algorithm allows an improved approximation of the 3D shape of a human person (see Fig. 5). Furthermore, the generation of 3D wireframes of two persons is possible (Fig. 6).


(a)

(b)

Fig. 4: Test sequences (CIF, 10Hz):
(a) *Akiyo* (7th frame),
(b) *Anis_Markus* (12th frame).

## 5. Conclusions

For automatic generation of 3D wireframes of human persons in head–and–shoulder videophone sequences, a new algorithm has been presented. First, a volume model is generated modelling the different parts of the human body by different ellipsoidal shapes. Therefore, the silhouettes and the location of the face of the persons in the scene are evaluated. Secondly, the volume model is approximated by a 3D wireframe with an incorporated 3D face model. Applying the presented algorithm to the videophone sequences *Akiyo* and *Anis_Markus* (CIF, 10Hz), an improved approximation of the 3D shape of the persons compared with [5] is possible. Due to the better modelling of the shape of a human person a coding gain can be expected by applying the proposed algorithm to a model–based coding scheme[5].

## 6. References


[1] K. Aizawa, T. Huang, "Model Based Image Coding: Advanced Video Coding Techniques for Very Low Bit–Rate Applications", *Proceedings of the IEEE*, Vol. 83, No. 2, February 1995, pp. 259–271.

[2] D. Pearson, "Developments in Model–Based Video Coding", *Proceedings of the IEEE*, Vol. 83, No. 6, June 1995, pp. 892–906.

[3] H. Li, P. Roivainen, R. Forchheimer, "3D Motion Estimation in Model–Based Facial Image Coding", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 6, June 1993, pp. 545–556.

[4] C. S. Choi, K. Aizawa, H. Harashima, T. Takebe, "Analysis and synthesis of facial image sequences in model–based image coding", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 4, June 1994, pp. 257–275.

[5] M. Kampmann, J. Ostermann, "Automatic Adaptation of a Face Model in a Layered Coder with an Object–based Analysis–Synthesis Layer and a Knowledge–based Layer", *Signal Processing: Image Communications*, Vol. 9, No. 3, March 1997, pp. 201–220.

[6] R. Rydfalk, "CANDIDE, A parameterised face", Internal Report Lith–ISY–I–0866, Linköping University, Linköping, Sweden, 1987.

[7] H. Busch, "Ein Verfahren zur Oberflächenmodellierung dreidimensionaler Objekte aus Kamerabildfolgen", Dissertation, Universität Hannover, 1992.

[8] J. Fischl, B. Miller, J. Robinson, "Parameter tracking in a muscle–based

analysis/synthesis coding system”, *Picture Coding Symposium* (PCS ’93), Lausanne, Switzerland, No. 2.3, March 1993.

[9] M. Kampmann, Ridha Farhoud, ”Precise face model adaptation for semantic coding of videophone sequences”, *Picture Coding Symposium* (PCS ’97), Berlin, Germany, September 1997.

[10] M. Kampmann, Geovanni Martinez ”Automatic adaptation of face models in videophone sequences with more than one person”, *International Workshop on Coding Techniques for Very Low Bit–rate Video* (VLBV97), Linkoeping, Sweden, July 1997.

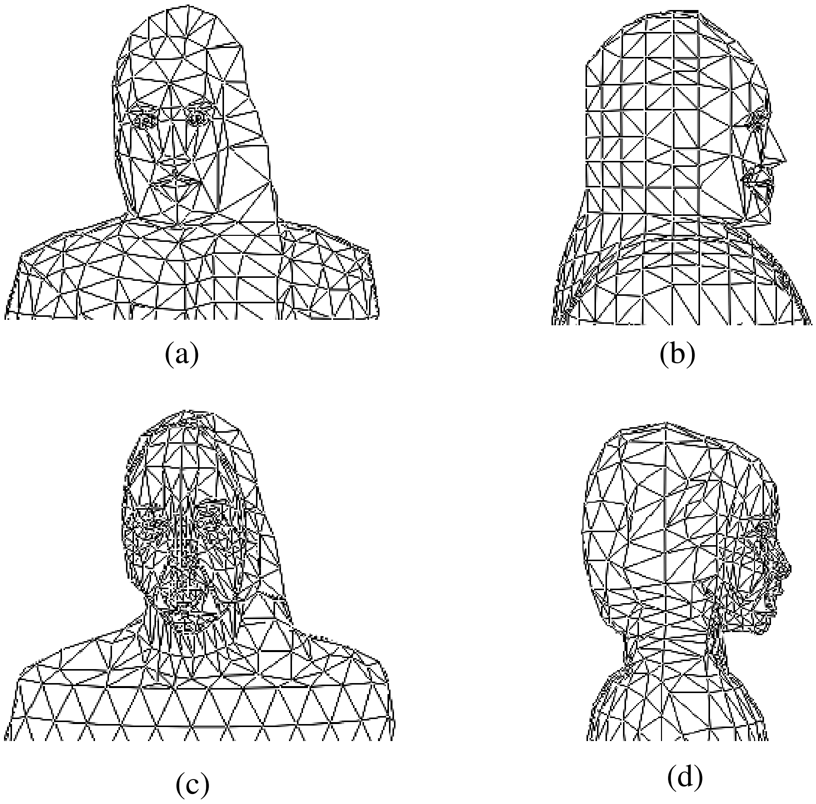


Fig. 5: Generated wireframes for *Akiyo* (7th frame) with the:
(a) algorithm in [5]: front view, (b) algorithm in [5]: side view,
(c) proposed algorithm: front view, (d) proposed algorithm: side view.

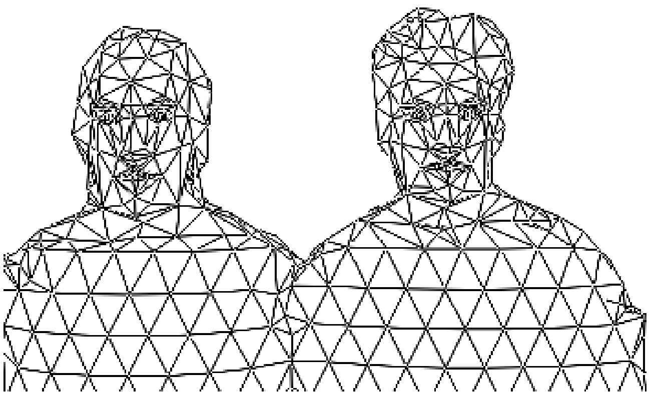

Fig. 6: Generated wireframe for *Anis_Markus* (12th frame) with the proposed algorithm