World Engineering & Applied Sciences Journal 7 (1): 44-57, 2016
ISSN 2079-2204

DOI: 10.5829/idosi.weasj.2016.7.1.22550

An Automatic Facial Localization Tracking Identification Biometric System Through PCA & Wavelet Distribution in 3D Mesh Environment

B. Rashida and Munir Ahamed Rabbani

Department of Computer Applications, BS Abdur Rahman,
University, Vandalur,Chennai. Tamil Nadu, India

Abstract: In this paper, the process of automatical facial region localization and tracking in video frames through 3D mesh model is estimated. The morphological region of the face are modeled into 18 geometry based regions based on the various shape and expressions. The feature are estimated based on the covariance matrix in high region space. Then, it undergoes to PCA to estimate facial deformations. The patterns feature are extracted and it mapped with the multi geometry mapping. Then, the corresponding wavelet transforms extracted region into various dimension for geometry matching for classification. The proposed method achieves roubustness based on its accurate transformation through wavele analysis.. The XM2VTSDB multi-modal face database is used to compare the real video sequence images. The experimental evaluation shows favarable results and yields 99% detection rate over 15000 video frame images. The frames in the facial tracking data are calculated through various parameters and compared with the ground truth against standard Euler angles of the muti-modal database.

Key words: 3D model · Face recognition · PCA · Wavelet transformation

INTRODUCTION

The robust monitoring of a face and its main characteristics is an important issue in the field of computer vision and has many applications. Examples include security (surveillance, personal identification), or man-machine interaction. The methods of object tracking deforming 3D vision typically based on a model of appearance and a model of temporal evolution of the objects. Appearance consideration may be global or local (encoded bye. g. In the form of points of interest, or contours). The appearance model is used to explore the search space considered, using matching methods, maximizing a similarity criterion.The monitoring of the face pose is present in many applications, playful perspectives as reality increased, for security protocols using facial recognition, through the man-machine interaction. Many approaches have been explored in the follow-of head movement. Some methods proceed monitoring of facial features to trace the information about the pose. Facial characteristic is the description of a specific area of interest of the face. These areas, such as the corners of the eyes, the nose tip, or the corners of the mouth are common to all faces. They, however, look and different position in different individuals. Locate and track these features on mobile and unknown faces are major challenges. Tong*et al*. Designed their system monitoring [1] with an active appearance model (AAM) hierarchical. [2] The facial features are areas represented by Gabor wavelets and profiles [3] Grayscale. Many other monitoring systems are based on purely two-dimensional representations of faces. Their effective operation implies that areas of interest are visible, which is binding on after laying the face. The three-dimensional approaches have been developed with a view to include variations of rotation of the larger faces. The object tracking is a widely treated in thematic vision in applications as diverse as video surveillance (monitoring of vehicles or people), the road safety (monitoring head of the driver attention control) or the human interaction (monitoring of hands for the interpretation of language sign.

Earlier Work: The active appearance models have been proposed as powerful tools for analyzing faces [4]. They can be used to track which faces both the form, the installation (often 2D) and appearance vary time course.

Corresponding Author: B. Rashida, Department of Computer Applications, BS Abdur Rahman, University, Vandalur, Chennai. Tamil Nadu, India.

Monitoring methods based on such models, however, remain robust when image acquisition conditions differ from the training set (lighting, camera properties, dimming, etc..). Only close enough look of faces that of faces belonging to the learned class can also be followed in good conditions accurately. In [5], the authors present the general problem of registration or alignment of the image, from a descent calculated gradient in an additive or compositional approach. In [6], a model based on appearance a mixture model is recommended for object tracking natural. It implements an estimate online model by EM algorithm. In [7], a method consisting of two consecutive steps was developed to monitor the 3D pose of the face and its deformations. The first step is learning the possible deformations3D faces continuing binocular data. The second step is simultaneously pursuing 3D pose and facial deformities by calculating the optical flow associated with primitive followed. In [8], the 3D pose of the face is estimated by readjusting the current texture of the face with respect to linear combination of texture templates and illumination. A stable tracking has been obtained by minimizing the least squares of the offset error. More complex models use close forms of the human face. Some forms may have a relatively appearance gross of only a hundred points describe the face, others may have a photo-realistic look. Basu *et al*. [9] Developed a system to measure the movements of the head, which is modeled by an ellipsoid. The three-dimensional nature of the model allows to disqualify non-visible parts of the face during the process of tracking. The texture model is described by optical flow, while the estimate optimal parameters, installation is done using the simplex algorithm. La Cascia *et al*. and Xiao *et al*. [10] proposed a similar approach with a cylindrical face model. La Cascia *et al*. Have chosen to optimize the illumination settings and installation simultaneously. Xiao *et al*. Applied a form of iterative weighting algorithm least squares [11] to update their reference texture, considering the probable occlusions and illumination changes. Our belief is that the models 3Ds extracts real faces will be increasingly tailored to the face tracking that simplified forms. We therefore propose here an alternative method. Although our approach is built around publications [12], we use a specific face model extracted from high resolution 3D scans of faces. Nowadays such a precise model Basel facial model (BFM) [13] is not applicable in its original form for tracking implementations face. The number of points describing each face implies too high dimensionality and quantities of too large for common calculations machinery. To track an object, one of the methods most used is the Kalman filter, which, however imposes restrictive conditions on the system linearity and requires special adaptation for a three-dimensional monitoring from multiple 2D views. [14] In this, Regions particulate filter tracking method that allows easily express tracking poses in 3D space merging the 2D information about different views. In this article, different particle filter algorithms are compared, including multitasks methods such as simulated annealing [15] and a variant taking into account the specificities of the face. These methods optimize the distribution of the particles by increasing the number of sampling for each frame.

Proposed Work: First, we present the appearance model of the face we use to represent facial structure. Then we propose a method face tracking and six elementary facial gestures, also called facial actions based on an algorithm of descent. This method extends the concept of online updated models described in [16] to the case of monitoring3D poses and motions on the face of the primitive (movements of the eyebrows, lips...) due by example, in facial expressions.

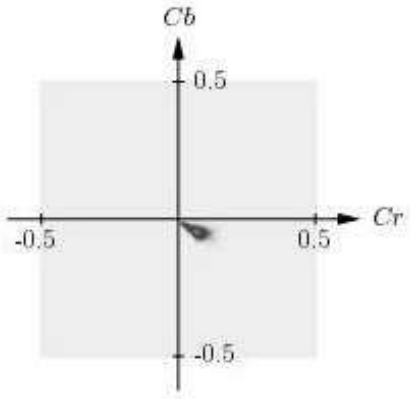


Fig 1(a) Color distribution of a face

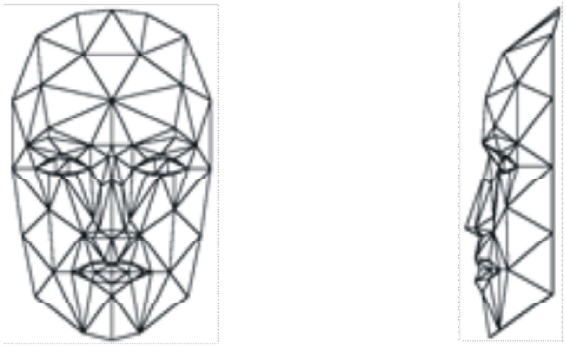


Fig 1(b) Segmented face

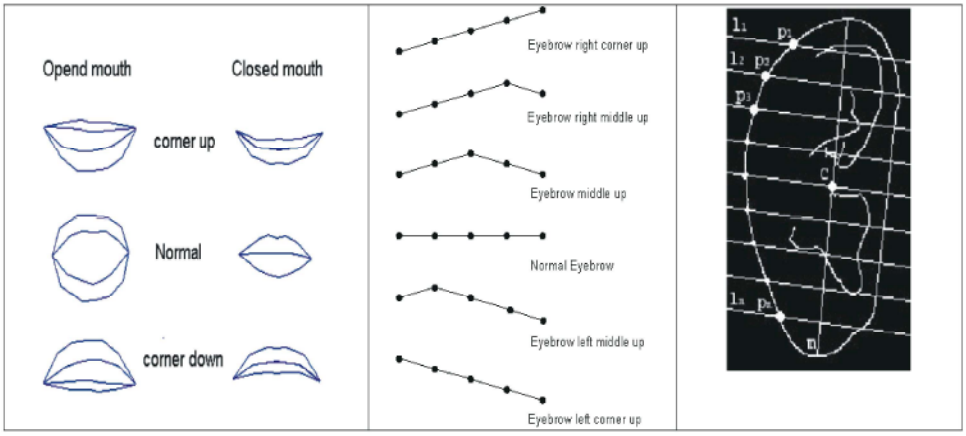


Fig 2: types of Mouth, Eyebrow and layers of Ear region

Tab 1: Parameter defining in face tracking

No	Model regions
0	Head height
1	Vertical position of the eyebrows
2	Vertical Position Eye
3	Eye Width
4	Eye Height
5	Horizontal separation of the eyes
6	Depth cheeks
7	Depth nose
8	Vertical position of the nose
9	Vertical position of nose
10	Vertical Position of mouth
11	Width of the mouth
12	Lift the upper lip
13	Lower the lower lip
14	Stretch horizontally lips
15	Lower eyebrows
16	Stretch vertically lips
17	Upper eyebrows

Columns matrices S and A are respectively the units of shapes and facial actions. The τS and τA vectors encode, respectively, for their part, the parameters detaining face following 18 modes (Table 1) and facial motion parameters according to six methods (From 12 to 17). All S-shaped units provides way to adapt the 3D model in the face of the subject. A form unit applies a displacement (encoded by a vector) on a reduced set of points that govern the width of the eyes, face height, separations, etc. All units Action A provides way to reproduce the 3D model movements face. Action unit applies a move on reducing set points that govern the lifting of the lip upper, the lowering of the lower lip, stretching eyebrows, etc.. Thus, the term SτS takes into account the variability while face, the term takes into account the intra-individual variability. S and A are constant over the model and τSτA encode variations. We assume that these two variations are decoupled, (IE) that the vector τA facial expressions will be supposed to be representative of all of the population and thus facilitate learning expressions.For a given person, all forms unitsτS is constant because it encodes the physiognomy of the face. In this part of this study, the τS vector is initialized manually, aligning the shape of the Candide model from the target face present in the first frame of the video. The depth of the face can be regarded as very small compared to the depth of the scene shot, the effects of perspectives can be neglected. That is why we have adopted an orthographic projection to scale (low perspective). The projection matrix size 2 × 4 depends on the settings, 3D pose of the face (rotations and translations) and internal camera parameters (scale). We are planning a summit3D model of Candide

Pi = [Xi Yi Zi] T ◎ g in the image at a point pi = [uivi] T by the following equation:

[ui vi] T = M [Xi Yi Zi 1] T

The state vector of Candide model consists of evolving parameters during follow: laying information3D (three rotations and two translations), the cam- data (scale) and facial gestures control vectorτA. This is represented by a vector b of dimension 18 (0 to 17 models):

b = [.θ.x,. θ.y, θz,tx, ty, s, τAT] T

3D Morphological Estimation: The objective of monitoring is the estimate of the system state (in our case it is encoded by the vector bit) from a set of images, y [t] = {$\mathbf{y}_1$Yt} arriving by set accordingly. Monitoring has two

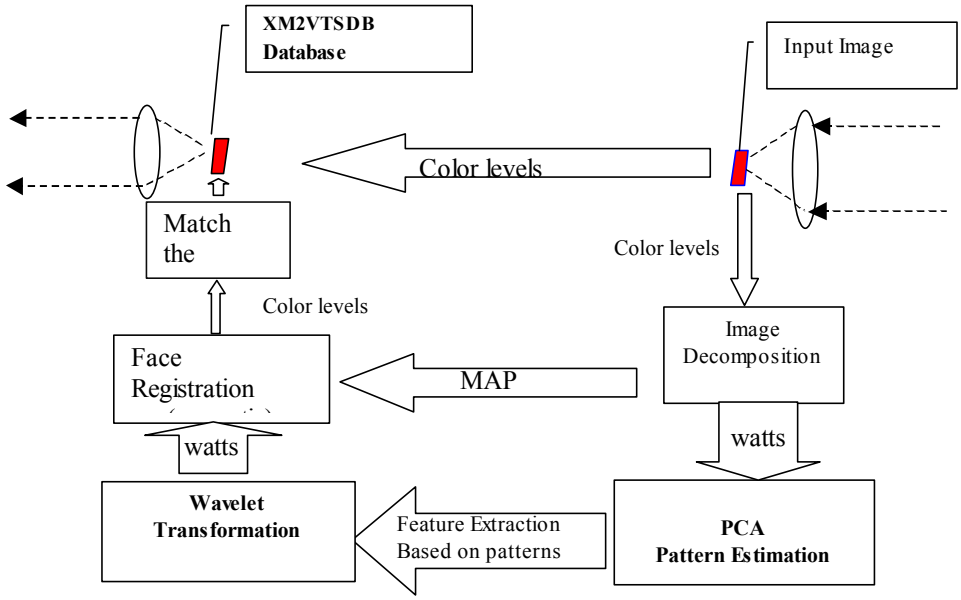


Fig. 3: An Automatic Facial localization Tracking Identification Biometric System Head tracking poses variation,

main components: (Tt) an observation model that characterizes the facial texture and a transition model that characterizes the kinetic (Dt) describing the evolution of the state between two observations. The texture modulates time t, Tt is the observation model. It models the texture of all the observations up to time t - 1. Its parameters vary over time, so this is an adaptive model. In each image, the observation is none other than the texture readjusted xτ (bt) = W (yτ, bt)[17-22].

Adaptation to the laying head tracking now describes the adaptation of the particulate filter in our context to track laying. Because the change in appearance of a face in the event of change of orientation, we consider simultaneously the three rotation angles and the position of the head. The status hidden xt looking at each moment is laying the face, observations, it is the images acquired at this time and x (t)represents the laying of the particle. The predicting step changes the state of each particle between, observations, It therefore characterized changes in position and orientation of the head between two consecutive acquisitions (the period being135 milliseconds). The acquisition rate is quite low, we choose a minimalist modeling change of poise during the advance. The functional equation 2 simply reflects the advance along the examine airlock. The rest of the installation changes between t – 1and t is expressed by adding Gaussian noise át.Once the state of the updated particle the second step is to reassess their weight according to observations made at time t. The update procedure to date is as follows

Selecting Patches: Given the state x (t) a particle, the 3D points of the model is first projected on each image. Patches are then extracted around these projections (Figure 5). By Moreover, it calculates the pose of the head in the mark each camera to generate the synthetic view is extracted and associated patches around the same 3D projected points. The correlation between patches of the model and those extracted observations can then be calculated as follows.

Similar in Texture Criterion: The aim being to simultaneously confirm the position and orientation of the face, pixel patches compared to pixel. Similarity is calculated by normalized cross correlation with zero mean (ZNCC) to be invariant to Face illumination changes during the sequence.

This system performs well, but as noted by its designers in the conclusions of their studies, it is derived occasionally the target object. In particular, care, especially, was given to the optimization of forgetting factors influence on both the stability and adaptability of the system 2D tracking. We have therefore modified the original algorithm on two points:

To minimize drift, the appearance model is a more updated view of the state related to the best particle. It is fed by detected by views face detector Viola-Jones [17] selected by spatial proximity and size criteria. - To preserve the robustness of the system outside the detection cone Viola-Jones - whose maximum angle is typically 40°- each likelihood of particle is then formed by

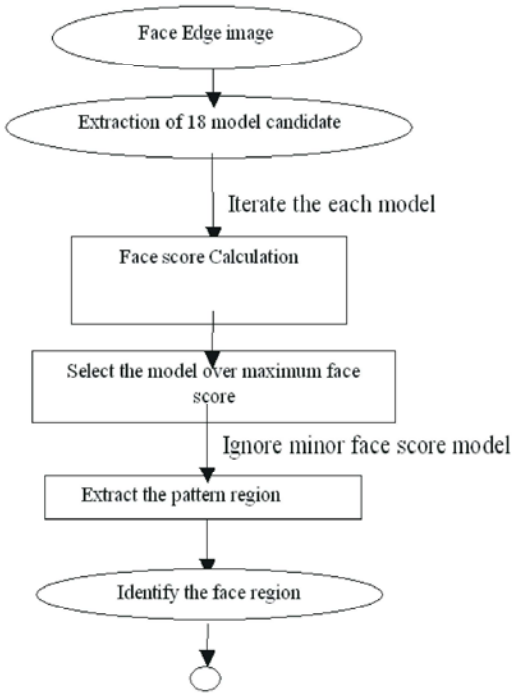


Fig 4: Level- 1 Data Flow Diagram of the model estimation

a joint probabilistic model which combines its similarity in appearance model and its auto similarites compared to the last observation, a Temporary template that evolves independently for each particle

The image frame S is said to be a MRF if

$$\forall s \in S, p(X_s \mid X_r, r \neq s) = p(X_s \mid X_{\delta_s}) \quad (1)$$

where s and r are pixel locations in the lattice and δ_s are the gray-level at the s and r locations, respectively; $p(X_S|X_{\delta s})$ is an approximation of the pixel at locations and X_S is the probability of finding the gray level at locations given the gray levels in the neighborhood δ_S [18] and the Gaussian parameter for the least square identification is given by

$$f_j = \frac{1}{u^2} \sum_{r, r \pm j \in R(s)} \left[y_r - \theta_j Q_j(r) \right]^2, j = 1, \ldots, 4 \quad (2)$$

where

$$Q(r) = \left[(y_{r+\tau_1} + y_{r-\tau_1}), \ldots, (y_{r+\tau_4} + y_{r-\tau_4}) \right]^T \quad (3)$$

τ Shows the position (0°, 45°, 90° and 135°) along the pixel location r and R(s) is the estimated frame in the image and the response of the image is given by,

$$\hat{\Theta} = \left[\sum_{r, r \pm j \in R(s)} Q(r) Q(r)^T \right]^{-1} \left[\sum_{r, r \pm j \in R(s)} Q(r) y_r \right] \quad (4)$$

Along with the four response images generated by f_j,

Principle Component Analysis: Usually applying a PCA on a set of N random variables X1,..., X N is known from a sample of achievements of these variables. This sample of the N random variables can be structured in a matrix M to K rows and N columns.

Var Ind	X_1	X_2	…	X_j	…	X_p
Ind.1	x_{11}	x_{12}	…	x_{1j}	…	x_{1p}
Ind.2	x_{21}	x_{22}	…	x_{2j}	…	x_{2p}
…	…	…	…	…	…	…
Ind.i	x_{i1}	x_{i2}	…	x_{ij}	…	x_{ip}
…	…	…	…	…	…	…
Ind.n	x_{n1}	x_{n2}	…	x_{ni}	…	x_{np}

Each random variable X n which X 1, n, . .., X K, n are independent realizations, has a mean \ X_n is a standard deviation V or σ of the data Xn.

$$V = \begin{bmatrix} \mathbf{Var(X_1)} & Cov(X_1,X_2) & \ldots & Cov(X_1, X_j) & \ldots & Cov(X_1,X_p) \\ Cov(X_1,X_2) & \mathbf{Var(X_2)} & \ldots & Cov(X_2, X_j) & \ldots & Cov(X_2,X_p) \\ \ldots & \ldots & \ldots & \ldots & & \ldots \\ Cov(X_1, X_j) & Cov(X_2, X_j) & \ldots & \mathbf{Var(X_j)} & \ldots & Cov(X_j,X_p) \\ \ldots. & \ldots & & & & \ldots \\ Cov(X_1,X_p) & Cov(X_2,X_p) & \ldots & \ldots & \ldots & \mathbf{Var(X_p)} \end{bmatrix}$$

If the achievements (the elements of the matrix M) are equal probability, then every achievement (a component X_ {I, j} of the matrix) has the same importance 1 / K_{in} the calculation of sample characteristics. One can also apply a weight p {i}to each different embodiment of joint variables (for samples recovered, pooled data,...). These weights, which are positive numbers sum 1 are represented by a diagonal matrix D of sizeR_{xy} :

$(R_{xy})_{ij} = \square (X_i, Y_j)$ i=1, …, p et j = 1, …, q :

$$R_{xy} = \begin{array}{c} \\ X_1 \\ X_2 \\ \dots \\ X_i \\ \dots \\ X_p \end{array} \begin{array}{c} \begin{array}{ccccccc} Y_1 & Y_2 & \dots & Y_j & \dots & Y_q \end{array} \\ \begin{bmatrix} \square(X_1,Y_1) & \square(X_1,Y_2) & \dots & \square(X_1,Y_j) & \dots & \square(X_1,Y_q) \\ \square(X_2,Y_1) & \square(X_2,Y_2) & \dots & \square(X_2,Y_j) & \dots & \square(X_2,Y_q) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \square(X_i,Y_1) & \square(X_i,Y_2) & \dots & \square(X_i,Y_j) & \dots & \square(X_i,Y_q) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \square(X_p,Y_1) & \square(X_p,Y_2) & \dots & \square(X_p,Y_j) & \dots & \square(X_p,Y_q) \end{bmatrix} \end{array}$$

The choice whether or not to reduce the point cloud (ie K random variable realizations (X 1, . .., X N)) is a model of choice. If it does not reduce the cloud: a high variant variable will be "pull" the whole effect of the PCA to it; if it reduces the cloud: a variable that is but a noise will end up with an apparent variance equal to an informative variable S^2.

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \qquad \text{ou} \quad S^2 = \sum_{i=1}^{n} p_i (x_i - \bar{x})^2 \tag{5}$$

To understand, imagine that the variance of u is equal to the variance of the cloud; we would have found a combination of X n which contains all the diversity of the original cloud (at least the part of all its diversity captured by the variance).A criterion commonly used is the variance of the sample (in order to maximize the variance explained by the vector u). For physicists, it rather has the sense of maximizing the inertia explained by u (that is to say to minimize the inertia of the cloud around u).First, the covariance matrix of the feature is extracted, then the eigenvectors are analyzed from a new linear transformation of the original attribute space:

$$PC_j = \alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 + \alpha_4 f_4 + \alpha_5 f_5 \tag{6}$$

where j stands for the jet principal component () and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ are the features contributions to forming the component. The diagonal of the correlation matrix (or covariance if we place ourselves in a non-reduced model), allowed us to write the vector which explains the more cloud inertia is the first eigenvector. Similarly the second vector which explains the bulk of the remaining inertia is the second eigenvector, etc.

$$\begin{bmatrix} PC_1(1,1) \dots PC_1(1,n) \\ \vdots \\ PC_1(n,1) \dots PC_1(n,n) \end{bmatrix} \tag{7}$$

In PCA, it is common that we want to introduce extra variables. For example, measurements of many quantitative variables on facial poses, the facial tracking poses have variables, for example the spaces to which the face tracking poses.

These data are subject to various quantitative variables. When analyzing the results, it is natural to try to connect the main components to the qualitative variable spaces. For this, the following results are produced. Identification on factorial designs, different species by representing, for example different colors. Representation on factorial designs, plant centers of gravity belonging to the same species. Indicates, for each center of gravity for each axis of a critical chance to judge the significance of the difference between a center of gravity and origin. In a series of P-frames, each pixel is regarded as a point of an affine space of dimension P, whose coordinates are the pixel value of each of P frames over time. The cloud, thus formed by all the image points can be analyzed by PCR, (it forms a hyper-ellipsoid dimensions P) which determines its principal axes.

Wavelet Transformation: A wavelet is a function based on the wavelet decomposition, decomposition similar to the Short-Time Fourier Transform [19], used in the signal processing. It corresponds to the intuitive idea of a function corresponding to a small oscillation based on its name

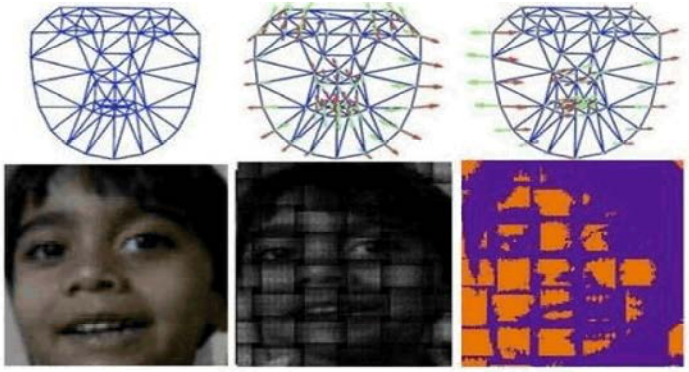


Fig. 5: Face Mashing based on 18 geometrical region

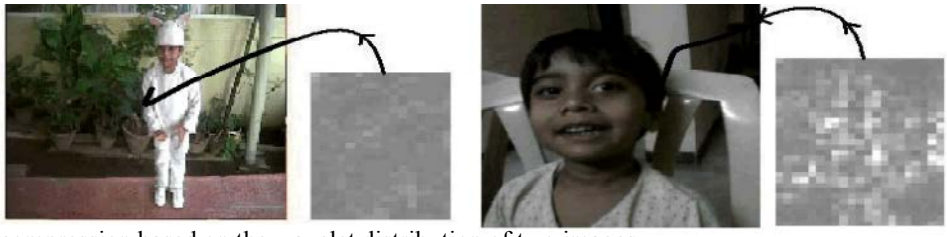


Fig 6: Image compression based on the wavelet distribution of two images.

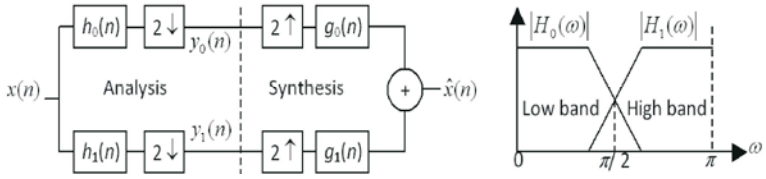


Fig. 7: Subband coding in wavelet transformed images

The z-transform shows the basic sampling theorem and it is shown below,

$$\hat{X}(z) = \tfrac{1}{2}\left[H_0(z)G_0(z) + H_1(z)G_1(z)\right]X(z) + \tfrac{1}{2}\left[H_0(-z)G_0(z) + H_1(-z)G_1(z)\right]X(-z) \quad (8)$$

The initial set is the set of integrable functions of a real variable x. The target set is the set of functions of a real variable. Concretely when this transformation is used in signal processing, it will be appreciated readily *t*

$$H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0$$
$$H_0(z)G_0(z) + H_1(z)G_1(z) = 2 \quad (9)$$

In multi resolution analysis, the technique reduces the size of digital information (quality of compressed information from the complete information), but also speed up the information display (display quality from a file compressed). The latter use is indispensable for cartographic where the quality and size of the information required are considerable.A signal $f(x)$ can be analyzed as a linear combination of expansion functions. The single-scale filter bank of Figure 5 can be "iterated" by tying the approximation. It is based on the use of wavelet compression for the elimination of non-perceptible high frequency information from the eye.

$$f(x) = \sum_k \alpha_k \phi_k(x) \quad (10)$$

Similarly, the expansion function is given by the space of resolution as,

$$\phi(x) = \sum_n h_\phi(n)\sqrt{2}\phi(2x - n) \quad (11)$$

the wavelet series expansion of function $f(x) \in L^2(\mathbf{R})$ relative to wavelet $\Psi(x)$ and scaling function $\phi(x)$. We can write

$$f(x) = \sum_k c_{j_0}(k)\phi_{j_0,k}(x) + \sum_{j=j_0}^{\infty}\sum_k d_j(k)\psi_{j,k}(x) \quad (12)$$

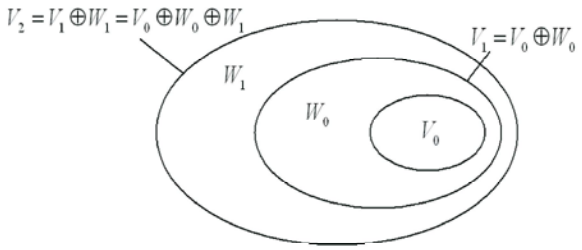


Fig. 8: The relationship between scaling and wavelet function spaces.

where J_0 is an arbitrary inital scale and the $c_{j0}(k)$'s are normally called the factors coefficients, the $d_j(k)$'s are the discrete wavelet coefficients. This expansion prominently used in facial image data compression and extraction,

$$c_{j_0}(k) = \left\langle f(x), \tilde{\phi}_{j_0,k}(x) \right\rangle = \int f(x)\tilde{\phi}_{j_0,k}(x)dx \tag{13}$$

$$d_j(k) = \left\langle f(x), \tilde{\psi}_{j,k}(x) \right\rangle = \int f(x)\tilde{\psi}_{j,k}(x)dx \tag{14}$$

The sequence of the continuous function $f(x)$ .shown below to compare transformation

$$W_\phi(j_0,k) = \frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} f(x)\tilde{\phi}_{j_0,k}(x) \tag{15}$$

$$W_\psi(j,k) = \frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} f(x)\tilde{\psi}_{j,k}(x) \tag{16}$$

for $j \geq j_0$ and

$$f(x) = \frac{1}{\sqrt{M}} \sum_k W_\phi(j_0,k)\phi_{j_0,k}(x) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j,k)\psi_{j,k}(x) \tag{17}$$

where $f(x), \phi_{j_0,k}(x)$ and $\psi_{j,k}(x)$ are functions of discrete variable $x = 0, 1, 2, .., M$ (1. The FWT shows the similarity of tracking, extraction in the facial image scheme of various section [20]. Consider again the multiresolution equation

$$\phi(x) = \sum_n h_\phi(n)\sqrt{2}\phi(2x-n) \tag{18}$$

Scaling x by $2j$, translating it by k and letting $m = 2k + n$ gives

$$\phi(2^j x - k) = \sum_n h_\phi(n)\sqrt{2}\phi\left(2(2^j x - k) - n\right) = \sum_m h_\phi(m-2k)\sqrt{2}\phi\left(2^{j+1}x - m\right) \tag{19}$$

Similarity,

$$\psi(2^j x - k) = \sum_m h_\psi(m-2k)\sqrt{2}\phi(2^{j+1}x - m) \tag{20}$$

Now consider the discrete wavelet transfrom. Assume $\tilde{\phi}(x) = \phi(x)$ and $\tilde{\psi}(x) = \psi(x)$, we substitute Eq. (20) Into Eq. (21), we get

$$W_\phi(j_0,k) = \frac{1}{\sqrt{M}} \sum_m h_\psi(m-2k)\sqrt{2}\phi(2^{j_0}x - m) \tag{21}$$

Then we substitute Eq. (20) into Eq. (21), we get

$$\begin{aligned} W_\psi(j,k) &= \frac{1}{\sqrt{M}} \sum_x f(x)2^{j/2}\psi(2^j x - k) \\ &= \frac{1}{\sqrt{M}} \sum_x f(x)2^{j/2}\left[\sum_m h_\psi(m-2k)\sqrt{2}\phi(2^{j+1}x - m)\right] \\ &= \sum_m h_\psi(m-2k)\left[\frac{1}{\sqrt{M}} \sum_x f(x)2^{(j+1)/2}\phi(2^{j+1}x - m)\right] \end{aligned} \tag{22}$$

where the bracketed quantity is identical to Eq. (22) with $j_0 = j + 1$. Therefore,

$$W_\psi(j,k) = \sum_m h_\psi(m-2k)W_\phi(j+1,m) \tag{23}$$

Similarity,

$$W_\phi(j,k) = \sum_m h_\phi(m-2k)W_\phi(j+1,m) \tag{24}$$

The below block diagram show the general way of decomposing of the facial image into various feature based vectors and the

Overall Face Model: The mixture models can easily integrate different components. They use it to make a Face model law where each probability density is distinct population. So we can have a model that explains both the texture and occlusions. Jepson *et al*. [21] and Hasler *et al*.

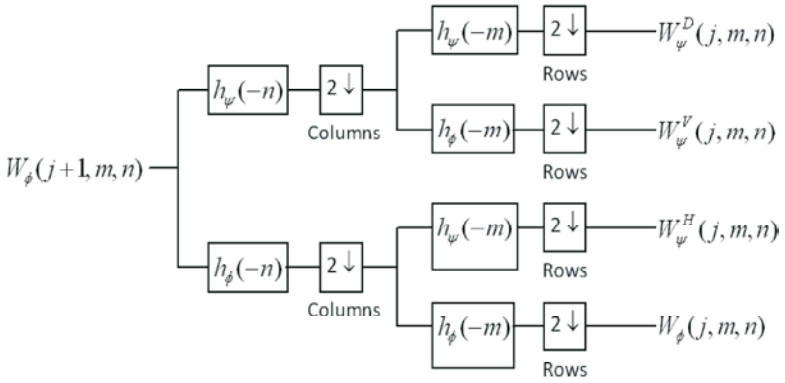


Fig. 5: The two-dimensional FWT ___ the analysis filter. bank

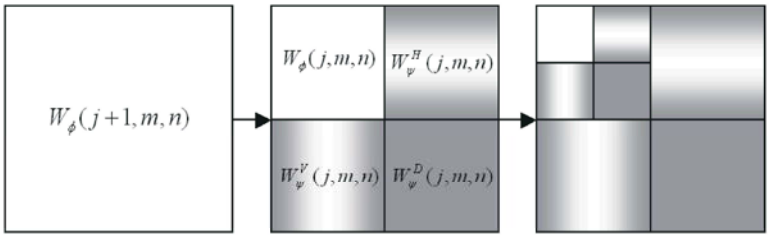


Fig. 9: Process of image decomposition

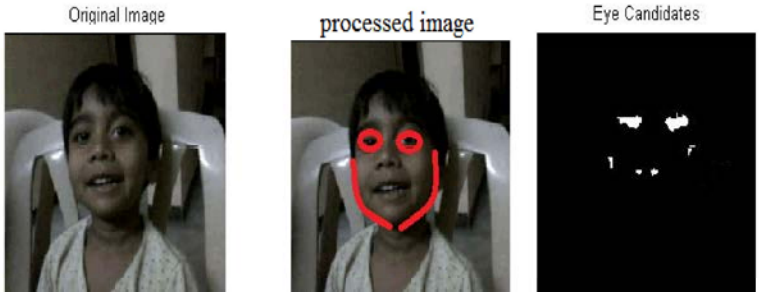


Fig. 10: Eye Candidate Extraction

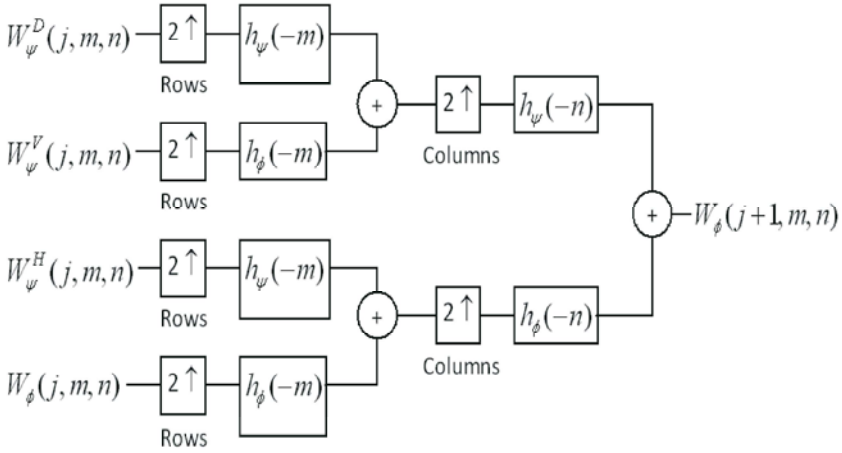


Fig. 11: The two-dimensional FWT synthesis filter bank.

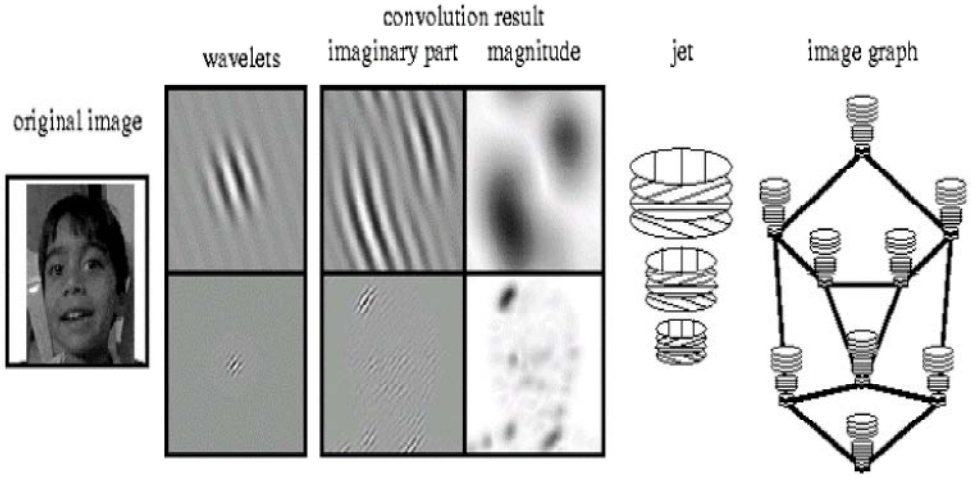


Fig. 12: Illustration of Wavelet Estimate the graphical region

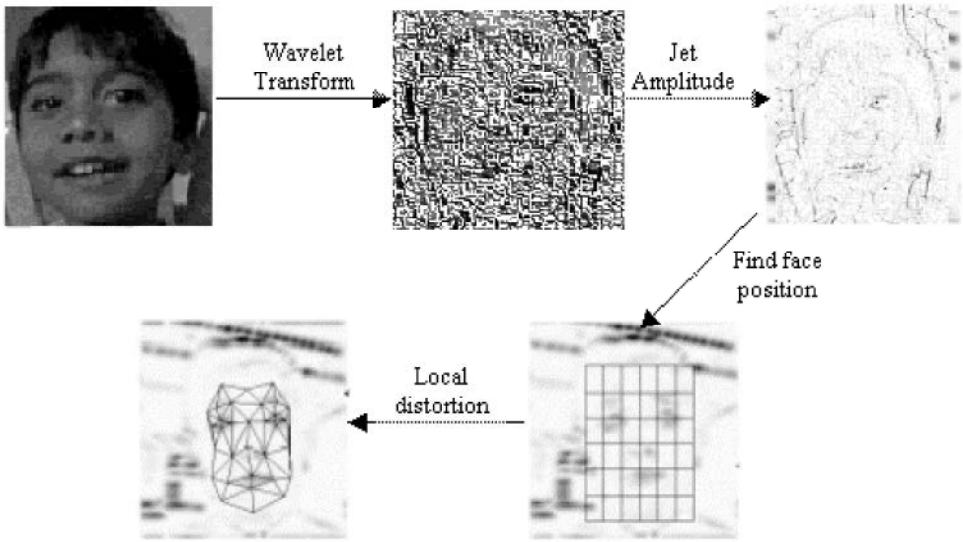


Fig. 13: Final 3D Mapping on Given image over Datasets

[22] Proposed use in connection with the block matching. Our second implementation is based on this approach applied to monitoring the position of the face and facial gestures.The observation xt undergoes a continuous and slow change when the subject changes the position his head in space or when changing expressionFacial. It is not the same with occlusions of the face, changes in appearance (the subject may withdraw his glasses), video noise. This brings us to build a mixture of laws containing two components: a stable component and a noise component. The Stable component: This component model the Sstable and constant phenomena of observation. Its probability density is normal ps (xt; microT, Σ2t). The parameters and micro Σ2t change slowly over time, they are respectively the mean and covariance that law. This component represents the texture modelTt the general method without the management of hidden. The noise component: This component model the Bunstable or sudden phenomenon. As the pixels are normalized in the range [-11], it is 12 in our study. These two components are combined with a model of mixture (Fig. 3).

Experimental Results

XM2VTSDB Multi-modal Face Database: It contains four records of 295 subjects taken over a period of four months. Each record contains a talking head shot and a rotary whim. Data sets taken from this database are

Table 2: Results of facial tracking of Data1 through texture based PCM biased parameters

DATA1	Minimum	Quartile 1	Median	Quartile 3	Maximum	STD
FRAME1	0.03	0.47	0.63	0.8	0.79	0.3
FRAME2	0.12	0.49	0.59	0.8	0.77	0.13
FRAME3	0.02	0.40	0.57	0.74	0.9	0.25
FRAME4	0.03	0.37	0.59	0.72	0.77	0.23
FRAME5	0	0.39	0.54	0.74	0.9	0.24

Table 3: Results of facial tracking of Data2 through texture based PCM biased parameters

DATE2	Minimum	Quartile 1	Median	Quartile 3	Maximum	STD
FRAME1	0.03	0.45	0.63	0.75	0.89	0.4
FRAME2	0.14	0.55	0.67	0.72	0.84	0.16
FRAME3	0	0.33	0.64	0.76	0.83	0.27
FRAME4	0.06	0.5	0.62	0.8	0.85	0.24
FRAME5	0.07	0.43	0.63	0.76	0.84	0.27

Table 4: Results of facial tracking of Data3 through texture based PCM biased parameters

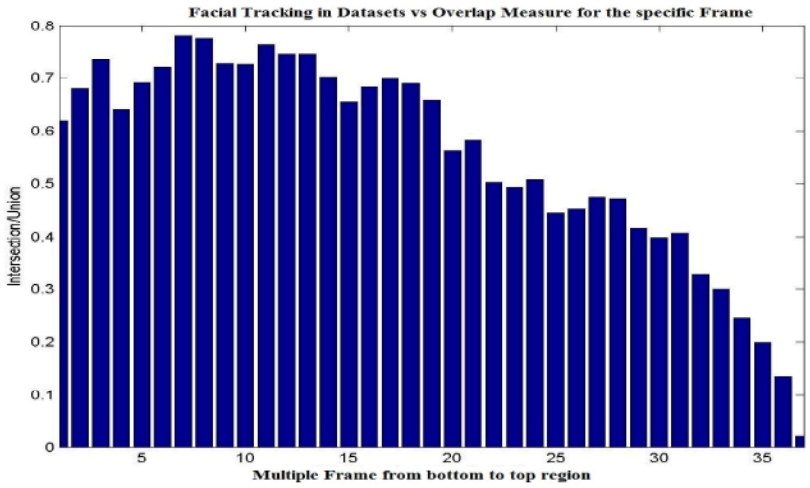
DATE3	Minimum	Quartile 1	Median	Quartile 3	Maximum	STD
FRAME1	0.03	0.45	0.64	0.8	0.79	0.3
FRAME2	0.14	0.6	0.5	0.8	0.79	0.17
FRAME3	0.02	0.37	0.58	0.77	0.9	0.25
FRAME4	0.05	0.38	0.55	0.76	0.79	0.24
FRAME5	0	0.39	0.59	0.75	0.9	0.27

Table 5: Results of facial tracking of Data4 through texture based PCM biased parameters

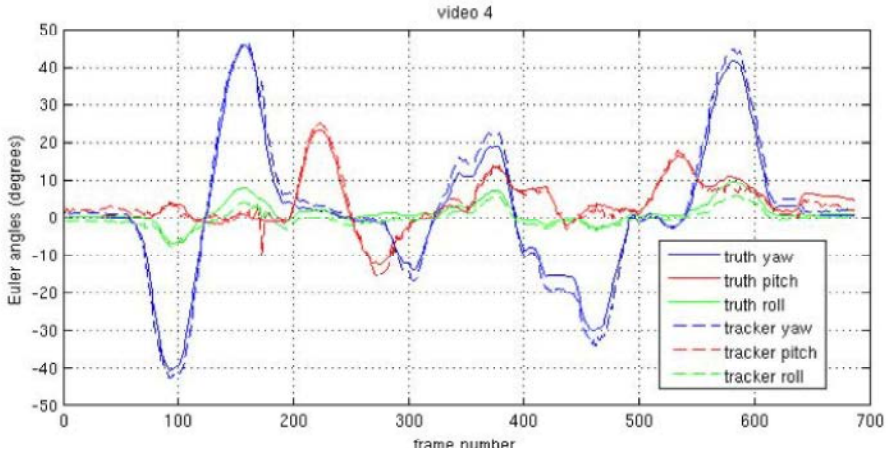
DATA4	Minimum	Quartile 1	Median	Quartile 3	Maximum	STD
FRAME1	0.03	0.47	0.69	0.75	0.81	0.3
FRAME2	0.15	0.56	0.67	0.74	0.84	0.15
FRAME3	0	0.5	0.62	0.77	0.83	0.26
FRAME4	0.07	0.5	0.64	0.75	0.85	0.24
FRAME5	0.07	0.40	0.65	0.77	0.83	0.24

available, including color images of high quality 32 kHz, 16-bit audio, video and a 3D model. For more information about the database and how to order it, follow the links at the side of this page.To ensure detection works despite the differences in brightness, it is necessary to normalize other images (that of departure, the template), such as those against which it tested. For this, it can already make a white balance all images. It must have a way to spread the spectrum of colors in the image (the color histogram), for it covers all the available values (there are the functions it need in MATLAB to access this histogram). After that, it can then change the color mode of representation to abstract lighting differences. So it takes the image in RGB, it equalized, as explained in on it, it passes in the HSV (or HSL, so is our case) and then the image HSV, it makes a Grayscale. In this case, it to write to passing loop colors_HSL -> Grayscale, RGB function because cutworms MATLAB is optimized for color perception of the human eye (that is not so not a mean beast of 3 channels but a weighted average, which would break completely in the case of HSV transition to gray). It can, for this conversion, make a pass on the channel 3 H, S and L, averaging a new Image single channel, which would be the image "HSV grayscale." The waterfalls Haardisenables with MATLAB to detect the eye correspond to an open eye. So if it wants to automatically build a template automatically, it will inevitably open eye template. If it knows that it arrange an eye picture closed, then it can build a closed eye template, it will actually easy. It can also detect the blinking action of the eye using a computing optical flow within the corresponding area detected at the head, as a normal person, this is the only movement that occurs on the face. To do this, it has obviously, even normalized of the images once. It can ultimately also do a subtraction between two face images (centered and normalized). If an image has an eye and the other eye closed, it are supposed to appear as a single large "blob", which corresponds to the difference between, pixel-by-pixel. Then, the latter two methods, although faster, are less robust. Data sequences used for the evaluation of our algorithms are derived from pre-acquisition systems of above. The ground truth is generated manually annotating nine points facial features two images for each moment. Their 3D positions (XIV) are then calculated knowing the constraints bipolar and are used to estimate the pose. This one is obtained by a method of minimization of least

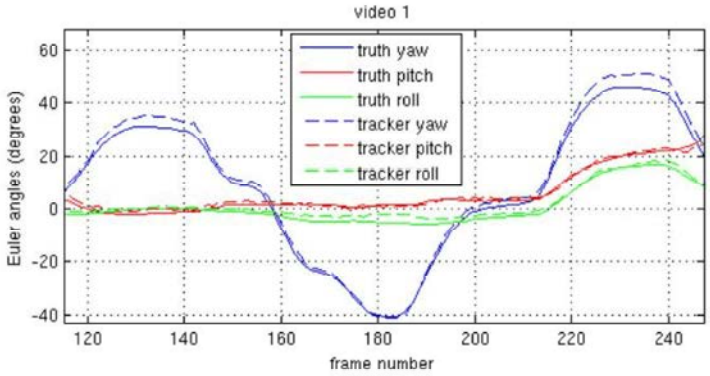
Processing Time: The influence of the number of views used. The calculation time is related to the number of views used to calculate the likelihood particles, we also evaluated the influence of this parameter on the quality of monitoring. The results given in Figure 10 were generated using the algorithm and 2PF 250 particles, for 2, 3 and 4 cameras. With only two views, there is an error on the rise throughout the sequence. Adding a third camera significantly reduces the error, which remains, then less than 2 cm on the first twelve frames. This is due to the fact that by using only the two upper cameras (used pair for performance with two cameras), there is no vertical disparity, leaving uncertainty about the estimate of the poses. Adding a third camera located below the first two to dispel this uncertainty. Taking consideration of the fourth camera only slightly improves performance because the information it carries is redundant in large part with the other views.



Graph 1: Frame Estimation over the tracking location in the Datasets



Graph 2: Ground truth estimation of the XM2VTSDB Database over the input frames in video



Graph 3: Ground truth accuracy of the XM2VTSDB Database over the input frames in video

DISCUSSION

In this study, we addressed the problem of monitoring face and facial gestures in a video sequence. It has presented two approaches based on a comprehensive model of the appearance of the face (facial texture) rather than local primitives. In the first approach, we use, robust statistical outlier pixels to manage due to occlusions. In the second, the model appearance which was given by a Gaussian distribution is replaced by a mixture model. These methods have both the advantage of being flexible since the model Appearance is learned

during the monitoring phase and not previously. They are therefore not specific to a given subject. We were able to treat long video pre-recorded (10 minutes or 15000 images) but also direct purchases on multiple faces in the limit of 4 at 5 frames per second. Quality monitoring, for extracting the position of the head and expressions facial, was obtained by the two methods developed,this even in the presence of occlusions and variations brightness, demonstrating their effectiveness. This article presents a new installation of monitoring methodBased on a 3D model that adapts to the specific face. The three-dimensional knowledge allows to synthesize the views of the head regions, which makes the monitoring more robust in spite of the variations in the orientation of the head (and thus the appearance in images). The precision gain is significant compared to a particulate filter using 2D extracts patches of the initial frame. Furthermore, the use of a particulate filter in a plurality of passes enhances the quality monitoring with calculation time similar. In particular, taking into account some characteristic for manual adjustment of likelihood functions(2PF) achieves better results than the simulated annealing two assists.The purpose of our application is to authenticate persons by means of a three-dimensional reconstruction of their faces, which may be refined over time by adding new information. (Equation 1) in the state cast hidden and filter values over the time. There will be jointly static and dynamic parameters to be evaluated, as done in [12] for simple geometric shapes. Of the due to the increase of the number of variables to be evaluated,new optimization methods particulate filter will be necessary to limit the number of particles.

CONCLUSION

We propose in this paper a stochastic approach to track changes in pose and appearance an almost frontal face in them sequences Clip. This approach based on the principle of particle filtering known as PCA-Wavelet transformation algorithm name in the computer vision field. Particle dynamics are adaptive, guide by a deterministic optimization. The distribution of observations is a derivative active appearance model and an estimated itintegrate robust measures. The outcomes are present's, including taking into account the facial blanking phases. In Secondly, we introduce a second approach consists in replacing the active appearance model by appearance, texture to the constraint of learning debt shelf the appearance has to follow and the problem of sensibilities model in appearance to the reporting requirements of learning images. The proposed method achieves roubustness based on its accurate transformation through wavele analysis.. The XM2VTSDB multi-modal face database is used to compare the real video sequence images. The experimental evaluation shows favarable results and yields 99% detection rate over 15000 video frame images. The frames in the facial tracking data are calculated through various parameters and compared with the ground truth against standard Euler angles of the muti-modal database.

REFERENCES

1. Tong, F., K. Nakayama, M. Moscovitch, O. Weinrib and N. Kanwisher, 2000. Response properties of the human fusiform face area. Cognitive Neuropsychology, 17(1-3): 257-280.
2. Saleh, A.A. and R. Valenzuela, 1987. A statistical model for indoor multipath propagation. Selected Areas in Communications, IEEE Journal on, 5(2): 128-137.
3. Lyons, M., S. Akamatsu, M. Kamachi and J. Gyoba, 1998. Coding facial expressions with gabor wavelets. In Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on IEEE, pp: 200-205.
4. Zhang, Z., M. Lyons, M. Schuster and S. Akamatsu, 1998. Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. In Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on IEEE, pp: 454-459.
5. Nair, P. and A. Cavallaro, 2009. 3-D face detection, landmark localization and registration using a point distribution model. Multimedia, IEEE Transactions on, 11(4): 611-623.
6. Tu, J., H. Tao and T. Huang, 2009. Online updating appearance generative mixture model for meanshift tracking. Machine Vision and Applications, 20(3): 163-173.
7. Abate, A.F., M. Nappi, D. Riccio and G. Sabatino, 2007. 2D and 3D face recognition: A survey. Pattern Recognition Letters, 28(14): 1885-1906.
8. Huang, J., B. Heisele and V. Blanz, 2003. Component-based face recognition with 3D morphable models. In Audio-and Video-Based Biometric Person Authentication Springer Berlin Heidelberg, pp: 27-34.
9. Basu, S., I. Essa and A. Pentland, 1996. Motion regularization for model-based head tracking. In Pattern Recognition, 1996., Proceedings of the 13th International Conference on IEEE, 3: 611-616.

10. La Cascia, M., S. Sclaroff and V. Athitsos, 2000. Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(4): 322-336.
11. Xiao, L., S. Boyd and S.J. Kim, 2007. Distributed average consensus with least-mean-square deviation. Journal of Parallel and Distributed Computing, 67(1): 33-46.
12. Rashida, B. and M.A. Rabbani, 2014. Face Recognition in 3d Facial Feature Using Nearest Neighbor Technique, International Journal of Applied Engineering Research, 9(22).
13. Paysan, P., R. Knothe, B. Amberg, S. Romdhani and T. Vetter, 2009. A 3D face model for pose and illumination invariant face recognition. In Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference On. IEEE, pp: 296-301.
14. Sappa, A.D. and M.A. García, 2000. Incremental multiview integration of range images. In Pattern Recognition, 2000. Proceedings. 15th International Conference on IEEE, 1: 546-549.
15. Goffe, W.L., G.D. Ferrier and J. Rogers, 1994. Global optimization of statistical functions with simulated annealing, Journal of Econometrics, 60(1): 65-99.
16. Zelinsky, A. and J. Heinzmann, 1996. Real-time visual recognition of facial gestures for human-computer interaction. In Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on IEEE, pp: 351-356.
17. Viola, P. and M.J. Jones, 2004. Robust real-time face detection. International Journal of Computer Vision, 57(2): 137-154.
18. Yeh, W.W.G., 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. Water Resources Research, 22(2): 95-108.
19. Grossmann, A. and J. Morlet, 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM Journal on Mathematical Analysis, 15(4): 723-736.
20. Dippel, S., M. Stahl, R.Wiemker and T. Blaffert, 2002. Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform. Medical Imaging, IEEE Transactions on, 21(4): 343-353.
21. Jepson, A. and M.J. Black, 1993. Mixture models for optical flow computation. In Computer Vision and Pattern Recognition, 1993. Proceedings CVPR'93., 1993 IEEE Computer Society Conference on IEEE, pp: 760-761.
22. Hasler, B.S., T. Buecheler and R. Pfeifer, 2009. Collaborative work in 3D virtual environments: A research agenda and operational framework. In Online Communities and Social Computing Springer Berlin Heidelberg, pp: 23-32.