MATCHING OF FACES IN CAMERA IMAGES AND DOCUMENT PHOTOGRAPHS

Valery Starovoitov[1] , Dmitry Samal[1], and Bulent Sankur[2]

1) Institute of Engineering Cybernetics
Surganova Str., 6, 220012 Minsk, Belarus

{ValeryS,samal}@newman.bas-net.by

2) Bogazici University, Dept. of Electrical Engineering, Istanbul, Turkey

sankur@boun.edu.tr

ABSTRACT

An automatic system for the person identification based on comparison of a camera image and a document photo presents several design difficulties. The main complication arises from the fact that the document photo may represent quite a different appearance of the test person, i.e., an older version, sometimes by years, or with a different hairstyle, facial expression and lighting conditions. In addition the scanned photo is usually of a much lower quality. In this paper elements of a system for a face comparison between camera images and archival photo albums are discussed and experimental results are given. Such a system may be instrumental for document control or for police investigations.

1. INTRODUCTION

There are various scenarios where the actual camera image of a person needs to be compared with his/her photo, for example, glued on a document. Routine security checks, access control, passport or driving license validation necessitate this type of check and verification. There are some advantages in the automation of such an "facial image matching" process for service efficiency and to preclude any subjectivity.

Such a system can address two problems: 1) Determination of whether the document, e.g., the passport, is forged or not, 2) To establish whether the passport bearer is actually presented in the photo or not. In the paper we concentrate on the second task.

At first glance the problem looks like a classical face identification problem. However the peculiarity of the problem is that the archival / document photograph is typically made at quite a different time, sometimes by several years. Furthermore hairstyle, expression and pose of the individual may all differ significantly. Finally one can expect the image qualities to be different due to, camera parameters, lighting conditions, smaller size of the document photos, and deterioration due to scanning conversion.

There are already systems developed solely for access control, or systems for browsing in a digitized face database. One can mention, for example, such experimental systems as "Photobook" [1], and others [2,3,4,5]. There are several commercial systems, like "FaceIt®" [6], "FLAVOR" [7], “ZN-Face®“ [8], “Face ID” [9], for which details are not available in the open literature.

The goal of our paper is to build a system that effects a robust and automatic comparison of an arbitrary document photo and an actual camera image of a person.

2. AN ALGORITHM FOR VIDEO-TO-PHOTO MATCHING

A fully automatic system for matching face images in document photographs to the camera images is described in the flow chart in Fig.1. Every face verification or recognition system performs more or less the same three essential functions:

- Face detection and eye localization.
- Feature extraction
- Evaluation of some similarity function.

Figure 2 illustrates several examples of actual passport and video camera images to be treated by this system.. Each one pair belongs to the same person.

Below we detail the algorithmic steps.

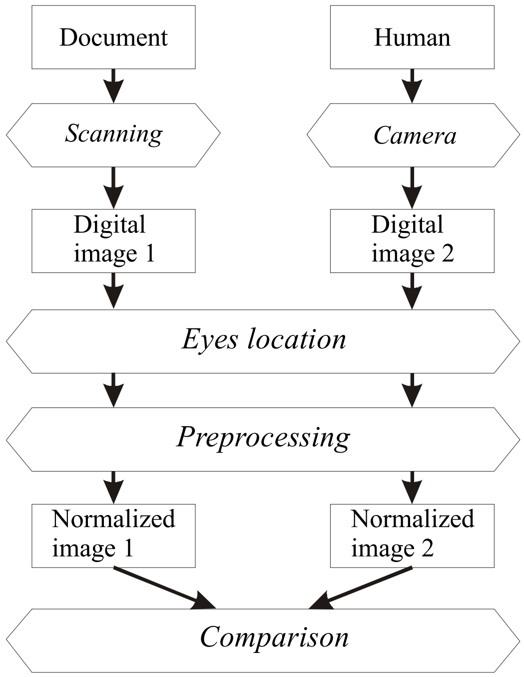

Figure 1. The flow chart of our system.

2.1. Eye localization

We first locate the eyes, which are assumed to be "open" in facial images. The gray-scale image is processed with the Sobel edge detector whose output is thresholded with Otsu's method to obtain a binary image. We calculate the vertical projection of this binary image and smooth the projection profile by averaging. The projection profile is searched for two large plateaus with a central high peak. The peak indicates the eye region in the image (the right part of Fig.3). To localize the eyes more precisely we apply the Hough transform to a small strip of the binary image (the shaded area in Fig.3) using a lower half circle mask as often the upper part of the iris is covered by the eyelids. To speed up the processing, we store several masks corresponding to different radii for the Hough transform. Among several peaks in the Hough space we find the two highest scoring candidates. The crosses in Fig.3 show the centers of eyes and some details of their location. The output of this step is the coordinates of the eyes.

Figure 2. Examples of face patterns obtained by passport scanning (right column) and by a camera (left column) .

2.2. Image conditioning

In the second step we condition the initial images, which involves rotation, scaling and cropping of the central part. We try to remove background and hair, and keep the most invariant in time part of face. In increasing order of fiducial patterns we have eyes and eyebrows, followed by nose. We have observed that the nose appearance varies a lot as a consequence of head rotation and lighting conditions. Finally mouth is the most variant part of the face center.

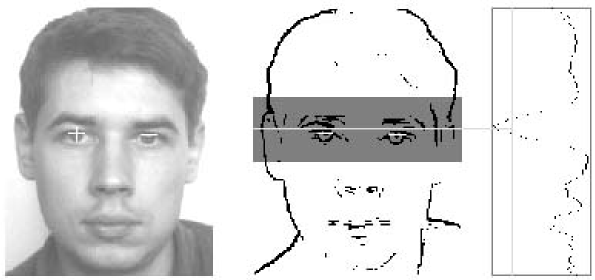

Figure 3. Original face image and located eye centers.

We can rely on the fact that in document photos face expression is usually minimal, at least discreet, while for the camera control image the test person can be requested to restrain from excessive expressions. Additional features from the original image can be gleaned in the form of a gray-scale edge map using the Deriche algorithm.

Examples of the conditioned images are given in in Fig. 4.

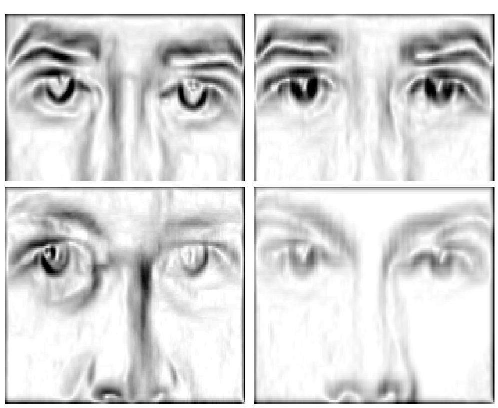

Figure 4. Cropped images of the faces presented in Fig.1.

Then we apply the ranking transform to the edge map and change the value of every map element by the rank of the edge value corresponding to the element. Histogram of the edge map gives ranking of all edge values. If there are, for example, 51 elements with value 0 in an edge map, their rank equals (51+1)/2= 26; let 3 map elements have value 1, their rank is 26+(3+1)/2=28, and so on...

The outcome of this step is the rank map of the cropped image.

2.3. Face comparison

Given the variability of a face, even under controlled conditions it is futile to try to compare directly original images or even their feature maps. In fact, we have got very low similarity scores with the popular method of mosaic images, and using linear correlation.

With the rank transformation described in Section 2.2, differences of luminance and contrast become compensated for. Furthermore to compensate for the local changes due facial expression and pose, we apply a nonlinear warping of one rank map vis-a-vis another map. Among the many algorithms to warp images we used a technique similar to that described in [10]. It turns out that comparison of one warped map with another non-warped one yields much better results than a direct comparison of the maps. However, the computational cost of warping is high, as we must test a large set of randomly generated local deformations. A practical way to maneuver in the warping parameter space consists in checking after every warping the score against the previously achieved level of similarity. We continue to warp the first map if the present score is higher than the previous one; otherwise we start warping with other parameters.

Note that the dissimilarity measure we used to compare two maps was the rank correlation coefficient. The smaller this score, the more similar are the two rank maps and, hence, the persons presented in the images. In fact, it equals 0 if the compared maps are identical, equals to 1 if they are very different, and assumes the value 2 if they are reverse images. When the score falls below a predefined threshold T, then one can conclude that the two images describe the same person.

3. RESULTS

Our database contains about 100 passport and camera images obtained under various illumination conditions. In the various experiments that we ran good results were obtained when the interocular distance was at least 40 to 50 pixels. This also sets a lower bound on the resolution level. The experimentally determined threshold on the rank correlation coefficient was T=0.19, at which value we obtained the lowest misclassification probability of about 9%.

The most difficult pairs of face images in our database are the ones presented in Fig 2: a and b, c and d, g and h. Notice the very different quality and contrast of the camera images and the passport photos (see also Fig.4). Run time of our non-optimized experimental procedure was 10-12 seconds on a Pentium PC (doing 800-1000 warpings). The system remains to be compared with its competitors, but we believe that due to the inherent nonlinearity in the rank correlation it should be more robust vis-a-vis scheme relying on linear correlation [4].

4. CONCLUSION

The proposed face similarity meter was found to perform satisfactorily in adverse conditions of different epochs of exposure, illumination and contrast levels, and face pose.

The preliminary recognition accuracy of above 91% can be improved if additional features are considered and the running time speeded up by cruising the warping space more efficiently, e.g., by allowing small face rotations (up to ±15 degrees).

REFERENCES

1. B. Moghaddam and A. Pentland, Face recognition using view-based and modular eigenspaces, *Automatic Systems for the Identification and Inspection of Humans*, SPIE Conf., Vol. 2277, 1994, pp. 1868-1876.
2. S. Gutta and H. Wechsler, Face recognition using hybrid classifiers, *Pattern Recognition*, Vol. 30, No.4, pp. 539 –553, 1997.
3. A. Samal and P.A. Iyengar, Automatic recognition and analysys of human faces and facial expressions: a survey, *Pattern Recognition*, 1992, Vol. 25, No 1, pp. 65-77.
4. R. Brunelli and T. Poggio, Face recognition: features versus templates, *IEEE Trans. on PAMI*, Vol.15, No.10, pp.1042-1052, 1993.
5. R. Chellapa, C.L. Wilson, S. Sirohey and C.S. Barnes Human and machine recognition of faces: a syrvey, Proc. of IEEE, Vol. 83, No. 5, pp. 705 - 739, 1995.
6. Visionics Corporation, http://www.faceit.com.
7. http://selforg.usc.edu:8376/frp.html.
8. ZN Gmbh., http://www.zn-gmbh.com.
9. Image Ware Software, http://www.iwsinc.com/crimes/faceid.htm.
10. J.C. Gee and D.R. Haynor Rapid coarse-to-fine matching using scale-specific priors. *Medical Imaging*, SPIE Conf., Vol.2710, 1996, pp. 416-427.