Animation of 3D Model of Human Head

Ján MIHALÍK, Viktor MICHALČIN

Laboratory of Digital Image Processing and Videocommunications, Dep. of Electronics and Multimedia Telecommunications, FEI TU Košice, Park Komenského 13, 041 20 Košice, Slovak Republic

Jan.Mihalik@tuke.sk, michalcin@vk.upjs.sk

Abstract. *The paper deals with the new algorithm of animation of 3D model of the human head in combination with its global motion. The designed algorithm is very fast and with low calculation requirements, because it does not need the synthesis of the input videosequence for estimation of the animation parameters as well as the parameters of global motion. The used 3D model Candide generates different expressions using its animation units which are controlled by the animation parameters. These ones are estimated on the basis of optical flow without the need of extracting of the feature points in the frames of the input videosequence because they are given by the selected vertices of the animation units of the calibrated 3D model Candide. The established multiple iterations inside the designed animation algorithm of 3D model of the human head between two successive frames significantly improved its accuracy above all for the large motion.*

Keywords

Animation, 3D model, human head, animation parameters, global motion, local motion, optical flow, estimation.

1. Introduction

MPEG-4 SNHC [1] uses the new methods of coding of the human head based on its modeling [2] and animation [3]. The animation enables the local motion of 3D model of the human head by using animation parameters including its global motion. The animation parameters can be generated synthetically or they are the result of analysis of the real human head. Except for the global motion as rotation and translation the human head in the input videosequence also has the local motion which is caused by a face mimic. The typical case of the local motion of the human head is opening of its mouth at generation of the speech. Then animation in the model based human head coding [4] generates by the local motion of 3D model the different expressions of the synthesized human head for verbal and nonverbal communication. The method of coding is an important part of the standard MPEG-4 SNHC, which allows such applications as the advanced communications using the cloned or virtual human heads.

The local motion or mimic of 3D model of the human head can be classically expressed by the method of clip-and-paste [5] for which the structure of 3D model is not deformed. First the texture of a characteristic part like a mouth and eyes of the human face is cut out from the frames of the input videosequence. The local changes of the human head, which participate on animation, have to be expressed in regard to the texture. We can see the texture as an image of the part of the human face what means that the vector quantizer [6] can be used for its coding. The code book of the vector quantizer involves the templates of the part with different expressions. Then by using the code book we can generate the finite number of possible expressions of the part of the synthesized human head. More complex expressions not for a part but for the whole synthesized human head are generated by combination of several templates from the code books of the different parts which have to be translated to the decoder. Therefore the code books include the big number of templates, the method of clip-and-paste is not very good for low bit rate coding of the human head. In addition the expressions have to be pasted on the right places in the synthesized frames what is not simple solution. Even small inaccuracy of their pasting affects busy in the synthesized frames. A better solution from this point of view is mimic expression by animation of 3D model of the human head on the basis of feature extraction and tracking [7]. The important features like points, edges and contours of the real human head in the input videosequence can be extracted by the deformable templates [8] of its parts or by application of the principal component analysis [9] on the ones. The method of optical flow [10], [11] using the parameterized 3D model of the human head enables excluding of the need of extraction of the features for its animation.

2. Parameterized 3D Model Candide

Each vertex $\mathbf{M}=(h,v,r)^T$ of 3D model Candide [12] has the initial position in the model coordinate system (MCS) defined by the horizontal (h), vertical (v) and depth (r) cordinates. The coordinates of its all vertices are normalized on the maximum value 1. The advantages of 3D model Candide are its availability and mainly simplicity referred to the relative small number of vertices and polygons. Next its advantage is that almost all feature points of the human head defined in the standard MPEG-4 SNHC have the corresponding vertices in 3D model Candide.

3D model Candide is parameterized by the following equation

$$\mathbf{x} = \bar{\mathbf{x}} + u\mathbf{S} + w\mathbf{A} \tag{1}$$

where **x** is the 3*N* dimensional vector composed from 3D coordinates of its *N* vertices, $\bar{\mathbf{x}}$ is the 3D model in the initial position, **S** the matrix of shape units, **A** the matrix of animation units, *u* the shaping parameters, and *w* the animation parameters.

The number of columns of the matrices **S** and **A** depends on the number of shaping and animation unites, respectively. The difference between shaping and animation units is that the shaping units define deformation of the geometrical shape of 3D model Candide while the animation units define its local motion in dependence on a mimic of the real human head. For the global motion of 3D model Candide given by rotation and translation of its all vertices and also its scaling further 7 parameters are needed. Then eq. (1) changes as follows

$$\mathbf{x} = m\mathbf{R}\left(\bar{\mathbf{x}} + u\mathbf{S} + w\mathbf{A}\right) + \mathbf{T} \tag{2}$$

where *m* is a scale, $\mathbf{R} = \mathrm{R}(\Theta h, \Theta v, \Theta r)$ is the rotation matrix with Euler´s rotation angles Θh, Θv and Θr, and $\mathbf{T} = \mathrm{T}(th, tv, tr)$ is 3D translation vector

From eq. (2) it follows out that shaping, animation and global motion including scaling of 3D model Candide are described by the parametric vector

$$\mathbf{D} = [v, u, w]^T = [\Theta_h, \Theta_v, \Theta_r, t_h, t_v, t_r, m, u, w]^T \tag{3}$$

where **v** is a vector of the global motion parameters and scale.

A change of the geometrical shape of 3D model Candide can be carried out by using the shaping units (SU) which number is 14. Each SU is labeled by its name derived from the kind of geometrical shaping.

Animation of 3D model Candide allows interpretation of different local motion. In general the local motion can be realized by the animation units (AU) which number is 11. Each AU is identified by its name derived from the kind of local motion and is defined by a grope of vertices, which enable the selected animation. The vertices of AU have animation vectors, which express the maximum changes Δh, Δv, Δr of their coordinates from the initial positions in MCS. In Fig.1a,c for 3D model Candide the vertices are outlined together with their animation vectors for AU of "Raising upper lip" and "Dropping of jaw". The maximum changes as components of the animation vectors are shown in Tab.1 and 2 for both AU, respectively.

Vertex indices	Animation vectors		
	Δ*h*	**Δ*v***	**Δ*r***
7	**0,000000**	**0,086957**	**0,021739**
33	**0,000000**	**0,065217**	**0,021739**
66	**0,000000**	**0,065217**	**0,021739**
79	**0,000000**	**0,050000**	**0,021739**
80	**0,000000**	**0,050000**	**0,021739**
81	**0,000000**	**0,050000**	**0,021739**
82	**0,000000**	**0,050000**	**0,021739**
87	**0,000000**	**0,065217**	**0,021739**

Tab. 1. AU "Raising upper lip".

Vertex indices	Animation vectors		
	Δ*h*	**Δ*v***	**Δ*r***
40	**0,000000**	**-0,260000**	**-0,050000**
8	**0,000000**	**-0,260000**	**-0,050000**
9	**0,000000**	**-0,260000**	**-0,100000**
10	**0,000000**	**-0,130000**	**-0,150000**
32	**0,000000**	**-0,150000**	**-0,130000**
65	**0,000000**	**-0,150000**	**-0,130000**
83	**0,000000**	**-0,200000**	**-0,050000**
84	**0,000000**	**-0,200000**	**-0,050000**
85	**0,000000**	**-0,200000**	**-0,050000**
86	**0,000000**	**-0,20000**	**-0,050000**

Tab. 2. AU "Dropping of jaw".

Then for the coordinates of the animated vertex $\mathbf{M}' = (h', v', r')^T$ in MCS are valid

$$\begin{pmatrix} h' \\ v' \\ r' \end{pmatrix} = \begin{pmatrix} h \\ v \\ r \end{pmatrix} + \begin{pmatrix} \Delta h \\ \Delta v \\ \Delta r \end{pmatrix} w \tag{4}$$

where *w* is the animation parameter, which having the value from the interval $w \in <0,1>$, determines the quantity of their changes from the initial values (h,v,r). It means that each AU has one animation parameter by using which its all vertices are currently affected. In Fig.1b the animated 3D model Candide is shown by using AU of "Lifting upper lip" and in Fig.1d by AU of "Dropping of jaw" assumed the maximum value $w=1$ for both AU.

For animation of 3D model Candide it is possible to use several AU which animate its different parts. The result such a combination of two before-given AU is an integrated expression as it is shown in Fig.1e. In addition, two vertices of one AU can belong to the different animation vectors from which it follows out that one AU can interpret more expressions of the human head. The fact is proved by the simple example when the AU for shutting of the lid will be the same as the AU for its full closing but their animation vectors are different.

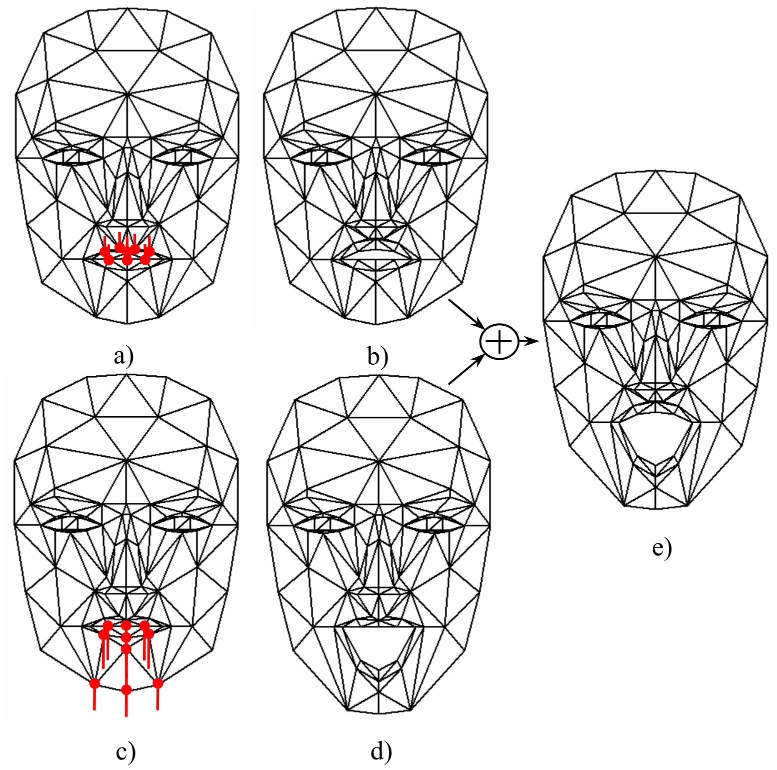


Fig.1. Animation of 3D model Candide by using AU "Raising upper lip" and "Dropping of jaw" a), c) the animation vectors, b), d) the animated 3D model by the separated AU e) the animated 3D model by combination of both AU.

3. Animation with Small Local Motion

The local motion of 3D model Candide is defined by AU, which is affected by the animation parameters $\overline{\mathbf{W}}=(w_1,...,w_n)$. Next we derive a formula for estimation of the animation parameter w for any AU assumed the small local motion between two successive frames. The first frame of the ones is a reference frame under which 3D model Candide and the camera are calibrated and the local motion is described by eq.(4). All vertices of AU from the point of view of estimation of w can be considered as the feature ones and their projections on the frame as the feature points. Then the coordinates of the animated feature vertex after their transformation to the camera coordinate system (CCS) [13] will be

$$x'=x\left[1+\frac{\Delta h}{x}w\right], \tag{5}$$

$$y'=y\left[1+\frac{\Delta v}{y}w\right], \tag{6}$$

$$z'=z\left[1-\frac{\Delta r}{z}w\right]. \tag{7}$$

If eq.(5) and (6) are divided by eq.(7) we get

$$\frac{x'}{z'}=\frac{x}{z}\frac{\left[1+\frac{\Delta h}{x}w\right]}{\left[1-\frac{\Delta r}{z}w\right]}, \tag{8}$$

$$\frac{y'}{z'}=\frac{y}{z}\frac{\left[1+\frac{\Delta v}{y}w\right]}{\left[1-\frac{\Delta r}{z}w\right]}. \tag{9}$$

Then after multiplying and by using the expression of perspective projection [13] for the components (u_i, u_j) of the vector of optical flow we have

$$u_j=\frac{\frac{J\Delta r+f_x\Delta h}{z}w}{1-\frac{\Delta r}{z}w}, \tag{10}$$

$$u_i=\frac{\frac{I\Delta r-f_y\Delta v}{z}w}{1-\frac{\Delta r}{z}w} \tag{11}$$

where $I=(i-i_0)$, $J=(j-j_0)$ are the centered coordinates of the feature point in the frame. Eq.(10) and (11) represent a non linear dependence of the components of the vector of optical flow on the animation parameter w. Assuming small local motion between two successive frames (w<<1) and the big distance d of the camera from 3D model Candide compared to its depth coordinates r for the denominator of eq. (10) and (11) it is valid

$$1-\frac{\Delta r}{z}w\cong 1 \tag{12}$$

where $z=d-r$. Afterwards from eq.(10) and (11) in regard to eq. (12) the linear dependences of the components (u_i, u_j) of the vector of optical flow on the animation parameter w are obtained

$$u_j = \frac{J\Delta r + f_x \Delta h}{d - r} w \, , \tag{13}$$

$$u_i = \frac{I\Delta r - f_y \Delta v}{d - r} w \, . \tag{14}$$

By substitution of eq.(13) and (14) into the optical flow equation [14] we get one linear equation for the feature vertex

$$\left(I_i \frac{I\Delta r - f_y \Delta v}{d - r} + I_j \frac{J\Delta r + f_x \Delta h}{d - r} \right) w = -\Delta I_t \, . \tag{15}$$

Then from eq.(15) the animation parameter w for the small local motion can be directly calculated. For better estimation of the animation parameter w it is needed to use more feature points above all for AU with a bigger number of the vertices of 3D model Candide. In matrix form eq.(15) for N feature vertices can be written as follows

$$\begin{pmatrix} e_{p1} \\ \vdots \\ e_{pN} \end{pmatrix} w = -\begin{pmatrix} \Delta I_t^{p1} \\ \vdots \\ \Delta I_t^{pN} \end{pmatrix} \tag{16}$$

where $e_{pk} = I_i \frac{I\Delta r_{pk} - f_y \Delta v_{pk}}{d - r_{pk}} + I_j \frac{J\Delta r_{pk} + f_x \Delta h_{pk}}{d - r_{pk}}$ and ΔI_t^{pk} are the members for the feature vertex pk, k=1,2,...N. On the basis of the least square method (LSM) a solution of eq.(16) for the animation parameter w of the small local motion will be

$$w = \frac{-\sum_{k=1}^{N} e_{pk} \Delta I_t^{pk}}{\sum_{k=1}^{N} e_{pk}^2} \, . \tag{17}$$

In eq.(16) the maximum number of the selected feature vertices is equal to the number of the vertices of AU. To achieve higher accuracy of estimation of the animation parameter w for the given AU it is suitable to take all its vertices.

4. Animation with Large Local Motion

In generality for estimation of the animation parameters of the large local motion it is possible to use a procedure in analogy with estimation of the parameters of the large global motion by using the multiple iterations [14]. Then calculation of the animation parameter for the feature vertex can be done by the following equation

$$\left(\hat{I}_i \frac{I\Delta r - f_y \Delta v}{d - r} + \hat{I}_j \frac{J\Delta r + f_x \Delta h}{d - r} \right) w = -\Delta \hat{I}_t \tag{18}$$

where the partial derivatives $\hat{I}_i = \partial I(i + \hat{u}_i, j + \hat{u}_j, t+1) / \partial i$, $\hat{I}_j = \partial I(i + \hat{u}_i, j + \hat{u}_j, t+1) / \partial j$ and the time difference $\Delta \hat{I}_t = I(i + \hat{u}_i, j + \hat{u}_j, t+1) - I(i, j, t) - \hat{I}_i \hat{u}_i - \hat{I}_j \hat{u}_j$. In the expressions the predicted components of the vector of optical flow are

$$\hat{u}_j = \frac{J\Delta r + f_x \Delta h}{d - r} \hat{w} \, , \tag{19}$$

$$\hat{u}_i = \frac{I\Delta r - f_y \Delta v}{d - r} \hat{w} \tag{20}$$

where $\hat{w}$ is the animation parameter from the previous frame as it is seen in Fig.2.

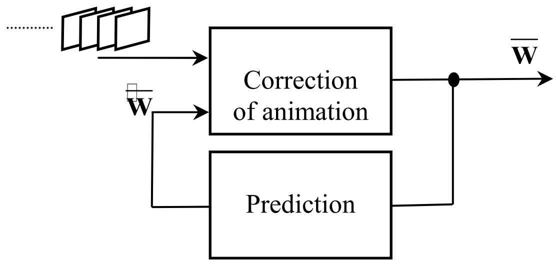


Fig.2. Estimation of the animation parameters of the large local motion.

In the case of using N feature vertices we get the linear system composed from eq.(18)

$$\begin{pmatrix} \hat{e}_{p1} \\ \vdots \\ \hat{e}_{pN} \end{pmatrix} w = -\begin{pmatrix} \Delta \hat{I}_t^{p1} \\ \vdots \\ \Delta \hat{I}_t^{pN} \end{pmatrix} \tag{21}$$

where $\hat{e}_{pk} = \hat{I}_i^{pk} \frac{I\Delta r_{pk} - f_y \Delta v_{pk}}{d - r_{pk}} + \hat{I}_j \frac{J\Delta r_{pk} + f_x \Delta h_{pk}}{d - r_{pk}}$ and $\Delta \hat{I}_t^{pk}$ are the members for the feature vertex pk, k=1,2,...N. Afterwards by using LSM a solution of eq. (21) for the animation parameter w of the large local motion will be

$$w = \frac{-\sum_{k=1}^{N} \hat{e}_{pk} \Delta \hat{I}_t^{pk}}{\sum_{k=1}^{N} \hat{e}_{pk}^2} \, . \tag{22}$$

The derived eq.(22) for estimation of the animation parameter of the large local motion is formally the same as eq.(17) valid only for the small local motion. In eq.(22) there is furthermore an information about the local motion from the previous frame which enable estimation of the animation parameter of the large local motion with less error. Then by application of the multiple iterations through the feedback of prediction in Fig.2 it is possible to increase the accuracy of estimation of the animation parameter w.

The result of estimation using the optical flow according to eq.(17) and (22) is the animation parameter w for AU which the given local motion realizes. The algorithm of estimation of the animation parameters of the local motion based on eq.(17) or (22) using the optical flow is very fast because it does not need extraction of the feature points in the frames. On the other side its calculation requirements depend just on the number N of the feature vertices.

5. Animation in Combination with Global Motion

The local motion of the human head in the real videosequence is often affected by its global motion and they create together the combination motion. Therefore for animation of 3D model of the human head we will use the designed algorithms not just for estimation of the animation parameters but also the parameters of the global motion [14]. In addition the estimation of the global motion parameters is carried out independently on the local motion what is done by selecting of the feature vertices without the local motion. On the other side it is not possible for estimation of the animation parameters because the selected feature vertices except for the local motion have the global motion too.

On the basis of the previous consideration about animation in combination with the global motion an estimator was designed which block diagram is in Fig.3. The first estimation of the global motion parameters is done from the input videosequence. Afterwards they are used for estimation of the animation parameters. Thereof it follows out that for estimation of the animation parameters the algorithm for the large local motion described by eq.(22) have to be used. In this case the global motion parameters together with the animation parameters from the previous frame are used for prediction of the components ($\hat{u}_i$, $\hat{u}_j$) of the vector of optical flow.

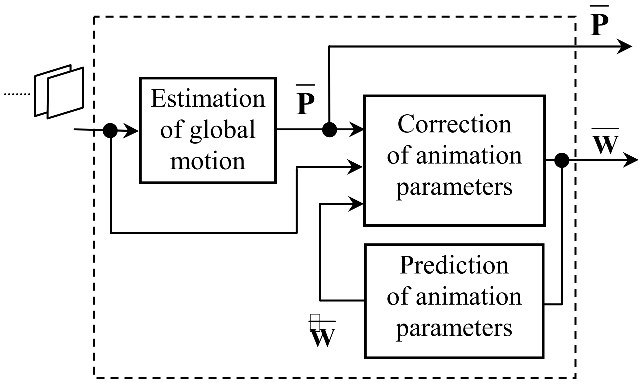


Fig.3. The block diagram of an estimator of the animation parameters with the global motion.

Next we derive the equations for estimation of the animation parameters in combination with the global motion between two successive frames. Assume that 3D model of the human head is calibrated according to the first one of the two frames and its vertex $\mathbf{M}=(h,v,r)^T$ in initial position in MCS. After animation with global motion its new coordinates $\mathbf{M}'=(h',v',r')^T$ will be calculated as follows

$$\begin{pmatrix} h' \\ v' \\ r' \end{pmatrix} = \underline{\mathbf{R}}\left[\begin{pmatrix} h \\ v \\ r \end{pmatrix} + \begin{pmatrix} \Delta h \\ \Delta v \\ \Delta r \end{pmatrix} w\right] + \overline{\mathbf{T}} = \underline{\mathbf{R}}\begin{pmatrix} h \\ v \\ r \end{pmatrix} + \overline{\mathbf{T}} + \begin{pmatrix} \Delta h^\circ \\ \Delta v^\circ \\ \Delta r^\circ \end{pmatrix} w \qquad 23)$$

where

$$\begin{pmatrix} \Delta h^\circ \\ \Delta v^\circ \\ \Delta r^\circ \end{pmatrix} = \underline{\mathbf{R}}\begin{pmatrix} \Delta h \\ \Delta v \\ \Delta r \end{pmatrix}. \qquad (24)$$

Then $\mathbf{M}'=(h',v',r')^T$ is the rotated, translated and animated vertex and (Δh, Δv, Δr) is the rotated animation vector belonging to the vertex **M**. From the definition of animation with global motion by eq. (23) it follows out that in this case the principle of superposition is not valid. It is caused by the fact that the animation vector of the vertex which except for the local motion has the global motion too is affected with its rotation matrix **R.** It has to be taken into account for estimation of the animation parameters. After transformation of the new coordinates $(h',v',r')^T$ of the rotated, translated and animated vertex **M**' from MCS into CCS and substitution for the rotation matrix **R** and the translation vector **T** we get

$$x' = x\left[1 + \frac{1}{x}\left(-\Theta_r y + \Theta_v (d-z) + t_h + \Delta h^\circ w\right)\right], \qquad (25)$$

$$y' = y\left[1 + \frac{1}{y}\left(\Theta_r x - \Theta_h (d-z) + t_v + \Delta v^\circ w\right)\right], \qquad (26)$$

$$z' = z\left[1 + \frac{1}{z}\left(\Theta_v x - \Theta_h y - t_r - \Delta r^\circ w\right)\right]. \qquad (27)$$

By application of the analogical derivation procedure as for the global or local motion it is possible from eq.(25), (26), (27) to derivate the linear formulas for the components of the vector of optical flow

$$u_j = \frac{J\Delta r^\circ + f_x \Delta h^\circ}{d-r} w - \frac{JI}{f_y}\Theta_h + \left(\frac{f_x r}{(d-r)} - \frac{J^2}{f_x}\right)\Theta_v + \frac{f_x I}{f_y}\Theta_r + \frac{f_x}{(d-r)} t_h + 0 t_v + \frac{J}{(d-r)} t_r = \frac{J\Delta r^\circ + f_x \Delta h^\circ}{d-r} w + \overline{\mathbf{V}}\overline{\mathbf{P}} \qquad (28)$$

$$u_i = \frac{I\Delta r^\circ - f_y \Delta v^\circ}{d-r} w + \left(\frac{f_y r}{(d-r)} - \frac{I^2}{f_y}\right)\Theta_h - \frac{JI}{f_x}\Theta_v - \frac{f_y J}{f_x}\Theta_r + 0 t_h - \frac{f_y}{(d-r)} t_v + \frac{I}{(d-r)} t_r = \frac{I\Delta r^\circ - f_y \Delta v^\circ}{d-r} w + \overline{\mathbf{U}}\overline{\mathbf{P}} \qquad (29)$$

Next suppose that the global motion parameters $\overline{\mathbf{P}}=(\Theta_h,\Theta_v,\Theta_r,t_h,t_v,t_r)$ are known between two successive frames which are estimated at first and separately. Then for prediction of the components of the vector of optical flow the following equation is valid

$$\hat{u}_j = \frac{J\Delta r^\circ + f_x \Delta h^\circ}{d-r}\hat{w} + \overline{\mathbf{V}}\overline{\mathbf{P}}, \qquad (30)$$

$$\hat{u}_i = \frac{I\Delta r^\circ - f_y \Delta v^\circ}{d-r}\hat{w} + \overline{\mathbf{U}}\overline{\mathbf{P}} \qquad (31)$$

where $\hat{w}$ is the animation parameter from the previous frame and Δh, Δv, Δr are the rotated components of the animation vector with the rotation angels of the global motion (eq.24). By substitution of eq.(28) and (29) into the optical flow equation [14] of the large motion for one feature vertex can be obtained the following linear equation with the unknown animation parameter w

$$\left(\hat{I}_i \frac{I\Delta r^\circ - f_y \Delta v^\circ}{d-r} + \hat{I}_j \frac{J\Delta r^\circ + f_x \Delta h^\circ}{d-r}\right) w = -\left(\Delta\hat{I}_t + \hat{I}_i \overline{\mathbf{UP}} + \hat{I}_j \overline{\mathbf{VP}}\right) \quad (32)$$

If N feature vertices are used then the system of N linear equations will be composed

$$\begin{pmatrix} e_{p1} \\ \vdots \\ e_{pN} \end{pmatrix} w = -\begin{pmatrix} \Delta I_{p1} \\ \vdots \\ \Delta I_{pN} \end{pmatrix} \quad (33)$$

where

$e_{pk} = \hat{I}_i^{pk} \frac{I\Delta r_{pk}{}^\circ - f_y \Delta v_{pk}{}^\circ}{d - r_{pk}} + \hat{I}_j \frac{J\Delta r_{pk}{}^\circ + f_x \Delta h_{pk}{}^\circ}{d - r_{pk}}$,

$\Delta I_{pk} = \Delta\hat{I}_t^{pk} + \hat{I}_i^{pk}\overline{\mathbf{UP}} + \hat{I}_j^{pk}\overline{\mathbf{VP}}$ are the members for the feature vertex pk, k=1,2,...N. After application of LSM on the solution of eq.(33) for the animation parameter w it is valid

$$w = \frac{-\sum_{k=1}^{N} e_{pk} I_{pk}}{\sum_{k=1}^{N} e_{pk}{}^2} \,. \quad (34)$$

If the local motion is realized by using more AU then for each of them the system of eq. (33) is composed and afterwards by application of LSM on its solution the animation parameter w belonging to the selected AU is obtained. The derived eq.(34) is formally the same as eq.(22) valid for estimation of the animation parameter of the large local motion. On the other side in eq. (34) there is furthermore information about the global motion which parameters are estimated independently on the animation parameter. It is clear that for animation of 3D model Candide with its global motion the accuracy of estimation of the animation parameters is dependent on that one of estimation of the global motion parameters before carried out. The accuracy can be further increased by the multiple iterations applied on eq.(30) until (34).

6. Experimental Results

The designed algorithm of animation of 3D model of the human head with the global motion has been experimentally verified for the testing videosequence "Miss America" with the frame rate 30 Hz, the size 288x352 pels and 8 bits per pel. As a specific 3D model we used 3D wire frame model Candide [12] which was calibrated by the first (reference) frame of the videosequence. For estimation of the animation parameters the camera parameters d, f_x, f_y, are also needed which were obtained by a calibration of the camera [14] according to the same reference frame. For the distance d=400 pels the scaled focal lengths f_x=354 pels and f_y=333 pels and together were used in all experiments. Verifying of the accuracy of animation of 3D model of the human head is very hard because in advance the exact values of the animation parameters are not known. As a suitable criterion of valuation of the accuracy we used the peak signal/noise ratio (SNR) for the human face region. To compare the results of animation of 3D model Candide by using the criteria in all experiments for its texturing the plane algorithm based on two dimensional affine transformation [13] was used. Note that the choice of the algorithm of texturing has not any direct impact on the accuracy of animation of 3D model Candide. Then conclusions in the paper are valid for any other algorithms of texturing [15].

For animation of 3D model Candide in combination with its global motion first the parameters of the motion are estimated with high accuracy. It was achieved such a way that from 3D model Candide in initial position its 35 feature vertices have been selected as it is seen in Fig.4a. After their projection on the frame we got the feature points in which the partial derivatives of the luminance function I(i,j,t) were calculated by using Sobel operator [16]. At selection of the feature vertices we proceeded under assumption that the ones of them corresponding points on the surface of the human head in the frame will have only the global motion. By such a procedure an influence of the local motion on the accuracy of estimation of the global motion parameters was eliminated. Following the high accuracy of the global motion estimation in the videosequence "Miss America" we used the algorithm on the basis of optical flow assuming large motion [14] for it. With reference to the character of the local motion in the videosequence we selected three AU, i.e. "Dropping of jaw", "Raising upper lip" and "Dropping of eyebrows" for which the three animation parameters $\overline{\mathbf{W}} = (w_1, w_2, w_3)$ were estimated. The feature vertices of the used AU, in which the partial derivatives of the luminance function $I(i,j,t)$ were calculated after their projection on the frame, are shown in Fig. 4b,c,d.

The local motion in the videosequence with the human head is considered as the large motion because in generally it has the global motion too. In addition the number of the feature vertices for AU is limited in regard to the finite number of its definition ones. Therefore in order to achieve the high accuracy of estimation of the animation parameters we used the multiple iterations in its block diagram (Fig. 3). In Fig. 5 there are the graphs of dependences of SNR of the synthesized frames after estimation of the global motion (Glob.M) with 3 iterations and the local motion (LM) with 1, 2, 3 and 10 iterations caused by animation of 3D model Candide using the selected three AU. From the graphs it is seen that using only one iteration by eq.(34) is not sufficient even such SNR is decreased. The decreasing is caused by the big inaccuracy of estimation of the animation parameters. Better results were achieved already for two iterations and for more then three iterations the accuracy of estimation of the animation parameters $\overline{\mathbf{W}} = (w_1, w_2, w_3)$ has not increased. For illustration in Fig.6 there are 4 successive frames with the projected 3D model Candide after its animation with the global motion using the three iterations.

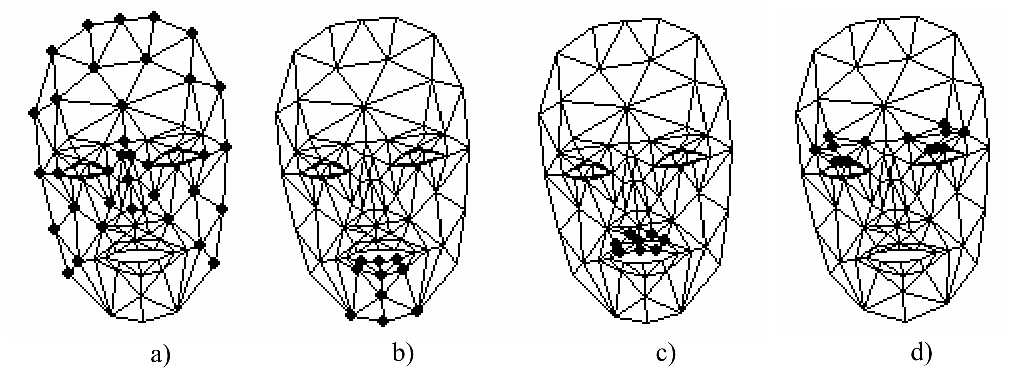


Fig. 4. The feature vertices for a) global motion, b) local motion of the jaw, c) local motion of the upper lip, d) local motion of the eyebrows.

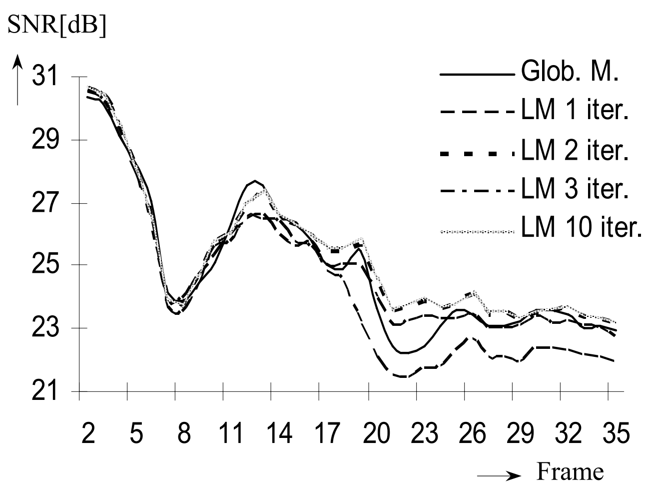


Fig. 5. SNR of the synthesized frames by animation with the global motion.

7. Conclusion

In the paper the animation of 3D model of the human head in combination with its global motion has been presented. The designed algorithm is very fast and with low calculation requirements because the estimation of the animation parameters is carried out directly from the input videosequence. In our experiments 3D model Candide was used which enables generation of different expressions by using of its AU. The estimation of the global motion parameters foregoes the estimation of the animation parameters and therefore affects the accuracy of the animation of 3D model of the human head. To increase the accuracy we have used the multiple iterations between two successive frames in the designed algorithm based on the optical flow. From the achieved results it follows out that they significantly particularized the estimated values of the animation parameters above all for the large motion. Of course the particularization is at the expense of the calculation requirements but on the other side it is necessary to suppress accumulation of the estimation errors in long videosequences. Last but not least the calculation of the partial derivatives of the luminance function using of Sobel operator contributed to the accuracy and stability of animation of 3D model Candide with the global motion. Objective valuation of the achieved results shows on the possibility of application of the designed algorithm of animation 3D model Candide in combination with its global motion in the model based human head coding with very high data compression. Regularly frame by frame in the method only the global motion and animation parameters are coded and transmitted with very low bit rate in the range of 300 till 2000 bits per second.

Acknowledgements

The work was supported by the Scientific Grant Agency of the Ministry of Education and the Academy of Science of the Slovak Republic under Grant No. 1/3133/06.

References

[1] The special issue of the *IEEE Trans.on Circuits and Systems for Video Technology on MPEG-4 SNHC*, July 2004.

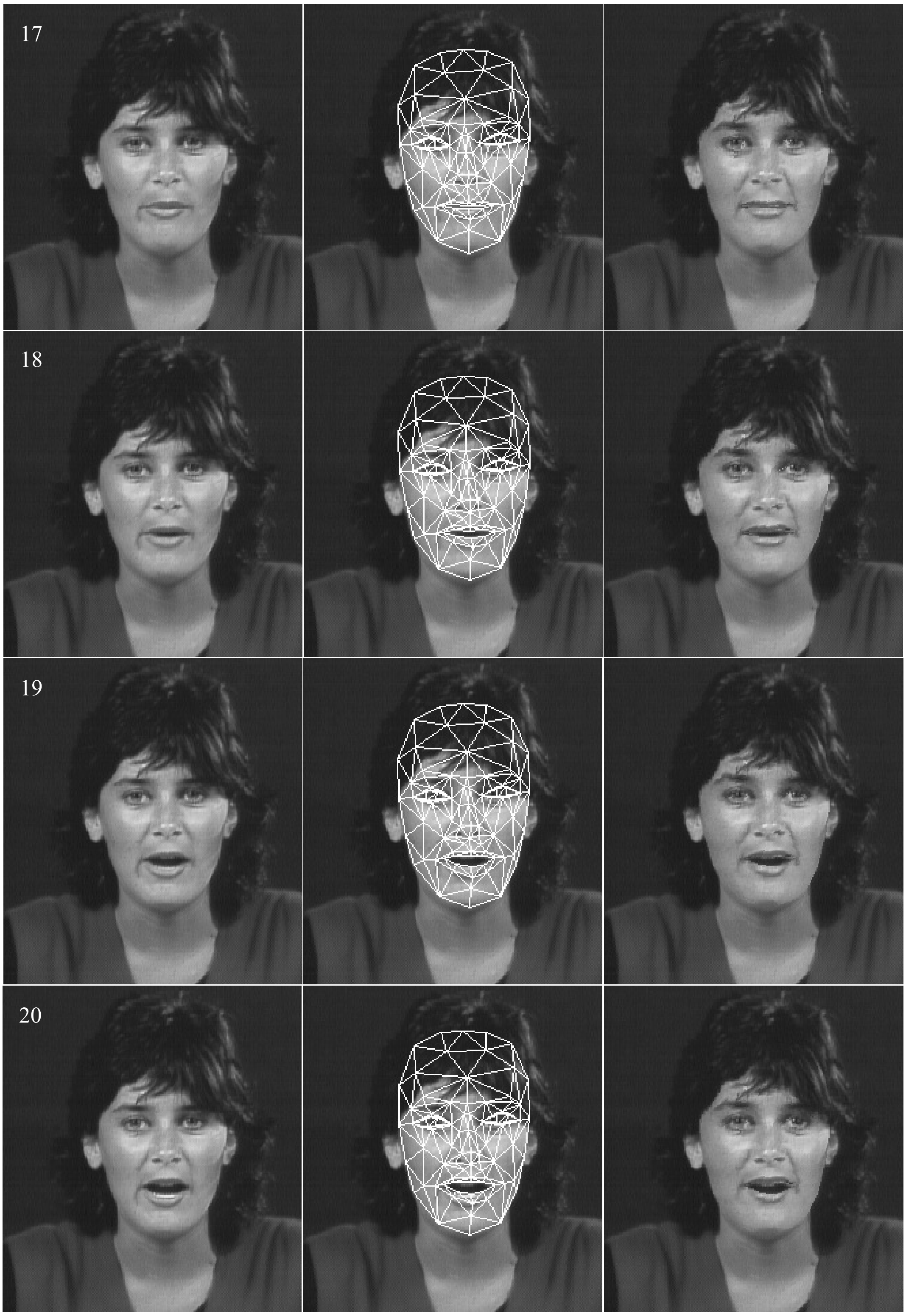
[2] STRINTZIS, M., SARRIS, N. (Ed.) *3D Modeling and Animation: Synthesis and Analysis Techniques for the Human Body.* IRM Press, Hershey, PA, July 2004.

[3] PANDZIC, I., FORCHHEIMER, R. *MPEG-4 Facial Animation: The Standard, Implementation and Appl.* John Wiley & Sons, 2002.

[4] MIHALÍK, J. *Image Coding in Videocommunnications.* Mercury-Smekal. ISBN-80-89061- 47-8, Košice, 2001. (In Slovak).

[5] CHAO, S., ROBINSON, J. Model-based analysis/synthesis image coding with eye and mouth patch codebooks. In *Proceedings of Vision Interface.* Banff Alberta, May 1994, p. 104-109.

[6] MIHALÍK, J. Neural network clustering vector quantizer design. *Neural Network Worl.,* 1993, vol. 2, no. 5, p. 197-207.



a) b) c)

Fig.6. a) The input frames of "MissAmerica" number 17, 18, 19, 20, b) the projected 3D model Candide after animation with the global motion, c) the synthesized frames number 17 (SNR=25dB), 18 (SNR=24,99dB), 19 (SNR=24,99dB), 20 (SNR=23,97dB) by using the planar texturing.

[7] ZHANG, L. Estimation of eye and mouth corner point position in a knowledge-based coding system. In *Proc. of SPIE.* 1996, vol. 2952, p. 21-28.

[8] YUILLE, A. L., HALLINAN, P. W., COHEN, D. S. Feature extraction from faces using deformable templates. *Int. Journal of Computer Vision,* 1992, vol. 8, no. 2, p. 99-111.

[9] ANTOSZCZYSZYN, P. M., HANAH, J. M., GRANT, P. M. A new approach to wire-frame tracking for semantic model-based coding moving image coding. *Signal Processing: Image Communication*. 2000, vol. 15, p. 567-580.

[10] EISERT, P., GIROD, B. Analysing facial expressions for virtual conferencing. *IEEE Computer Graphics and Applications,* 1998, p. 70-78.

[11] DAVIS, M., TUCERYAN, M. Coding of facial image sequences by model-based optical flow. In *Proc. of the Int. Workshop on Synthetic-Natural Hybrid Coding and 3D Imaging.* Rhodes (Greece), 1997, p. 192-194.

[12] AHLBERG, J. *Candide-3: An Updated Parameterised Face.* Rep. No. LiTH-ISY-R-2326, January 2001.

[13] MIHALÍK, J., MICHALČIN, V. 3D motion estimation and texturing of human head model. *Radioengineering,*2004, vol.13, no.1, p.26-31

[14] MIHALÍK, J., MICHALČIN, V. 3D motion estimation of human head by using optical flow. *Radioengineering,* 2006, vol. 15, no.2, p. 37- 44.

[15] MIHALÍK, J., MICHALČIN, V. Texturing of surface of 3D human head model. *Radioengineering,* 2004, vol. 13, no. 4, p. 44- 47.

[16] PRAT, W. K. *Digital Image Processing.* John Wiley & Sons, 1978.

About Authors...

Ján MIHALÍK graduated from the Technical University in Bratislava in 1976. In 1979 he joined the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, where he received his PhD degree in Radioelectronics in 1985. Currently, he is Full Professor of Electronics and Telecommunications and the head of the Laboratory of Digital Image Processing and Videocommunications at the Department of Electronics and Multimedia Telecommunications. His research interests include information theory, image and video coding, digital image and video processing and multimedia videocommunications.

Viktor MICHALČIN was born on 1976 in Ukraine. He received the Ing degree from the Technical University of Košice in 2000. He is a PhD student at the Department of Electronics and Multimedia Telecomunications of the Technical University, Košice. His research is focused on model based and very low bit rate video coding. Currently he is working as a developer of VRVS/EVO videoconferencing system in Caltech-VRVS-SK team at the University of P. J. Šafarik in Košice.

RADIOENGINEERING REVIEWERS

April 2007, Volume 16, Number 1