Visitor Identification - Elaborating Real Time Face Recognition System

Georgy Kukharev, Adam Nowosielski
Technical University of Szczecin
Faculty of Computer Science and Information Technology
ul. Żołnierska 49
70-210 Szczecin, Poland
{gkoukharev, anowosielski}@wi.ps.pl

ABSTRACT

In this paper requirements and conditions for the visitor identification system are outlined and an example system is proposed. Two main subsystems: face detection and face recognition are described. Algorithm for face detection integrates skin-colour, mask analysis, facial features (fast and effective way of eyes localization is presented), reductors, knowledge and template matching. For face recognition a three stage algorithm is proposed. It utilizes well known methods connected in a sequential mode. To improve accuracy and speed some modifications to original methods were proposed and new one presented.

The aim was to build a visitor identification system which would be able to operate in mode with a camera and present results in real-time. The emphasis on speed and accuracy was stressed.

Keywords

face recognition, face detection, eyes localization, DCT, colour segmentation, skin-colour, template matching, real-time

1. INTRODUCTION

Visitor identification system [Kuk03, Sim99, Suk99] belongs to class of face recognition systems. This particular system addresses problems connected with real-life: guests (visitors) identification. The camera (which is a part of the security system) registers the entrance of the building and takes pictures of the scene. For every face detected in the input image, the face recognition algorithm is executed.

Face recognition systems have different requirements and operate in different conditions. Basic systems have simple structure whereas complex contains more sophisticated subsystems. However, the core is invariable and complex systems are formed by connecting additional blocks.

WSCG SHORT Communication papers proceedings
WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.

The basic system for face recognition consists of database and comparator [e.g. Kuk03]. The first contains a set of templates corresponding to known people. Face of a single person can be represented by few templates – models which represents different variants (e.g. different lighting condition, different pose, presence or absence of glasses, etc.). The task of the comparator is to evaluate the similarity between the given face and every face from the database. In such a basic system comparisons between two face images are not efficient. Face images include unnecessary information (e.g. elements of background). It is necessary to extract such features which unambiguously characterize the given face. This task is realized by the features extractor. Face might be represented by a set of features of lower dimensionality. Thus, reduction of dimensionality is obtained. Database is smaller and comparisons are performed faster.

Feature extractor and comparator form a structure which in complex systems is duplicated. In that way, additional accuracy is obtained. There are generally two solutions. In the first, one method performs the task of face recognition and the other method is used for verification (usually in sequential systems). The other is when two (or more) approaches are equivalent and final result is constructed on the base of individual approaches (e.g. voting, weighted summation).

There exists number of methods for face recognition. Generally two main categories are distinguished: analytical (geometrical) and global (holistic, pictorial) approaches [Lam98]. The first group capture the characteristic set of points on a face and relative distances between them. In the second group the aim is to code face image (transform it to face space) as a whole. In both cases dimensionality reduction and better description of face are obtained.

To build fully automatic face recognition system it is necessary to add face detection subsystem. Its aim is to determine if the presented image contains faces or not. If there are any faces, face detector returns their locations and their sizes [Yan02]. The complexity of face detection systems is connected with problems which might belong to one of the class: occlusions, pose, scale, location, facial expressions, disturbing factors (glasses, make-up), scene conditions (lighting, background). To summarize, in the input image number of faces, their sizes and locations in the image space are not known. Because of this, there exists a number of approaches to the task of face detection. Classification of face detection methods can be found in [Yan02].

The basic system for face detector consists of two stages. In the first step most face-like regions are looked for. They are labeled as face candidates and might not contain faces. In the second step, verification process checks which face candidate really contains a face. In practice, face detection is preceded by input image normalization. Its task usually limits to histogram stretching or equalization.

The output of a face detector subsystem are geometrical coordinates corresponding to location of every face in the image or a set of face images.

When there is uniform background and distance from the capture device is fixed, then the task of face detection is simple. However, when there are more faces, their distance is not known and the background is complex, task of face detection is complicated. Moreover, heads have different sizes and normalization process is required. Usually, there is a search for eyes and distance between them is used for size normalization.

When system works in unconstrained environment additional problems might appear. In that case, in normalization block, it is necessary to deal with directional lighting or face rotations.

The remainder of the paper is structured as follows. Section 2 outlines the rules of the visitor identification system, structure of an example system is presented. In section 3 the face detection algorithm is proposed. This section also outlines fast and efficient method of eyes localization and the verification process of face detection. The face recognition algorithm is presented in section 4. The article ends with summary where conclusions are presented.

2. VISITOR IDENTIFICATION SYSTEM

The visitor identification system performs its work in real-world conditions. It is expected to work in a stable and reliable fashion. It should operate fast. Face detection and recognition are time restricted. Slow algorithms are therefore useless in spite of their high accuracy.

The face database cannot be specified once as new faces are frequently added and old removed. The system should be easily adapted. It is clear that with such conditions methods which require training on the full face database are very difficult to implement.

The proper verification of face recognition is required. It should ensure that a person who does not have his or her image in the database is rejected. In a specific installation of the system it is possible that people who appear in the scene are not known and only few of them were registered in the database. But if such a person walks through the door, the system should properly identify him or her.

When elaborating visitor identification system one should take into consideration some characteristic situations. People, when enter a building, usually look straight ahead. It allows to constrain face detection system only to frontal faces with small deflections. Face recognition from such pictures is easier too. The head scale in the image is not known. As person approaches the entrance (with a camera) the face becomes bigger. Face size cannot be therefore restricted. Moreover, the analysed scene might contain more than one walker. People walk in groups frequently. In case the camera is installed in front of an office building, there are times of particular intensity (work beginning at set hours).

For visitor identification system it is characteristic that cost of incorrect identification is not high. Face recognition is performed under conditions when security requirements are not strict. People are identified only to inform others that guests arrived or to collect information how often specified person visit controlled place. But the higher recognition accuracy the better.

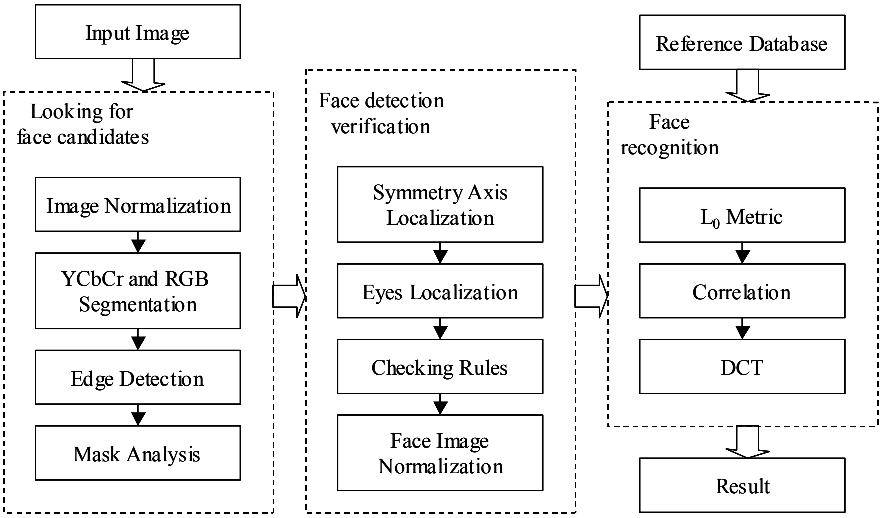

Figure 1. Scheme of the proposed system.

Structure of an example system is depicted in Fig. 1. There are three compound block: face candidate localizator and block for verification (both form face detector) and block responsible for recognition.

The face detection system can generate two types of errors [Yan02]. One is when it does not detect a face and second - when it erroneously classifies as a face a region without a face. By adjusting parameters of one approach one can achieve reduction of the first kind of errors but increasing the second types of errors. Such a solution is favourable in a visitor identification system. The regions which were incorrectly detected can be verified in following stages. Eventually, the image which does not contain a face will not be similar to any image from the database of recognition subsystem. In this way very high detection level is assured.

However, to avoid unnecessary computation, verification process is included. It consists of localization of symmetry axis and position of eyes. If their localization fulfil defined rules, then it is assumed that region contains a face. Additionally, distance between the eyes is used to normalize face size.

A collection of image regions which contain faces constitute output of face detector. These are the input of the second subsystem – face recognition. Face recognition utilizes three stage algorithm. It is based on recognition methods: L_0 metric, correlation and cosine transform. These methods are connected in a sequential order. The idea is to perform fast but inexact approaches to reduce searching space for slower but more reliable methods. Eventually, the list of peoples' names is the output of the system. If a face is not similar to any from the database, it is labelled as "not recognized".

3. FACE DETECTION

Reduction of searching area

First method used in proposed scheme – very popular skin-colour method – belongs to feature invariant approaches [Yan02] and it utilizes colour information. To perform accurate skin-colour detection HSV and YC_bC_r are mainly used. The problem with the HSV colour space is connected with initial non-linear transmission from RGB colour space which is time consuming. It appears, that it is not necessary to perform complicate conversion but it is possible to transfer thresholds directly to RGB colour space. For component *H* of HSV colour space equal to [0; 25] and [335; 360] (skin-colour detector from [Nik00]) equations can be derived:

$$\begin{cases} r > g \\ if\ (g > b)\ then\ (12g - 7b \le 5r). \\ else\ (12b - 7g \le 5r) \end{cases}$$

In addition to skin region detector in RGB, YC_bC_r detector is used. Taking advantages of more than one colour space gives additional accuracy (see Fig. 2). Skin-colour detector in YC_bC_r was built on the base of 25 face examples from different races (people of black skin-colour were not considered). Here are adequate equations:

$$\begin{cases} 135 < Cr < 180 \\ 85 < Cb < 135 \\ Y > 80 \end{cases},$$

where: *Y, Cb, Cr* = [0; 255].

As can be seen from Fig. 2 regions erroneously classified as skin are separate in different colour spaces. Combining the results with AND operator enables the increase in accuracy. Usage of more then one colour space was already used (e.g. [Kuk03, Tan99]).

Figure 2. Skin-colour detection. Input image, final mask (top row). YC_bC_r, RGB mask (bottom row).

Figure 3. Example images (left column), skin-colour masks (right column).

Figure 4. Skin-region mask (left) and final mask after combination with edges (right).

Determination of face candidates

The output of the face segmentation algorithm is a mask which contains sets of connected components. These are the regions which correspond to skin-colour model but they do not necessarily correspond to faces (see Fig. 3). Other parts of the human body (hands) or other objects in the scene might have similar colour distribution as the face model. Moreover, connected components of different objects might unite in one object. It usually happens when background around face is similar to face or person wears clothes which are skin-like. In that case it is necessary to separate face. Generally, verification process is indispensable.

In our algorithm we propose to perform edge detection. To accelerate process the input image is scaled with coefficient equal to 0.5. The influence of small edges that exists on faces is diminished. Experiments showed that changing the image resolution do not have great impact on the results. The final mask is generated by combination of the skin and the edge mask. If skin-colour regions are denoted as 1 and other regions by 0 in the first mask and in the second mask 0 indicate edges, then by using AND operator the final mask is generated. In this mask connected regions of different objects are separated (head is separated from torso, see Fig. 4).

To determine face boundaries scan-line contouring algorithm was adapted [Abl96]. The main feature of the algorithm is that two lines of the input image are processed and conclusions are drawn on the base of existing situations in the current and the former image lines. This algorithm is very fast and efficient. It requires only one pass and allows to detect connected components which are inside other connected components (face might be surrounded by skin-colour background).

There are usually regions in the skin-colour mask that cannot correspond to faces. They are too small or have improper proportions. To eliminate these regions following reductors are proposed: size reductor, vertical edge reductor, horizontal edge reductor, proportion reductor.

Size reductor discards all regions with height or width smaller than 32 pixels. Vertical edge reductor and horizontal edge reductor work on the edge of the regions. They traverse to the centre of the face candidate image (from top, bottom, left and right) until number of skin-colour pixel in line to the edge length is greater than 50%. If the region contains face, connected components have shape similar to ellipse (faces are oval). When reductors are applied, size of bounding rectangle decreases but the face is still distinguishable. When reductors are applied to regions which do not contain faces they can totally eliminate some erroneous regions (see Fig. 5).

The last reductor – proportion reductor – deletes regions which width is greater than length or length is two times greater than width. At last size reductor is executed once again.

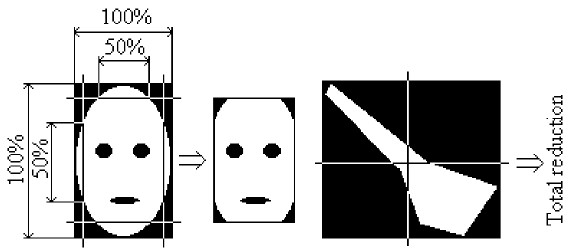

Figure 5. The idea of the reductors on the face and non-face region.

Verification process

After mask analysis is performed it is assumed that every region contains a face. In the next step, the best hypothetical eyes locations are looked for. First, the method of projection is used to find eyes line (the horizontal line which connect two eyes).

Projections are effective way to find facial features [e.g. Kuk03]. The condition which must be fulfilled is to properly select subregion which will be processed. In case of inaccurate adjustment interference from other facial features might appear.

Before eyes line is found the image must be correctly cropped (see Fig. 6). Thanks to horizontal and vertical reductors (see Fig. 5) every candidate contains only central part of hypothetical face. The analysis consists of minima finding in the upper half of the horizontal projection. Some authors uses filtration (e.g. mask of the size of three pixels in [Tse98]).

To make algorithm even more stable the maximum-minimum luminance in every region's row is calculated. Eyes have dark iris and pupil, and white outside. Therefore, the line where the difference between the brightest and the darkest pixels is the largest is the eyes line.

If the individual results of eyes line finding converge, *y* estimation of eyes is found. Otherwise, if the difference is small, the average value is chosen. If the *y* approximations differ substantially, the image strip is taken into consideration for further analysis.

The next step consists of vertical projection calculation - on the very small neighbourhood of eyes line (10% of its length). The maximum value of that profile (only 50 % central part is taken into consideration) indicates the centre of the face. In the method presented in [Kot97] minima from every side of central maximum indicate the *x* coordinates of eyes location. However, when direct light is present, area around the eyes might be shadowed. The same situation occurs when person wears glasses. In both cases *x* coordinates of the eyes might be displaced. Therefore some researchers treat this step as preliminary before proper eyes localization [e.g. Tse98]. To find exact geometrical coordinates different calculations are performed. In [Tse98], for example, repeated multiplications with weighting trigonometric function (sine) are required. In other solutions [Kuk03] template matching is used. These are complex and mainly time consuming operations. Here, we propose to use geometrical moments but first we introduce advantages of using *R-B* colour space.

Figure 6. Eyes localization on the base of profiles in R-B colour space.

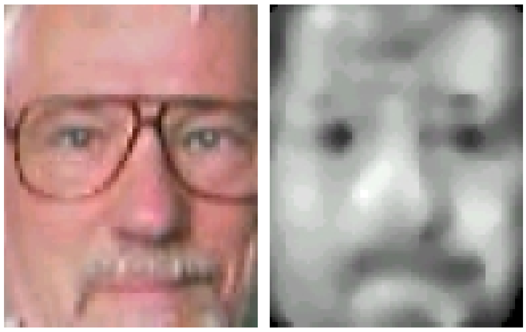

Figure 7. Reduction of the influence from the glasses in R-B colour space.

By subtracting channels *R* and *B* (of the original image) undesirable features are eliminated. Output image has dark regions only around the eyes (see Fig. 6). Influence of the glasses is also reduced (see Fig. 7).

To exactly determine eyes locations we proceed to moments calculations. For the image region *I(y, x)=R(y, x)-B(y, x)* constrained by a rectangle: $x_1 \le x \le x_2$ and $y_1 \le y \le y_2$ (see Fig. 8) zero and first order moments can be calculated from equations [Bra98]:

$$M_{00} = \sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} I(y,x),\ M_{10} = \sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} xI(y,x)\,.$$

The M_{10}/M_{00} quotient might also be used to describe face symmetry ([Bra98]). In proposed scheme it is used for verification of central axis of the face (found by method of profiles)

Fig. 8. Eyes localization searching area.

Eyes coordinates are searched in regions on the left and right side of the symmetry axis. But to find the location successfully, pixels values must be inversed according to equation: max(*I*)-*I(y,x)*. The proposed eyes locator for image *I(y,x)=R(y,x)-B(y,x)* constrained by the rectangle: $x_1 \le x \le x_2$ and $y_1 \le y \le y_2$ can be written as:

$$\begin{cases} X_C = M_{10} / M_{00} \\ m = MAX(I(y,x)) \\ X_1 = \sum_{x=x_1}^{X_C} \sum_{y=y_1}^{y_2} x(m - I(y,x)) / \sum_{x=x_1}^{X_C} \sum_{y=y_1}^{y_2} (m - I(y,x)) \\ X_2 = \sum_{x=X_C}^{x_2} \sum_{y=y_1}^{y_2} x(m - I(y,x)) / \sum_{x=X_C}^{x_2} \sum_{y=y_1}^{y_2} (m - I(y,x)) \end{cases}$$

where: M_{00}, M_{10} - zero and first order moments.

Fig. 9 presents example results of finding eyes locations by a given scheme. Proposed algorithm allows correct determination of eyes geometrical coordinates even in difficult conditions. Drawbacks of original methods are decreased. Profile analysis of the *R-B* image allows to eliminate negative influence from glasses. Use of geometrical moments guarantee proper eyes finding even when local minimum is shifted from the centre of the eye.

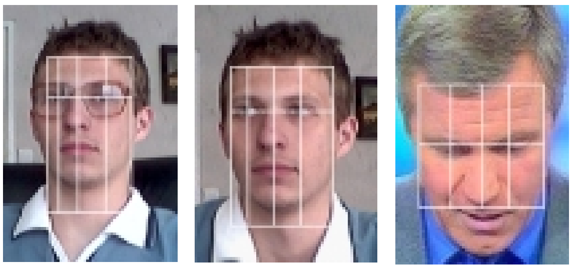

Figure 9. Eyes detection. Face with glasses, rotated face, bowed head.

Here, it was assumed that head is not rotated round *Z* axis, and therefore there is one *y* coordinate for both eyes. However, by calculating moments for *y* value (separately for left and right area of symmetry axis) our algorithm might be extended to that case.

That way for every region which passed process of mask analysis and passed through reductors eyes are located. They are used in face normalization process. Size of the face depends on distance between the eyes, and area which embrace face depends on eyes locations.

In the first step of recognition process, face image is compared with database templates. This additionally perform the task of verification.

4. FACE RECOGNITION

Face recognition in the proposed scheme consists of three stages. In the first step the recognition process is based on the L_0 metric and face images of reduced resolution (16x16). This solution is similar to systems described in [Kuk03, Sim99]. The L_0 metric between two images (described as vectors $\vec{x}$ and $\vec{y}$) is defined as [Sim99]: $L_0(\vec{x} - \vec{y}) \equiv \sum_{|x_i - y_i| > \delta} 1$, where δ is the threshold. If the difference between the intensity value of the two corresponding pixels is smaller then the threshold, it is assumed that these pixels are similar. In practice, when the threshold is too high, errors occurs. On the other hand, if the threshold is too low, similar faces might not be matched. Such situations may appear when considerable differences between images exist (lighting conditions, viewing angle). The threshold in presented system was set to 25. The lower is the L_0 distance between two examined faces the more they are similar. In practice, it is more convenient to use fraction value, obtained by division L_0 by the number of pixels in the image).

In case of erroneous face detection some of the regions might not contain faces. They will be characterized by high distance with any face from the database. When the minimum distance is greater than 0.4, then region which is being tested is discarded as non face-like or as the depicted person is not similar to any from the database. If this distance is smaller then 0.4, then analysis of a given face is continued but only with those faces from the database which have similarity measure smaller then 0.4 too. Proposed solution thus realizes two task. First, it performs face verification. Secondly, it prepares face database for further (more precise) examination.

In the second step of face recognition the method of face correlation is employed [Bru93]:

$$C(X,Y) = \frac{\overline{X \circ Y} - \overline{X}\,\overline{Y}}{\delta(X)\,\delta(Y)},$$

where $\circ$ is the pixel by-pixel product, and δ is the standard deviation.

For the correlation coefficient the threshold was set to 0.7. Experiments suggest that this value cannot be lower. For a safety reasons it should be set at 0.8. If the maximum correlation coefficient between the given face and any face from the database is lower than 0.7, then it is assumed that face which is being examined does not belong to known person.

Otherwise, from classes with the highest similarity the winner is chosen. The winner is a class with the highest similarity between vectors resulting from the cosine transform of examined images. In the third step the time-consuming operations are performed only if a face which is being examined is similar to any face from the database (determined in two preceding stages) and comparisons are carried out on a small subset of record from original database.

The discrete cosine transform (DCT) is used as the extraction method of the most significant facial features. The main advantage of the transform is its high compression rate (exploited for example in the JPEG standard). In presented system the DCT is performed on the full face image. It is similar to solution described in [Haf01] but owing to proposed scheme last stage's comparisons do not include the full database.

A face image of dimensions *MxN* is scaled to *NxN* under assumption that $M > N$ and N is the width of the image. Then the coefficients of cosine transform are calculated and only small subset of them (low and middle frequencies) is saved as a feature vector. They have the greatest variance [Rus99]. There again, to accelerate the process it is sufficient to calculate only the coefficients which are needed. That is for u, v = 0..7, according to equation [Rus99]:

$$\begin{cases} F(u,v) = \dfrac{4c(u,v)}{N^2} \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} I(y,x) \cos\alpha \cos\beta \\ \alpha = \dfrac{(2x+1)\pi u}{2N}, \ \beta = \dfrac{(2x+1)\pi v}{2N} \\ if \ u \ or \ v = 0 \ then \ c(u,v) = 0.5 \ else \ c(u,v) = 0 \end{cases}$$

Obtained 8x8 matrix is transformed into a vector and the Euclidean distance classifier is used to compare given face ($\vec{a}$) with a database face ($\vec{b}$):

$$E = \sqrt{\sum_{i=0}^{63} (a_i - b_i)^2} \ .$$

The winner is the class whose member has the lowest distance with the examined face. Moreover, the threshold value equal to 2025 was set here, to ensure correct classifications and rejections.

Proposed scheme has two main advantages. It reduces total amount of calculations. It utilizes discrete cosine transform which is characterised by high accuracy. Presented visitor identification system works therefore nearly in real-time and performs its job reliable.

To test the proposed solution The ORL Database of Faces (10 pictures of 40 person each) was used. First, the database was sorted (like in [Kuk03]) in a way that faces with the highest variance were assigned the smallest indexes. This way the database is representative. Database of known individuals was built from the first 30 classes (8 images were used). The remaining 60 images (2 faces, 30 classes) were used to test recognition accuracy. The remaining 10 classes (100 images) were used to test the false acceptance rate. The results are 96.67% for recognition accuracy (RA) and 5% for FAR.

5. SUMMARY

In this article elements of face recognition systems have been presented. From basic blocks it is possible to build any kind of face recognition system. Requirements for a visitor identification system were specified and an example system has been presented.

The visitor identification system built upon proposed scheme might be used to identify people entering buildings, in stores or to recognize frequent visitors in different kind of institutions. To make system work efficiently and reliably some conditions must be fulfilled.

In proposed solution there are no methods that eliminate any kind of directional light. When such a illumination exists in the scene, shadows might occur on the face. It significantly complicate the process of face detection and recognition. Algorithms that solve above problem are slow and usually do not cope with it satisfactorily. To add such an algorithm would render the proposed scheme slow. It is much cheaper to install suitable illumination and a cover in front of the door which would limit the influence of the sun.

However, the best solution would be to implement an algorithm that analyses the scene whether or not directional light exist. If the answer to above question is positive then additional light would be turned on or system would try to reduce its negative influence.

The visitor identification system, built on the base of proposed algorithm, operates approximately in real-time. The processes of detection and recognition takes 0.11 second (it is average time for images of dimensions 320x240 and 99 faces database) achieving 9 images per second. It is a good result especially when taking into consideration computer speed of 900 MHz on which the tests were performed. Figure 10 shows the main window of the visitor identification system. On the right hand side of the window there is an input from the camera. In the centre there is the database image of the best recognized person from the scene. On the left hand side there is a list of all recognized persons.

Contemporary recognition systems reached very high recognition rate (nearly 100% [e.g. Kuk03]). In comparison with the human recognition ability they still have to be improved. A human takes into consideration a lot of visual factors like way of walking, clothes, etc. In a computer security system, however, face might be utilized as an additional factor which might increase the credibility. On the other hand, when the safety requirements are not high computer systems might fully substitute a human.

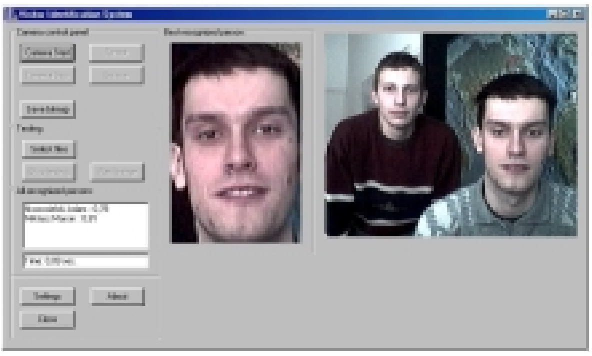

Figure 10. Visitor identification system.

The visitor identification system solves the task of guests identification. The combination of methods proposed in the article might be used in another face recognition system. Presented scheme might be used (after adaptation and threshold adjustment) to solve the task of authentication or be used as an independent component in the access control system. Some of presented tasks can be utilized in current version of realized system but only in the manual mode.

6. REFERENCES

[Abl96] S. Ablameyko. Recogniton of graphic images. Belarussian Academy of Sciences Institute of Engineering Cybernetics. Minsk. pp. 83-85. 1996.

[Bra98] G. Bradski. Computer Vision Face Tracking For Use in a Perceptual User Interface. Intel Tchnology , 2nd Quarter, 1998.

[Bru93] R. Brunelli, T. Poggio. Face Recognition: Features versus Templates. IEEE Transactions on Pattern analysis and machine intelligence. pp. 1042-1052. October 1993, 15(10).

[Haf01] Z. M. Hafed, M. D. Levine. Face Recognition Using the Discrete Transform. International Journal of Computer Vision, 43(3). pp. 167-188. 2001.

[Kot97] C. Kotropoulos, I. Pitas. Rule – Based Face Detection in Frontal Views. In Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP'99). pp. 21-24. Phoenix, Arizona, 1997.

[Kuk03] G. Kukharev, A. Kuźmiński. Techniki Biometryczne Część 1 Metody Rozpoznawania Twarzy. Wydział Informatyki Politechnika Szczecińska, 2003.

[Lam98] K. Lam, H. Yan. An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View. IEEE Transactions on pattern analysis and machine intelligence, vol. 20, No. 7. pp.673-686. July 1998.

[Nik00] A. Nikolaidis, I. Pitas. Robust Watermarking of Facial Images Based on Salient Geometric Pattern Matching. IEEE Transactions on Multimedia, vol. 2, no. 3. pp 172-184. September 2000.

[Rus99] J. C. Russ. The Image Processing Handbook, Third Edition. pp. 136-155. 1999.

[Sim00] T. Sim, R. Sukthankar, M. Mullin, S. Baluja. High-Performance Memory-based Face Recognition for Visitor Identification. In Proceedings of the 4tn International Conference on Face and Gesture Recognition, Grenoble, France, March 2000.

[Suk99] R. Sukthankar, R. Stockton. ARGUS: An Automated Multi-Agent Visitor Identification System. Proceedings of AAAI-99, July 1999.

[Tan99] J. Tang, S. Kawato, J. Ohya. Face Detection Form a Complex Background. International Workshop on Very Low Bitrate Video Coding. Kyoto Research Park on October 29-30, 1999, Japan.

[Tse98] S. Tsekeridou, I. Pitas. Facial Features Extraction in Frontal Views Using Biometric Analogies. Proc. IX European Signal Processing Conference, 1998, vol. I.

[Yan02] M. Yang, D. Kriegman, N. Ahuja. Detecting Faces in Images: A Survey. IEEE Transactions on pattern analysis and machine intelligence, vol. 24, No. 1. pp. 34-58. January 2002.