ISSN : 0974 - 7435 Volume 10 Issue 24

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(24), 2014 [16081-16098]

Face alignment based on semi-active appearance model

Yan shi-liang
Engineering and Technology Center, SouthwestUniversity of Science and Technology, Mianyang 621010, (CHINA)
School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, 611731,(CHINA)

ABSTRACT

In the information era, the technology of biological characterrecognition has attracted more and more attentions. In this paper, by investigating theories of active appearance model and inverse compositional image alignment algorithm, we mainly proposed a semi-active appearance model for face alignment based on improving the classical models in the aspects of computation complexity, easily suffering from light, angle and expression, and so on. Firstly, the model of active appearance and the algorithm of alignment are investigated. For the inefficiency of classic gradient descent method in the matching process, the inverse compositional image alignment algorithm is proposed. Then, through combining the active appearance model and Grey Level Co-occurrence Matrix, a novel semi-active appearance model is proposed, which has an simple calculation and higheraccuracy of identification. Finally, experiments were designed to demonstrate the effectiveness of the proposed algorithms.

KEYWORDS

Image fusion; Target recognition; Compressed sensing; Wavelet transform.

INTRODUCTION

In the modern information society, recognitions are often required, such as identity[1-4] and authentication[5-8]. The traditional recognition depends on certification tools, such as IC cards, ID papers, passwords, and so on. Furthermore, in the areas of online bank, e-commerce, and timely communication, user name and passwords are needed for login. Unfortunately, these traditional methods cannot satisfy the requirements of identification system, such as portability, security, effectiveness, and privacy. In a word, the application of these traditional identity authentication technologies suffers from inconvenience and insecurity. From the current view, the best way to solve this problem is biometric identification technology. As the prediction by Bill Gates, the biologicalrecognition technology, which verifies the personal identity using the person's physiological characteristic, will be the important innovation development of IT industry in the future.

Biological recognition technology is founded on pattern recognition, machine vision, and data mining. The specificdefinition is using human physiological characteristics and behavioral characteristics to realize identification. Physical characteristics include fingerprint, iris, face, retina, palm, etc. Behavioral characteristics include speech, gait, signature, etc. The biological recognition technology has improved the traditional recognition mode, which has the properties of universality, uniqueness, stability, and collection. These advantages greatly reduce the possibility of a forge or lost, thus preventing the theft and forgery, which is of great importance in social security, public security, financial security, and human computer interaction. These features have the following definition:universality, uniqueness, and stability. Biometric technology provides people a new way to identity recognition[9-13], which is convenient, stable, not easily lost, not easy to be faked crack etc. In recent years, biological recognition technology has become an international research hot spot. Compared to other biological identification systems, face recognition system has the advantages of f more practical and simple, which mainly reflects the following aspects. First, the sampling is realized in a non-contact manner, in which the contact between human and equipment can be avoided. It can be more easily accepted by the people. Second, the operation of recognition can be done in a hidden way, which does not need the cooperation of the testers. This makes it more practical in the application in catching criminals and security. Third, the cost of face recognition system is much lower than most of other recognition systems, such as iris recognition system, retina system, and fingerprint acquisition. At last, the information of human's expression that can reflect the human's psychological information can be simultaneouslyobtained while recognition, which is cannot be easily realized in other recognition systems.

The research on recognition theory and its application started in 1960 s. Based on some achievementsprovided by researches, the semiautomatic face recognition system has been built. Face recognition research is based on the facial features. For the input face information or video streams, the judgment to determine whether there exists a face or not is should be done at first. If existing, the size of face and the locations of main facial organs need to be determined. Then, according to the above information, extract the essential characteristics. These characteristics will be matched with the known faces in data base to identify the identity of each face. The complete face recognition system includes face detection, data preprocessing, feature extraction, feature matching, and identification. By far, a lot of face alignment algorithm are proposed. According to the difference of input method, these algorithms can be classified with two types, i.e., recognition based on static images and recognition with video streams.

The algorithms based on static images mainly have algorithms based on geometry character, sub linear space algorithms, algorithms with wavelet transform, and recognition based on neural networks. Compared with algorithms based on static images, the algorithms based on video streams have the better performance. In[14] and[15], it has been proved that the dynamic information is helpful for face alignment when the face to be recognized is inverse. Although latter type of algorithm is founded on the static method, it is more difficult to deal with, the main problems of which have two aspects. First, the quality of collected images is poor according to the cost of video camera and some influence of outer environment, which would cause a great impact on recognition performance. Second, as most photographing take place outdoors, the environment of shooting is uncontrollable. As a result, the images may suffer from the changes of attitude, light, and barrier, which will reduce the efficiency of recognition.

In recent years, although the technology of face alignment /recognition has been improved greatly, there are still lots of problems need to be solved, such as the challenge of light changes, different face expressions, and complicated background. In this paper, based on investigating the classical active appearance model and inverse compositional algorithm, we will propose an improved face recognition algorithm based on semi-active appearance model.

OVERVIEW OF FACE ALIGNMENT BASED ON ACTIVE APPEARANCE MODEL

Active apperance model theory

Active appearance model, i.e., AAM, is developed from active shape model, which is firstly proposed by Tim in 2001. The improvement of AMM to active shape model is that it not only uses the shape of face information, but also uses the texture information. In this model, the characters can be used are increased. The AMM algorithm mainly contains two contents: modeling and alignment. Therein, the modeling process includes three parts: shape modeling, texture modeling, and combination modeling.

(a) Shape modeling

In shape modeling, we mainly focus on making a model to describe the outer shape and the key character points of input images. Suppose there are *N* samples of training image. We use *n* character points to describe each sample, which is can be expressed in the vector form:

$$s_i = \left(x_{i1}, x_{i2}, ...x_{in}, y_{i1}, y_{i2}, ...y_{in}\right)^{\mathrm{T}} \quad \textbf{(1)}$$

In (1), s_i represents the shape characteristic vector of ith sample, the jth characteristic point of which can be expressed as (x_{ij}, y_{ij}). Moreover, n stands for the numbers of characteristic point, which usually are assigned with the value of 58. The detailed positions of these points are shown in Figure 1.

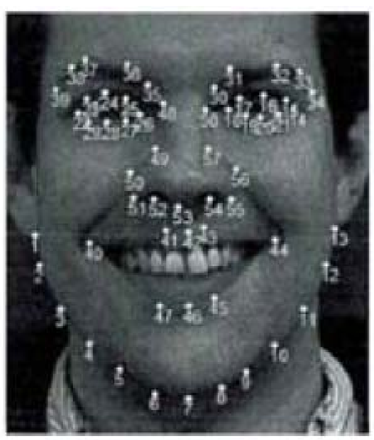

Figure 1: The distributes characteristic ponits of human face

However, the primary shape characteristic points are not in a same template, which makes the differences of position, size, and direction. The algorithm of Procrustes analysis should be used to translate them into a uniform coordinate system. After uniform operation, the PCA algorithm is needed for reduction dimension, the calculation process of which is shown as following. First, calculate the average vector and covariance matrix of shape vectors of *N* samples, i.e.:

$$\bar{s} = \frac{1}{N} \Sigma_{i=1}^{n} s_i \tag{2}$$

$$C = \frac{1}{N} \Sigma_{i=1}^{n} \left(s_i - \bar{s}\right)\left(s_i - \bar{s}\right)^{\mathrm{T}} \tag{3}$$

Then, through eigenvalue analysis of the above equation, the eigenvalue and eigenvector can be obtained, i.e.,$(\lambda_1, \lambda_2, ...\lambda_{2n})$, $(p_1, p_2, ...p_{2n})$. From the theory of PCA, the larger eigenvalue means represents that the more information of human face shape is included in the corresponding eigenvector. Thus, we select the *J* numbers of larger eigenvectors to describe the shape information of human face, i.e.:

$$\frac{\Sigma_{i=1}^{J} \lambda_i}{\Sigma_{i=1}^{2n} \lambda_i} \geq t \quad \left(0 \leq t \leq 1\right) \tag{4}$$

In (4), *t* is usually assigned the value of 0.95, which is enough to describe the most information of human shape. Then, an arbitrary human face can be expressed in a vector form:

$$\begin{aligned} s &= \bar{s} + \left(p_1, p_2, ...p_J\right) b_s \\ &= \bar{s} + \mathbf{p}_s b_s \end{aligned} \tag{5}$$

The vector b_s stands for the parameters of human face. A new vector of face shape can be available by changing the value of b_s . Of course, the value range should be restricted to maintain the efficiency of face shape. Due to PCA theory and 3σ principle, the value range of each component of b_s should be maintained in $(-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i})$.

(b) Texture modeling

The texture of AMM $A(u)$ indicates the pixel points in mapping region of shape model, which also represents the combination of shape and texture of face image. The particular method to get texture information is founded on texture mapping algorithm in computer graph theory. After shape modeling, the average shape $\bar{s}$ can be obtained. By Delaunay triangle dissection, the benchmark grid is available, as shown in Figure 2.

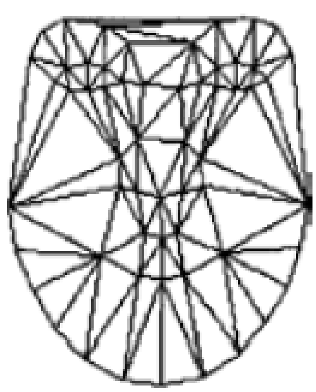

Figure 2: Delaunay triangle dissection of average shape

Then by a sublevel affine operation, the pixel points are assigned on the proper positions of input human face image, the examples of which are shown if Figure 3. Using $W(u, p)$ to stands for the sublevel affine operation, the process of which is translating the points $u = (x, y)$ in the average shape onto the target image, and p is weighting coefficient. Then make the PCA dimension reduction on pixel again to get the average texture $A_0(u)$. The same as shape modeling method above, the texture model can be expressed as:

$$A(u) = A_0(u) + \Sigma_{i=1}^{J} b_i A_i \tag{6}$$

In this equation, A_i is the eigenvector of $A(u)$, and the component b_i represents model parameter. Given a group of model parameters, an example model can be determined.

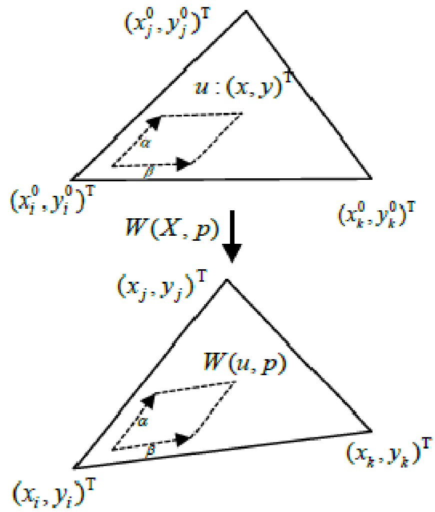

Figure 3: Sublevel affine operation

As shown in Figure 3, the main calculation process of sublevel affine operation can be described as follows. Suppose that the average shape is obtained. After Delaunay triangle dissection, usedata (x_i^0, y_i^0), (x_j^0, y_j^0) and (x_k^0, y_k^0) to represent vertex coordinate of a triangle. Then, every point coordinate in the triangle can be expressed with these data:

$$\begin{aligned}(x,y)^{\mathrm{T}} &= (x_i^0, y_i^0)^{\mathrm{T}} + \alpha\left[(x_j^0, y_j^0)^{\mathrm{T}} - (x_i^0, y_i^0)^{\mathrm{T}}\right] \\ &\quad + \beta\left[(x_k^0, y_k^0)^{\mathrm{T}} - (x_j^0, y_j^0)^{\mathrm{T}}\right]\end{aligned} \tag{7}$$

In (7), the value of α and β can be calculated using the values of vertex coordinate.

(c) Combination modeling

After complete the modeling process of shape and texture, we can find that an arbitrary face image can be expressed by shape parameter b_s and texture parameter b_i. In order to get a simpler and precise expression and reduce the correlations between shape and texture, PCA is used again to combination the shape parameter and texture parameter:

$$b = \begin{pmatrix} \Lambda b_s \\ b_i \end{pmatrix} = \begin{pmatrix} \Lambda P_s^{\mathrm{T}}(s - \bar{s}) \\ P_i^{\mathrm{T}}(A_i - A_0) \end{pmatrix} \tag{8}$$

In this equation, the parameter Λ is a diagonal scaling matrix that is used to omit the difference of dimensions between shape and texture parameters. Reduction the dimension of parameter b in combination model by PCA, the shape model and texture model in the combination form can be expressed as:

$$\begin{cases} b = \mathbf{p}_c c \\ \mathbf{p}_c = \left(\mathbf{p}_{c,s} \quad \mathbf{p}_{c,t}\right)^{\mathrm{T}} \\ s = \overline{s} + \mathbf{p}_s \Lambda^{-1} \mathbf{p}_{c,s} c \\ A(u) = A_0(u) + \mathbf{p}_t \mathbf{p}_{c,t} c \end{cases} \tag{9}$$

(d) Model alignment

Alignment calculation is an important part in face recognition based on AMM. Essentially, the process of modeling of face is location the characteristic points of face image. If we want to determine that the input face is the given face or the face in data base, the alignment process is needed, which is the precondition to obtain a satisfaction performance. In model alignment, the subtraction operation between the input image $I(u)$ and template $A(u)$ is needed at first. Use δI to stand for the difference between them, i.e.:

$$\delta I = I_i - I_m \tag{10}$$

In (10), I_i is the shape expression of input image and I_m represents the shape information of template face image. We change the parameter c in (9) to minimize the value of $\|\delta I\|^2$. When this minimum value $\|\delta I\|^2$ is less than a given threshold, then the two faces are the same. The essential process is an optimizing problem. As the number of parameters in AAM is much more, it needs a complex calculation. Fortunately, theory analysis and a lot of experiments have both shown that there exists a linear relationship between δI and δc in a certain range in AAM. Then the multivariate regression method can be used to solve this problem in an easier way:

$$\delta c = A\delta I = A\left(I_i - I_m\right) \tag{11}$$

Inverse compositional alignment algorithm

The AAM and the similar model to it are types of generated models based on visual phenomenon. Although the shape and appearance is linear, generally, AMM is a nonlinear model of pixel grey. In the modeling process, the number of arguments is large, and in the alignment calculation, the classic gradient descent method that used to solve the nonlinear optimizing problem causes much lower alignment efficiency. In order to improve the efficiency, we will focus on inverse compositional alignment algorithm and its application in AAM. This algorithm changes the updating process of parameters in matching, which will greatly reduce the amount of calculation and improve the efficiency of alignment significantly.

(a) Generation of model

Suppose the parameter vector of shape is defined by b, and denote the vector of appurtenance model by λ:

$$\begin{cases} b = \left(b_1, b_2, \ldots b_n\right)^{\mathrm{T}} \\ \lambda = \left(\lambda_1, \lambda_2, \ldots \lambda_n\right)^{\mathrm{T}} \end{cases} \tag{12}$$

According to (12), the eigenvector of shape s and the appearance $A(u)$ based on benchmark shape s_0 can be obtained. In order to get the particular position of characteristic points of the average shape under the corresponding parameters, the affine transform of s_0 is needed. Define the affine transform by $W(u,p)$. As the parameter s represents the face shape after reducing translation, rotation, and scaling under model framework, the standard transform of s is needed to get the actual face shape in actual coordinate. Suppose the rotation angle is θ, the position is $t=(t_x,t_y)$, and the scaling is k, then the parameter vector of attitude is $q=(a,b,t_x,t_y)$, $a = k\cos\theta - 1$, and $b = k\sin\theta$. For a point (x,y), the standard translation function $N(x,q)$ can be expressed as:

$$N(x,q)=\begin{pmatrix}1+a & -b\\ b & 1+a\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+\begin{pmatrix}t_x\\ t_y\end{pmatrix} \tag{13}$$

Therefore, the final face shape can be expressed as $N(W(u,p);q)$. For an arbitrary point u_0 in shape s_0, the final position of it in human face is $N(W(u_0,p);q)$, with the texture value of $A(u_0)$. So, the final model of AAM can be denoted as:

$$M\left(N\left(W\left(u,p\right),q\right)\right)=A(u)\ \ \forall u\in s_0 \tag{14}$$

(b) Problem Set

Define the error between the input image and image model as $E(u)$:

$$\begin{aligned}E\left(u\right)&=A_0\left(u\right)+\Sigma_{i=1}^{m}\lambda_i A_i\left(u\right)-I\left(N\left(W\left(u,p\right),q\right)\right)\\ &=E\left(b,\lambda,q\right)\end{aligned} \tag{15}$$

The alignment process of AAM is a nonlinear optimizing problem; the traditional method to solve this problem is based on gradient descent, which will reduce the alignment efficiency. In[16], a improved algorithm is proposed, in which it suppose that the relationship of $E(u)$ and $(\Delta b,\Delta\lambda)$ is linear, i.e., $(\Delta b,\Delta\lambda)=ME(u)$, in which the matrix M can be obtained by multivariate regression method. As the linear relationship only holds in a small range, the pyramid search method is used by Cootes, the principle of which is that, search in a larger range at the beginning, and then reduce this range until getting the minimum value. This algorithm can get a satisfied alignment performance. However, the convergence rate is slow, and it did not consider the impact of attitude parameter q on the algorithm performance. To solve this problem, an improved algorithm was proposed in[16], which added a precondition that the gradient matrix is constant near the optimal value. Because the gradient matrix is correlative with some particular region in texture space, this gradient matrix is not the best estimate of the real value when the target texture is disagreement with the corresponding region, as shown in[17].

For the shortages in the above algorithm, replacing the estimate of gradient matrix with a better value that can be strictly proved in mathematics, and improving the alignment efficiency are still needed further research. As the inverse compositional alignment algorithm can avoid calculating the image gradient repeatedly, Jacobianmatrix, and Hessian matrix, the shortages in above algorithm can be solved with application of inverse compositional alignment in AAM alignment calculation.

(c) Steps of inverse compositional alignment

(A) Pretreatment

(1) Calculate gradient ∇A_0 of the template image A_0;

(2) Calculate Jacobianmatrix $\partial W/\partial p$, $\partial N/\partial p$ at $(u,0)$;

(3) Calculate the fastest image $SD_j(u)$:

$$\begin{cases} SD_j(u) = \nabla A_0 \dfrac{\partial N}{\partial q_j} - \Sigma_{i=1}^{m}\left[\Sigma A_i(u)\cdot\nabla A_0 \dfrac{\partial N}{\partial q_j}\right]A_i(u) \\ SD_{j+4}(u) = \nabla A_0 \dfrac{\partial W}{\partial p_j} - \Sigma_{i=1}^{m}\left[\Sigma A_i(u)\cdot\nabla A_0 \dfrac{\partial W}{\partial p_j}\right]A_i(u) \end{cases} \quad \textbf{(16)}$$

(4) Calculate Hessian matrix:

$$H_{j,k} = \Sigma_{u\in s_0} SD_j(u)\cdot SD_k(u) \quad \textbf{(17)}$$

(B) Iterative process

(1) Calculate $I\left(N\left(W\left(u,p\right),q\right)\right)$;

(2) Calculate the distance of the transformed input image and the template image:

$$\Delta\Omega = I\left(N\left(W\left(u,p\right),q\right)\right) - A_0(u) \quad \textbf{(18)}$$

(3) Calculate Ξ

$$\Xi = \Sigma_{u\in s_0} SD_j\left(u\right)\cdot\left[I\left(N\left(W\left(u,p\right),q\right)\right) - A_0(u)\right] \quad \textbf{(19)}$$

(4) Calculate Δq and Δp:

$$\left(\Delta q, \Delta p\right) = H^{-1}\Xi \quad \textbf{(20)}$$

(5) Update the parameters:

$$N\left(W\left(u,p\right),q\right) \leftarrow N\left(W\left(u,p\right),q\right)\circ N\left(W\left(u,\Delta p\right),\Delta q\right)^{-1} \quad \textbf{(21)}$$

Challenges of current algorithm

In recent years, the technology of human face alignment is improved greatly. A lot of excellent algorithms have been proposed by researchers. At the same time, a lot of enterprises increased investment in biology area, especially in human face alignment, which led a large amount of human face recognition system in business and other application fields. But there are still lots of technologyproblems in human face recognition system; for example, these systems are easilysufferedfrom the changes of light, expression of human face, and so on.

The active shape model and active appearance model are both founded on classic active contour model, which are widely used in the current face recognition systems. These models can provide a good description of the shape and texture of human face. Unfortunately, as the numbers of parameters of shape and texture in these models are much large, the alignment efficiency is much low. Moreover, the algorithm may easily converge to a wrong point, which will lead a wrong recognition result. Furthermore, these models are also sensitive to the changes of light, which will cause a bad performance. Therefore, more researches are needed to solve these problems.

FACE ALIGNMENT BASED ON SEMI-ACTIVE APPEARANCE MODEL

Modeling of semi-active appearance model

To solve the problems indicated above in the traditional active appearance model, an improved model called semi-active appearance model, i.e., SAAM will be proposed in this section. This model is founded on the combination of active shape model and Grey Level Co-occurrence Matrix[18 - 20]. As there are dynamic parameters in active shape model, and has non dynamic parameter in Grey Level Co-occurrence, the model proposed in this paper can avoid processing a large amount of dynamic parameters, which is the reason that we call it semi-active appearance model. Of course, this model can be also used in CT image location, gesture recognition, and other applications. In the alignment process, the same as the application of inverse compositional alignment in AAM, we also study the application of it in model.

(a) Extraction of characteristic of texture

The texture is a description of the distribution of pixel grey. In some images, such as bricks and grid windows, there are several images appearing in periodically, which are usually called determination texture image. Otherwise, the images appearing not in a periodical way are called stochastic texture images, such as finger print and human face.

The characteristic of texture is a statistic of local property of image. From the experience of human, the texture is the main differences among different images. In these years, researchers have established lots of texture algorithms to measure the properties of texture, which can be mainly divided into two classes, stochastic analysis and structure analysis. The main algorithms for texture analysis at present are based on stochastic, such as Grey Level Co-occurrence Matrix and Laws texture measurement.

Suppose that the characteristic of texture to be analyzed is a matrix, in which there are N_c pixels in the horizontal direction and N_r pixels in the vertical direction. Then divide all the pixels in the texture images into N layers. Use Z_c to express the horizontal space, i.e., $Z_c = \{1,2,...,N_c\}$, use Z_r to describe the vertical space, i.e., $Z_r = \{1,2,...,N_r\}$, and use $G = \{1,2,...,N_q\}$ to stand for the quantization grey set. Then the function of an image can be defined as: assign each pixel in the image a value of G in grey set $f: Z_r \times Z_c \to G$.

In the texture image, as the stochastic disciplinarian of the pair of pixels with a particular distance in some direction can express the characteristic of image texture, the Grey Level Co-occurrence Matrix can provide a good description of this stochastic disciplinarian. The Grey Level Co-occurrence Matrix is an indeed description of a matrix under an emergencefrequency with the direction of θ, the distance of d, and the grey level of i and j, any an element of which can be expressed as $P(i, j \mid d, \theta)$. If the parameters d and θ are determined, this matrix can be also expressed as $P_{i,j}$:

$$P(i, j \mid d, \theta) = \#\left\{\left[(k,l),(m,n) \mid f(k,l) = i; f(m,n) = j\right]\right\} \quad \textbf{(22)}$$

In (22), $(k,l),(m,n) \in (Z_r \times Z_c)$, the item $f(k,l) = i$ means that the position in matrix of the point with grey level i. The numbers of grey between i and j with distance of d and direction of θ are denoted by $\#$.

As the difference exiting in texture scale, there has a large difference in Grey Level Co-occurrence Matrix among different kinds of images. Even for images with the same kind, the differences still exist. The most usually used parameters to describe the characteristic are shown as follows:

(A) Energy

$$E(d, \theta) = \Sigma_{i,j}\left(P(i, j \mid d, \theta)\right)^2 \quad \textbf{(23)}$$

The energy can give an expression of the uniformity of texture change. Also it can measure the coarseness of texture and distribution of grey.

(B) Entropy

$$H(d, \theta) = -\Sigma_{i,j} P(i, j \mid d, \theta) - \log\left(P(i, j \mid d, \theta)\right) \quad \textbf{(24)}$$

The entropy of image can describe the content of information, also represents the confusion degree of images.

(C) Inertia Moment

$$I(d,\theta) = -\Sigma_{i,j}(i-j)^2 P(i,j \mid d,\theta) \tag{25}$$

The inertia moment stands for distance between the elements in Grey Level Co-occurrence Matrix and the main diagonal.

(D) Correlation

$$C(d,\theta) = \frac{\Sigma_{i,j}(i-u_x)(j-u_y)P(i,j \mid d,\theta)}{\sigma_x \sigma_y} \tag{26}$$

(E) Local balanced

$$L(d,\theta) = \Sigma_{i,j}\frac{1}{1+(i-j)^2}P(i,j \mid d,\theta) \tag{27}$$

$$\begin{cases} u_x = \Sigma_i i \Sigma_j P(i,j \mid d,\theta) \\ u_y = \Sigma_j j \Sigma_j P(i,j \mid d,\theta) \\ \sigma_x = \Sigma_i (i-u_x)^2 \Sigma_j P(i,j \mid d,\theta) \\ \sigma_y = \Sigma_j (j-u_y)^2 \Sigma_j P(i,j \mid d,\theta) \end{cases} \tag{28}$$

As the active shape model can give a satisfied description of shape of images, and Grey Level Co-occurrence Matrix has the abilities to counteract the changes of light also with a lower computation burden, in this section, we will describe the shape of image by computing the Grey Level Co-occurrence Matrix under shape alignment to establish the semi-active appearance model.

First, extract the shape characteristic of face image. For the face contour, eyes, nose, mouse, and other key components, the PDM model is used in our algorithm. In order to reduce the influence of position, size, and angle, we unify the face alignment to a uniform coordinate framework with Procrustes analysis. Then model the shape characteristic with active shape model to obtain the average face model and the model expression of each face sample. Make an affine transform of face sample onto the average shape. At last, use the texture characteristic and the corresponding pixel matrix to calculate the Grey Level Co-occurrence Matrix, which is used to describe the outer shape of images.

Modeling of local semi-active appearance model

In section A, the face shape is described by a Grey Level Co-occurrence Matrix under the benchmark alignment grid, which mainly represents the overall relationship of face image grey. But if light on some parts of the image is not uniform, there will be a big difference between thecorrelations of grey in input face and the standard face image, which will cause a bad performance of the semi-active appearance model, as shown in Figure 4.

Therefore, the semi-active appearance model needs to be improved according to the above problem. In order to solve the problem of non-uniform light, we divide the face in benchmark grids into four parts, as shown in Figure 5.

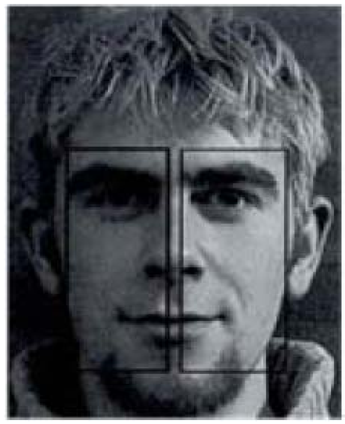

Figure 4: Human face under non-uniform light

Figure 5: The divided face image under benchmark grids

After the divided operation, semi-active appearance model is used to modeling each part of the image, in which the grey level is decided by the pixel in each part. Through this operation, the influence of non-uniform light on human face image can be reduced. The model can be expressed as:

$$M\left(A(u)\right)=\begin{pmatrix} G_1\left(A_1\left(u\right),d,\theta\right) & G_2\left(A_2\left(u\right),d,\theta\right) \\ G_3\left(A_3\left(u\right),d,\theta\right) & G_4\left(A_4\left(u\right),d,\theta\right) \end{pmatrix} \quad \textbf{(29)}$$

This is a partitioned matrix, each element of which represents the model of corresponding part in human face. In this local semi-active appearance model, it not only has considered the overall correlations of image grey, but also has considered the local correlations. Next, we will discuss the problem of alignment calculation in semi-active appearance model.

Similar to the alignment calculation in active appearance model, the alignment process in semi-active appearance model can be also treated as an optimizing problem. To determine the identification of input face needs alignment calculating process, which is the important precondition to produce a right recognition result. The principle of alignment in semi-active appearance model is making a subtract operation of input image $I(u)$ and the template image $T(u)$, and finding out the minimum value of matrix norm, and getting the parameters that can minim the norm:

$$\arg\min\left\|G\left(T\left(u\right),d,\theta\right)-G\left(I\left(W\left(u,p\right)\right),d,\theta\right)\right\|_F^2 \quad \textbf{(30)}$$

In (30), the main unknown parameter is p. The values of θ and d can be determined at the beginning. Of course, they can also be seemed as the unknown parameters, which will increase the computation. This is an optimizing problem without any constraint, which can be solved by gradient descent method.

In the alignment process, the simple updating strategy based on affine transform is adopted, i.e.:

$$p \leftarrow p + \Delta p \tag{31}$$

In (31), we suppose that the difference of input image shape and template image shape is linear with Δp, i.e.:

$$\begin{cases} \Delta p = R \cdot r(p) \\ r(p) = G_T - G_I \end{cases} \tag{32}$$

In (32), Δp represents a little change on p, and R stands for the linear relationship between parameter G_I and parameter $r(p)$, which can be regarded as a fixed gradient matrix and can be estimated by multivariate regression method.

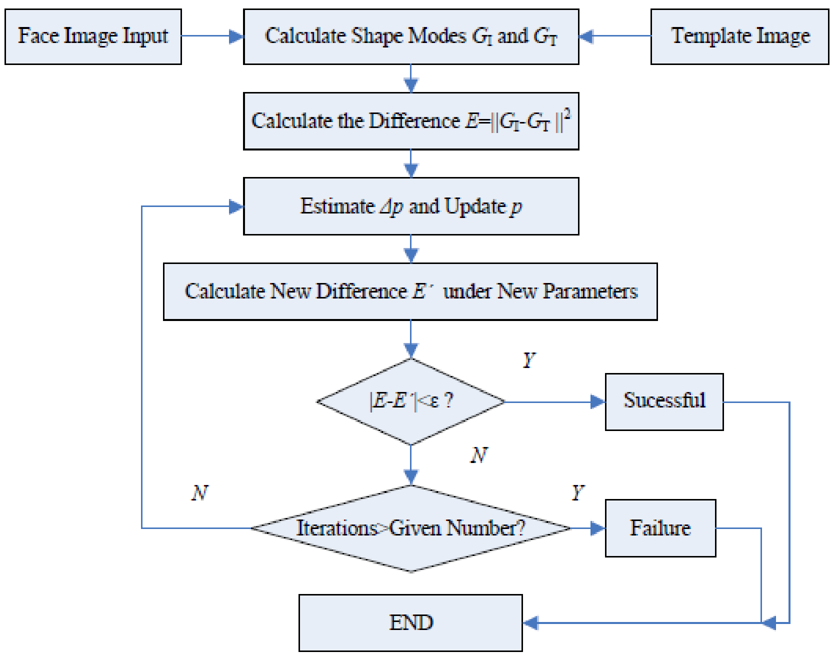

Figure 6: The flowchart of alignment process

As the value of R can be estimated at the beginning, thus the process of alignment calculating can be simplified as an iterative process, the steps of which are shown as follows:

(1) Take texture samples of input human face image and template image, and calculate their shape models G_I and G_T .

(2) Calculate the difference of image G_I and template image G_T . By calculating the matrix norm of this difference, evaluate the alignment precise, i.e., $E = \| r \|_F^2$.

(3) Estimate Δp and update p .

(4) Calculate the shape G_I' and G_T' under new parameters.

(5) Calculate the difference $E' = \| G_I' - G_T' \|_F^2$.

(6) If the condition $| E' - E | < \varepsilon$ is satisfied, then the two faces belong to the same people, otherwise go to step 3.

(7) If the difference above is still larger than the threshold value when the iterations exceed a given number, then the decision that the two faces belong to different persons is made.
The flowchart of above algorithm is shown in Figure 6.

The application of inverse compositional in SAAM

As shown in section II, the inverse compositional can be used for solving the problem of alignment between input and template images. As indication in previous content, the shape of template image can be described by (d,θ), this is alterable. But the change of them will greatly increase the computation. Usually, we can choose some typical values, such as $(d=0,\theta=0°,45°,90°,135°)$. Using the variable values will improve the recognition result but also will reduce the efficiency. How to balance the two aspects is still needs further researches.

The optimizing problem of alignment in SAAM can be translated to make the follow item take the minimum value:

$$\left\| G\left(T\left(u\right),d^{u},\theta^{u}\right)-G\left(I\left(N\left(W\left(u,p\right),q\right),d^{u},\theta^{u}\right)\right)\right\|_{F}^{2} \quad \textbf{(33)}$$

Solving this optimizing problem by inverse compositional alignment algorithm usually needs two changes. First, the transforms of input and template images are should be done to minimize the following value:

$$\left\| G\left(I\left(N\left(W\left(u,p\right),q\right),d^{u},\theta^{u}\right)\right)-G\left(T\left(W\left(u,\Delta p\right)\right),d^{u},\theta^{u}\right)\right\|_{F}^{2} \quad \textbf{(34)}$$

Second the updating strategy is improved. For the introduced similarity transformationfunction $N(u,p)$, the updating process is expressed as:

$$N\left(W\left(u,p\right),q\right)\leftarrow N\left(W\left(u,p\right),q\right)\circ N\left(W\left(u,\Delta p\right),\Delta q\right)^{-1} \quad \textbf{(35)}$$

As $W(u,p)$ represents a piecewise affine transform, the condition of affine set should be group domain can be satisfied, so $N(W(u,p),\nabla q)^{-1}$ cannot be calculated directly. From[17], we have the following equation:

$$\begin{aligned} &N\left(W\left(u,p\right),q\right)\circ N\left(W\left(u,\Delta p\right),\Delta q\right)^{-1} \\ &\approx N\left(W\left(u,p\right),q\right)\circ N\left(W\left(u,-\Delta p\right),-\Delta q\right) \end{aligned} \quad \textbf{(36)}$$

The parameters Δp and Δq can be calculated as follows:

$$\left(\Delta p,\Delta q\right)=H^{-1}SD_{j}\left(u\right)\left[I\left(N\left(W\left(u,p\right),q\right)\right)-T\left(u\right)\right] \quad \textbf{(37)}$$

$$\begin{cases} SD_{j}\left(u\right)=\nabla T\dfrac{\partial N}{\partial q_{j}}-\left[T(u)\cdot\nabla T\dfrac{\partial N}{\partial q_{j}}\right]T(u) \\ SD_{j+4}\left(u\right)=\nabla T\dfrac{\partial W}{\partial p_{j}}-\left[T(u)\cdot\nabla T\dfrac{\partial W}{\partial p_{j}}\right]T(u) \end{cases} \quad \textbf{(38)}$$

After calculation of Δp and Δq, calculate the shape of s_0 under $N(W(u,p),q)\circ N(W(u,-\Delta p),-\Delta q)$, i.e., s^{+}, and then according the following equation to update Δp and Δq.

$$N\left(W\left(u,p\right),q\right)=s^{+} \quad \textbf{(39)}$$

The computation of s^{+} can be decomposed as two steps. First, calculate the shape of s_0 under $N(W(u,-\Delta p),-\Delta q)$, which can be expressed as $N(W(s_0,-\Delta p),-\Delta q)$.

$$W\left(s_0,-\Delta p\right)=s_0-\Sigma_{i=1}^{n}\Delta p_i s_i \quad \textbf{(40)}$$

Then translate $N(W(s_0,-\Delta p),-\Delta q)$ using the current affine transform $W(u,p)$ and current standard transform $N(u,p)$.

EXPERIMENT TEST AND VALIDATION

Image fusion based on compressed sensing

In order to test face alignment algorithm proposed in this paper, we mainly investigate the semi-active appearance model and the application of inverse compositional. Our experiments are based on the images in IMM face database. In IMM database there are lots of face images from 40 different people. For each person, there are 6 face images from different angles, i.e., there are totally 240 face images.

At first, making the active appearance model and the semi-active for 120 face images from 40 persons, in which 58 characteristic points are marked. Then the 120 remained images are used to test the performance. First, we use the gradient descent method for alignment, and then the inverse compositional alignment algorithm is used. The standard test images with light and angle changes are used.

The experiment results with original active appearance model and semi-active appearance model are shown in Figure 7 and Figure 8.

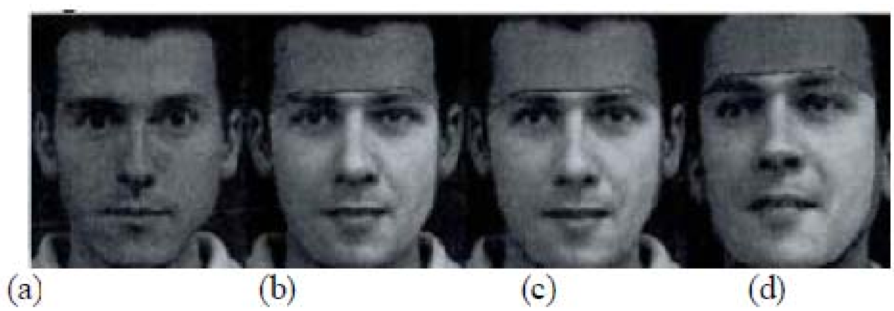

Figure 7: (a) Original Image, (b) AAM Model, (c) Gradient Descent Alignment. (d) Inverse Compositional Alignment. Results with active appearance model

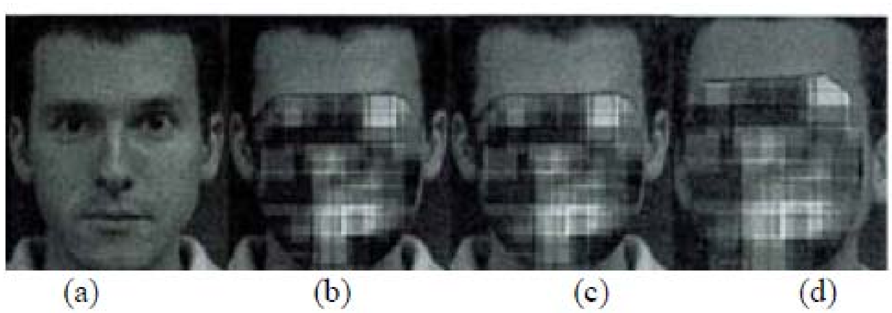

Figure 8: (a) Original Image, (b) SAAM Model, (c) Gradient Descent Alignment. (d) Inverse Compositional Alignment. Results with semi-active appearance model

In TABLE 1, the alignment computational time and recognition accuracy with active appearance model (AAM) and semi-active appearance model (SAAM) by gradient descnt (GD) alignment and inverse compositional (IC) alignment are listed.

In TABLE 1, *N* stands for the number of samples, ME is the shape error rate, which is defined as follows in (40), and the demension of computation time is millisecond.

$$ME = \left\| G_I - G_T \right\|_F / G_T \tag{41}$$

TABLE 1: Composition of recognition results

Algorithm		AMM-GD		SAAM-GD		AMM-IC		SAAM-IC	
		ME	TIme	ME	Time	ME	Time	ME	Time
N	10	0.1	530	0.08	350	0.1	130	0.07	68
	20	0.11	602	0.09	390	0.11	150	0.07	70
	30	0.13	605	0.1	420	0.12	188	0.08	75
	40	0.15	780	0.13	478	0.14	243	0.09	78

Next, we will show the experiment with light changes, the resluts of which are shown in Figure 9, Figure 10, and TABLE 2.

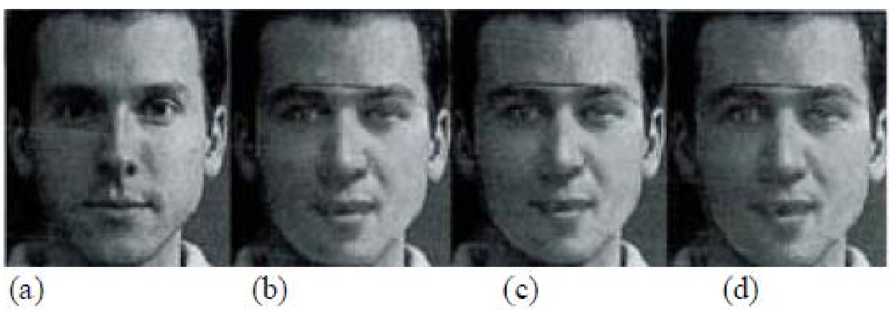

Figure 9: (a) Original Image, (b) AAM Model, (c) Gradient Descent Alignment. (d) Inverse Compositional Alignment. Results with active appearance model under light changes

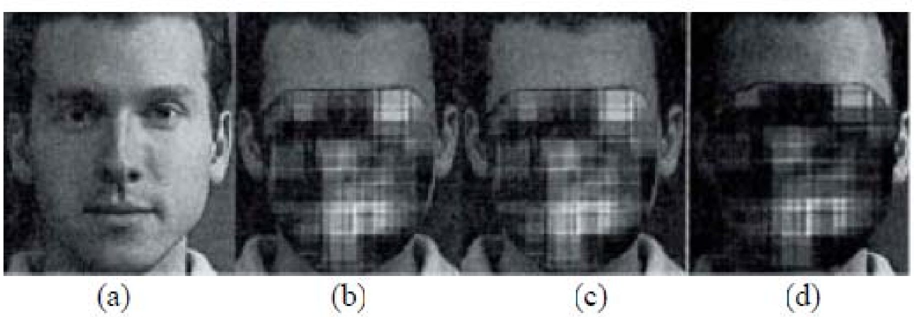

Figure 10: (a) Original Image, (b) SAAM Model, (c) Gradient Descent Alignment, (d) Inverse Compositional Alignment. Results with semi-active appearance model under light changes

TABLE 2: Composition of recognition results under light changes

Algorithm		AMM-GD		SAAM-GD		AMM-IC		SAAM-IC	
		ME	TIme	ME	Time	ME	Time	ME	Time
N	10	0.15	543	0.08	350	0.15	139	0.08	67
	20	0.18	622	0.09	390	0.17	161	0.09	71
	30	0.21	663	0.1	420	0.20	190	0.11	74
	40	0.25	798	0.13	478	0.23	252	0.12	78

Finnally, we will show the experiment with angel changes, the resluts of which are shown in Figure 11, Figure 12, and TABLE 3.

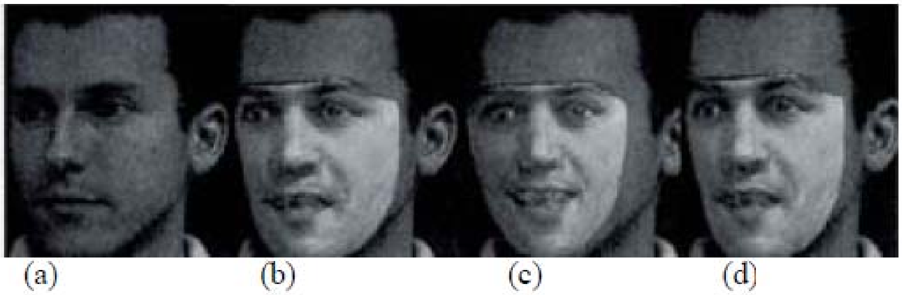

Figure 11: (a) Original Image, (b) AAM Model, (c) Gradient Descent Alignment, (d) Inverse Compositional Alignment. Results with active appearance model under angle changes

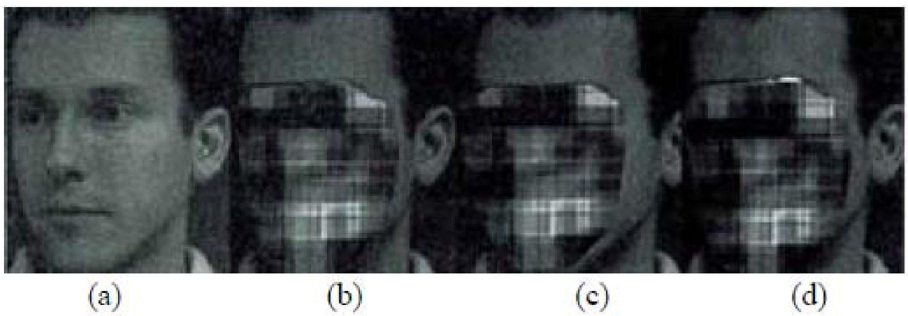

Figure 12: (a) Original Image, (b) SAAM Model, (c) Gradient Descent Alignment, (d) Inverse Compositional Alignment. Results with semi-active appearance model under angle changes

TABLE 3: Composition of recognition results under angle changes

Algorithm		**AMM-GD**		**SAAM-GD**		**AMM-IC**		**SAAM-IC**	
		ME	**TIme**	**ME**	**Time**	**ME**	**Time**	**ME**	**Time**
N	10	0.16	545	0.08	348	0.16	136	0.08	66
	20	0.19	624	0.09	388	0.18	158	0.09	69
	30	0.21	657	0.10	426	0.22	187	0.11	72
	40	0.26	792	0.13	469	0.25	249	0.12	76

From the results, we can see that either with gradient descent or inverse compositional algorithms, the semi-active model that is proposed in this paper has provided a better alignment performance. And this conclusion also holds even if the light or angle changes.

CONCLUSION

Active appearance model is an important computer visual method, in which the deformable targets can be described with fewer characteristics. This model has been widely used in the application of pattern recognition and computer technology, such as gesture recognition and face alignment.

For the low alignment efficiency in the traditional active appearance model, we propose an improved semi-active appearance model based on the investigation of active appearance model and inverse compositional alignment algorithm. This model needs less computation, and can overcome the influences light, angle, and expression. First, we made a deeper research on active appearance model, in which includes the modeling of shape model, texture model and combination model. The classic gradient descent algorithm in active appearance model is to solve a nonlinear optimizing problem, in which the eigenvalue, eigenvector and gradient are needed to be calculated. As a result, this model is easily suffered from the changes of light, angle, and expressions. Then, the inverse compositional alignment algorithm is studied, which can be used in active appearance model. This application in active appearance model will improve the computation efficiency and accuracy of alignment. And then based on the active appearance model, an improved model called semi-active appearance model, which is a combination of active shape model and Grey Level Co-occurrence Matrix, in which the eigenvalue, eigenvector, and other parameters under the average texture image are not needed. Moreover, the inverse compositional alignment algorithm is used the proposed semi-active appearance model, which can improve the recognition efficiency further.

However, the theory of semi-active appearance model is still in the developing stage, it needs more exploration in many application areas. Furthermore, the semi-active appearance model is the combination of active appearance model and grey Level Co-occurrence Matrix, in which the detailed shape and texture of face image can be recovered as the grey Level Co-occurrence Matrix represents the correlation of overall grey. As a result, this model can be directly used in expression analysis, which needs further research.

ACKNOWLEDGEMENT

A Project Supported by Scientific Research Fund of SiChuan Provincial Education Department 13ZB0176.

REFERENCES

[1] N.S.Nourabadi, K.G.Dizaji, S.A.Seyyedsalehi; "Face pose normalization for identity recognition using 3D information by means of neural networks,"5th Conference on Information and Knowledge Technology, doi: 10.1109/IKT.2013.6620106, 432 – 437 **(2013)**.

[2] C.Ching-Kun, L.Chun-Liang, L.Shyan-Lung Lin, C.Yen-Ming, C.Cheng-Tang; "A Chaotic Theorectical Approach to ECG-Based Identity Recognition," IEEE Computational Intelligence Magazine, doi: 10.1109/MCI.2013.2291691, **9**, 53 – 63 **(2014)**.

[3] L.Jiwen Lu, W.Gang Wang, P.Moulin; "Human Identity and Gender Recognition From Gait Sequences With Arbitrary Walking Directions," IEEE Transactions on Information Forensics and Security, doi: 10.1109/TIFS.2013.2291969, **9**, 51 – 61 **(2014)**.

[4] L.Feng Liu, D.Zhang, G.Zhenhua; "Distal-Interphalangeal-Crease-Based User Authentication System," IEEE Transactions on Information Forensics and Security, doi: 10.1109/TIFS.2013.2272787, **8**, 1446 – 1455 **(2013)**.

[5] M.S.Umar, M.Q.Rafiq; "Select-to-Spawn: A novel recognition-based graphical user authentication scheme," 2012 IEEE International Conference on Signal Processing, Computing and Control, doi: 10.1109/ISPCC.2012.6224382,1-5 **(2012)**.

[6] A.Drosou, D.Tzovaras, K.Moustakas, M.Petrou; "Systematic Error Analysis for the Enhancement of Biometric Systems Using Soft Biometrics," IEEE Signal Processing Letters, doi: 10.1109/LSP.2012.2221701, **19**, 833-836 **(2012)**.

[7] P.Tsai, M.K.Khan, J.Pan, B,Liao; "Interactive Artificial Bee Colony Supported Passive Continuous Authentication System," Systems Journal, doi: 10.1109/JSYST.2012.2208153, 395 – 405 **(2014)**.

[8] N.N.Erdogmus, J.L.Dugelay; "3D Assisted Face Recognition: Dealing With Expression Variations," IEEE Transactions on Information Forensics and Security, doi: 10.1109/TIFS.2014.2309851, **9**, 826 – 838 **(2014)**.

[9] S.J.Elliott, S.A.Massie, M.J.Sutton; "The Perception of Biometric Technology: A Survey," 2007 IEEE Workshop on Automatic Identification Advanced Technologies, doi: 10.1109/AUTOID.2007.380630, 259 – 264 **(2007)**.

[10] R.A.Rashid, N.H.Mahalin, M.A.Sarijari, A.A.Abdul; "Security system using biometric technology: Design and implementation of Voice Recognition System (VRS)," International Conference on Computer and Communication Engineering, doi: 10.1109/ICCCE.2008.4580735, 898 – 902 **(2008)**.

[11] L.Simon, M.Silverman; "A practical guide to biometric security technology," IT Professional, doi: 10.1109/6294.899930, **3**, 27 – 32 **(2001)**.

[12] D.C.Leonard, A.P.Pons, S.S.Asfour; "Realization of a Universal Patient Identifier for Electronic Medical Records Through Biometric Technology," IEEE Transactions on Information Technology in Biomedicine, doi: 10.1109/TITB.2008.926438, **13**, 494 – 500 **(2009)**.

[13] M.C.Fairhurst; "Signature verification revisited: promoting practical exploitation of biometric technology," Electronics & Communication Engineering Journal, doi: 10.1049/ecej:19970606, **9**, 273 – 280 **(1997)**.

[14] M.Zaied, O.Jemai, C.Ben Amar; "Training of the Beta wavelet networks by the frames theory: Application to face recognition," Workshops on Image Processing Theory, Tools and Applications, doi: 10.1109/IPTA.2008.4743756, 1 – 6 **(2008)**.

[15] B.Knight, A.Johnston; "The role of movement in face recognition," IEEE Transactions on Pattern Analysisand Machine Intelligence, **4**, 265 – 273 **(1994)**.

[16] T.F.Cootes, G.J.Edwards, C.J.Taylor; "Active Appearance Models," Proceedings of the European Conference on Computer Vision, **2**, 484 – 498 **(1998)**.

[17] T.TJolliffe; "Principal Component Analysis (2nd Hd.)"Springer **(2002)**.

[18] T.Jing Yi, Y.H.Tay, L.Phooi Yee Lau; "One-dimensional Grey-level Co-occurrence Matrices for texture classification," International Symposium on Information Technology, doi: 10.1109/ITSIM.2008.4631992, **3**, 1 – 6 **(2008)**.

[19] A.Juraiza Ishak, M.M.Mustafa, A.Hussain; "Gradient Field Distribution and Grey Level Co-occurrence Matrix techniques for automatic weed classification," International Symposium on Mechatronics and Its Applications, doi: 10.1109/ISMA.2008.4648846, 1 – 5 **(2008)**.

[20] S.A.Banday, A.H.Mir, F.Khursheed; "Multi-unit iris biometric fusion using gray level co-occurrence matrix features," International Conference on Advanced Electronic Systems, doi: 10.1109/ICAES.2013.6659397, 225 – 229 **(2013)**.