

Specific Sensors for Face Recognition

Walid Hizem, Emine Krichen, Yang Ni, Bernadette Dorizzi,
and Sonia Garcia-Salicetti

Département Electronique et Physique, Institut National des Télécommunications,
9 Rue Charles Fourier, 91011 Evry France
Tel: (33-1) 60.76.44.30 , (33-1) 60.76.46.73
Fax: (33-1) 60.76.42.84
{Walid.Hizem, Emine.Krichen, Yang.Ni, Sonia.Salicetti,
Bernadette.Dorizzi}@int-evry.fr

Abstract. This paper describes an association of original hardware solutions associated to adequate software software for human face recognition. A differential CMOS imaging system [1] and a Synchronized flash camera [2] have been developed to provide ambient light invariant images and facilitate segmentation of the face from the background. This invariance of face image demonstrated by our prototype camera systems can result in a significant software/hardware simplification in such biometrics applications especially on a mobile platform where the computation power and memory capacity are both limited. In order to evaluate our prototypes we have build a face database of 25 persons with 4 different illumination conditions. These solutions with appropriate cameras give a significant improvement in performance (on the normal CCD cameras) using a simple correlation based algorithm associated with an adequate preprocessing. Finally, we have obtained a promising results using fusion between different sensors.

1 Introduction

The face recognition systems are composed of a normal video camera for image capturing and a high speed computer for the associated image data processing. But this structure is not well suited for mobile device such as PDA or mobile phone configuration where both computation power and memory capacity are limited. The use of biometrics in mobile devices is becoming an interesting choice to replace the traditional PIN code and password due to its commodity and higher security.

The high complexity of face recognition in a cooperative context comes largely from the face image variability due to illumination changes. Indeed, a same human face can have very different visual aspects under different illumination source configurations. Research on face recognition offers numerous possible solutions. First, geometric feature-based methods [3] are insensitive to a certain extent to variations in illumination since they are based on relations between facial features (eyes, nose, mouth); the problem of these methods is the quality of the detection of such features, which is far from being straightforward, particularly in bad illumination conditions. Also, statistical methods like Principal Components Analysis [4], Fisherfaces [5], and Independent Components Analysis [6] emerged as an alternative

to a certain variability of facial appearance. Such methods, despite success in certain conditions, have the drawback of being reliable only when the face references used by the system and the face test images present similar illumination conditions, which is why some studies have proposed to model illumination effects [7]. So large computation power and memory capacities have to be dedicated to compensate this variability. Consequently reducing this image variability at the face image capturing stage can result in a significant both hardware and software simplification.

In this paper, we present an association of hardware and software solutions to minimize the effect of ambient illumination on face recognition. We have used two dedicated cameras and an appropriate pre-processing to suppress the ambient light. We have also built a database under different illumination conditions and with different cameras. Then a pixel correlation algorithm has been used for testing purpose.

In the following sections, we will present the two cameras. Then, we show the influence of illumination on face recognition. And finally, we describe our protocols and the results of our method.

2 Camera Presentation

2.1 Active Differential Imaging Camera - DiffCam

In a normal scene there is not a big illumination variation between two successive frames: the illumination remains static. So to eliminate it, a differentiation operation can be used in this case. We have applied inside a specially design CMOS image sensor with an analog memory in-situ in each pixel (Fig. 1). The integration of this in-situ analog memory permits a parallel image capture and further an on-chip differentiation computation.

The working sequence is the following: 1) the first image is captured by illuminating the subject's face with an infrared light source and 2) the second is captured by turning this light source off. The two captured images will be subtracted from each other during the image readout phase by using on-chip analog computation circuits on the sensor chip as shown Fig. 2. We have designed and fabricated a prototype CMOS sensor with 160*120 pixels by using a standard 0.5 μ m single poly CMOS technology. The pixel size is 12 μ m. A 8-bit ADC converter has been integrated equally on the sensor chip, which reduces considerably the system design complexity[1].

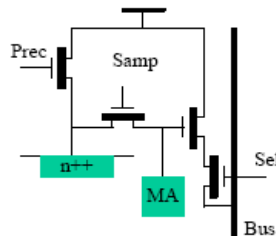


Fig. 1. Structure of the pixel

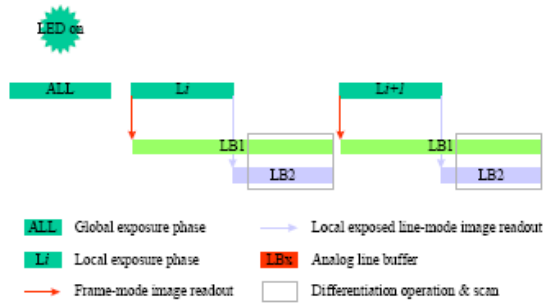


Fig. 2. The function principle and sequence of the active differential imaging system[1]

Compared to other analog/digital implementations such [8] [9], our solution requires not only single analog memory in each pixel, which gives an important pixel size reduction, but also neither off-chip computation nor image frame buffer memory.

A prototype camera with parallel port interface has been built by using two micro controllers. The infrared flash has been built with 48 IR LEDs switched by a MOSFET. A synchronization signal is generated from the microcontroller controlling the sensor. The pulse length is equal to the exposure time (50 μ s, the frame time is 10ms). The peak current in the LEDs is about 1A but due to the small duty cycle (1/200), the average current is low.

2.2 Synchronized Flash Infrared Camera – FlashCam

Another possible way to attenuate the ambient light contribution in an image is to use a synchronized flash infrared illumination. As shown in (Fig. 3), in classic integration-mode image sensor, the output image results from a photoelectric charge accumulation in pixels.

As has been indicated, the stationarity of ambient light makes its contribution proportional to its exposure time. So the idea here is to diminish the ambient light contribution by reducing the exposure time and at the same time using a powerful infrared flash, synchronized with this short exposure time. The images obtained by this imaging mode result mostly from the synchronized flash infrared light. This imaging mode has the advantage to work with a standard CCD sensor.

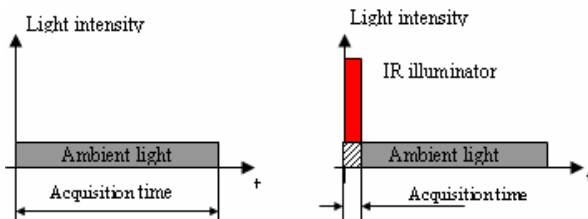


Fig. 3. Principle of the synchronized pulsed illumination camera[2]

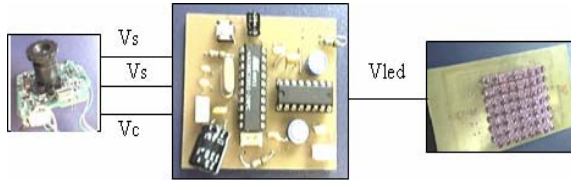


Fig. 4. The functional architecture of the prototype

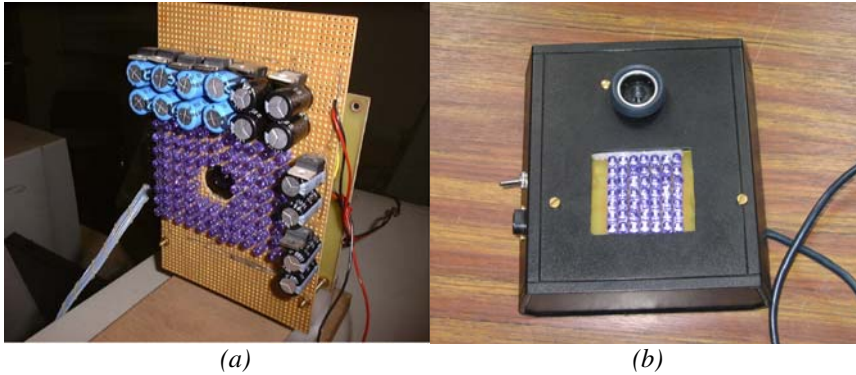


Fig. 5. (a) The active Differential Imaging System (b) The Synchronized pulsed flash camera

An experimental camera has been built by modifying a PC camera with a CCD sensor. CMOS sensor based PC cameras cannot be used here, because the line sequential imaging mode used in APS CMOS image sensors is not compatible with a short flash-like illumination. The electronic shutter and synchronization information has been extracted from the CCD vertical driver. This information is fed into a micro controller which generates a set of control signals for infrared illuminator switching operation as shown in Fig. 4. The same LED based infrared illuminator has been used for this prototype camera. Fig. 5 shows the two prototype cameras.

3 Database

3.1 Description

To compare the influence of the illumination on faces, a database with 25 persons has been constructed by using three cameras: DiffCam, FlashCam and also a normal CCD camera. There are 4 sessions in this database with different illumination conditions: Normal light (base1), no light (base2), facial (base3) and right side illumination (base4). In the last two sessions, we have used a desk lamp to illuminate the face. In each session we have taken 10 images per person per camera. So we have 40 images per person per camera. The resolution of the images from the DiffCam is 160×120 , the resolution obtained from the FlashCam and the normal CCD Camera images are



Fig. 6. Samples of the face database



Fig. 7. Samples of the face expression

320×280. The captured images are frontal faces; the subject was about 50cm from the device. There are small rotations of the faces on the three axes and also expression on faces. Indeed, anyone could wear glasses, regardless of whether spot reflections obscured the eyes. Face detection is done manually using the eyes location. Samples of this database are shown in Fig.6. (for the same person and different illumination conditions). Samples of different face expressions are shown in the Fig.7.

3.2 Protocol

For the experimentation, we have chosen 5 images for each person as test images and 5 as reference ones. We have two scenarios: The first consists in comparing images from the same camera and the same illumination condition. The second compares images from the same camera but from different session (illumination conditions change): there are six comparisons in this scenario: Normal light versus no light (base 1 vs base 2), Normal light versus facial illumination (base 1 vs base 3), Normal light versus right side illumination (base 1 vs base 4), No light session versus facial light (base 2 vs base 3), no light versus right side illumination (base 2 vs base 4) and facial illumination versus right side illumination (base 3 vs base 4).

4 Preprocessing and Recognition Algorithm

First the faces are detected and normalized. We have performed a series of preprocessing to attenuate the effect of illumination on face images. The best result has been found with a local histogram equalization associated with a Gaussian filter. In order to take benefits from the face symmetry and to reduce the effect of lateral illumination, we have added a second preprocessing calculating a new image.

$$I_s = \frac{I(x, y) + I(-x, y)}{2} .$$

We have applied this preprocessing to the images acquired with the normal CCD cam as they will be more perturbed by illumination effects. For the other images we've applied only an histogram equalization.

The verification process is done by computing the Euclidian distance between a reference image (template) and a test image.

4.1 Experimental Result

We have splitted our database into 2 sets, the templates set and the test set. As 10 images are available for each client and each session, we consider 5 images as client's templates, and the remaining 5 images as test images. Each test image of a client is compared to a template of the same client using the preprocessing and recognition algorithms above described, and the minimum distance between each test image and the 5 templates is kept. We obtain this way 125 intra class distances. Each test image is also compared to the other sets of 5 templates of the other persons of the database in order to compute the impostor distances. So we have 3000 inter class distances. The following tables (Tab.1 and Tab.2) compare the performance (in terms of EER) in function of the type of the camera. For the first Camera, we have two results: the first corresponds to preprocessed images. The second one uses images without preprocessing.

The first scenario (images from the same session) shows a general good and equivalent performance for each camera for the different illumination conditions.

In the second scenario, the reference images are taken from one session and the test images are taken from another session (different illumination conditions). Using the images, taken from the first camera without preprocessing, gives 50% of EER in nearly all the tests. Using a preprocessing improves significantly the results, which proves its usefulness for the Normal CCD Camera in order to attenuate the illumination effects. Comparing the normal camera and the FlashCam, we notice that

Table 1. Scenario 1

| EER | Base 1 | | Base 2 | | Base 3 | | Base 4 | |
|------------|--------|----|--------|----|--------|------|--------|------|
| Normal CCD | 3.4% | 6% | 6% | 6% | 3.2% | 5.5% | 4.5% | 4.7% |
| FlashCam | 5% | | 4.2% | | 3.2% | | 2% | |
| DiffCam | 5,6% | | 2% | | 3.5% | | 4.1% | |

Table 2. Scenario 2

| EER | Base 1vs2 | | Base 1vs3 | | Base 1vs4 | | Base 2vs3 | | Base 2vs4 | | Base 3vs4 | |
|----------|-----------|-----|-----------|-----|-----------|-----|-----------|-----|-----------|-------|-----------|-------|
| | 20% | 39% | 38% | 53% | 24.5% | 54% | 40% | 56% | 30% | 50,7% | 25% | 37,6% |
| FlashCam | 26% | | 27% | | 22% | | 28% | | 22% | | 23% | |
| DiffCam | 15,7% | | 14% | | 21% | | 9.5% | | 13% | | 15% | |

flashCam gives an improvement of the EER especially in the tests: Base1vs3, Base2vs3 and Base2vs4. In all these tests we observe a stable EER for the flashCam: this suggests a stronger similarity between the images acquired under different illumination conditions than the ones from the normal CCD. The relative high EER of the FlashCam is due to the quality of some images for which the flash did not give a sufficient light due to battery weakness. The correlation algorithm might be not suitable for the flashCam.

We have tried the eigenfaces algorithms but it gives worse results. We have to investigate other methods.

Comparing the FlashCam and the DiffCam, we observe that the second camera gives better results in all tests. The noticeable improvement is on tests: Base 2vs3, Base 1vs2 and Base 3vs4. This indicates the existence of residual influence of ambient light on the output images from FlashCam. On the contrary, we confirm real suppression of the ambient light by the differentiation operation.

4.2 Fusion Results

We have done other tests to know if the three cameras can be associated to give better results. For this purpose, we have done a simple mean between the scores given by the three cameras, (after some normalization) Table 3 shows the results of this fusion scheme and compares them to the best single camera performance.

We notice that in most cases, the fusion improves the best single camera results: this is due to the complementarities between the infrared images that eliminate the ambient light and the details in faces and the normal camera that compensates this lack of details.

Table 3. Fusion result of the three cameras

| EER | Base 1vs2 | Base 1vs3 | Base 1vs4 | Base 2vs3 | Base 2vs4 | Base 3vs4 |
|--------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 3 cameras fusion | 11.2% | 10.7% | 18% | 13.9% | 14.3% | 9.6% |
| Best single camera | 15,7% | 14% | 21% | 9.5% | 13% | 15% |

5 Conclusion

In this paper, we have presented two specialized hardware developed in our laboratory dedicated to face recognition biometric applications. The first one is based

on temporal differential imaging and the second is based on synchronized flash light. Both cameras have demonstrated a desired ambient light suppression effect. After a specific preprocessing, we have used a simple pixel-level correlation based recognition method on a database constructed with varying illumination effects. The obtained performance is very encouraging and our research direction in the future is focused on a SoC integration of both sensing and recognition functions on a same smart CMOS sensor targeted for mobile applications.

References

1. Y. Ni, X.L. Yan, "CMOS Active Differential Imaging Device with Single in-pixel Analog Memory", Proceedings of IEEE European Solid-State Circuits Conference (ESSCIRC'02), pp. 359-362, Florence, Italy, Sept. 2002.
2. W. Hizem, Y. NI and E. Krichen, "Ambient light suppression camera for human face recognition" CSIST Pekin 2005
3. R. Brunelli, T. Poggio, "Face Recognition: Features vs. Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, N° 10, pp. 1042-1053, October 1993.
4. M. A. Turk and A. P. Pentland. Face Recognition Using Eigenfaces. In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pages 586 – 591, June 1991.
5. Jian Li, Shaohua Zhou, Shekhar, C., "A comparison of subspace analysis for face recognition", Proceedings of ICASSP'2003 (IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing), 2003.
6. M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, "Face recognition by Independent Component Analysis", IEEE Transactions on Neural Networks, Vol. 13, N°6, pp. 1450-1464, Nov. 2002.
7. Athinodors S Georghiades, Peter N.Belhumeur, David J.Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, pp. 643 – 660.
8. Hiroki Miura & al. "A 100Frame/s CMOS Active Pixel Sensor for 3D-Gesture Recognition System", Proceeding of ISSCC98, pp. 142-143
9. A. Teuner & al. "A survey of surveillance sensor systems using CMOS imagers", in 10th International Conference on Image Analysis and Processing, Venice, Spet. 1999.