

# Ambient Illumination Variation Removal by Active Near-IR Imaging

Xuan Zou, Josef Kittler, and Kieron Messer

Centre for Vision, Speech and Signal Processing,  
University of Surrey, United Kingdom  
{x.zou, j.kittler, k.messer}@surrey.ac.uk

**Abstract.** We investigate an active illumination method to overcome the effect of illumination variation in face recognition. Active Near-Infrared (Near-IR) illumination projected by a Light Emitting Diode (LED) light source is used to provide a constant illumination. The difference between two face images captured when the LED light is on and off respectively, is the image of a face under just the LED illumination, and is independent of ambient illumination. In preliminary experiments across different illuminations, across time, and their combinations, significantly better results are achieved in both automatic and semi-automatic face recognition experiments on LED illuminated faces than on face images under ambient illuminations.

## 1 Introduction

Face has been widely adopted as a useful biometric trait for personal identification for long time. However, for practical face recognition systems, several major problems remain to be solved. The effect of variation in the illumination conditions is one of those challenging problems [10].

Existing approaches addressing this problem fall into two main categories. The first category includes methods attempting to model the behaviours of the face appearance change as a function of illumination. However, the modelling of the image formation generally requires the assumption that the surface of the object is Lambertian, which is violated for real human face. In the other category, the goal is to remove the influence of illumination changes from face images or to extract face features that are invariant to illumination. Various photometric normalization techniques have been introduced to pre-process face images, and a comparison of five photometric normalisation algorithms used in a pre-processing stage for face verification on the Yale B database, the BANCA database and the XM2VTS database can be found in [7]. Face shape (depth map or surface normal) [1] or face images in multiple spectra [5] are used in face recognition as illumination invariant features. However, face shape acquisition always requires additional devices and is usually computationally expensive. The problem with using multi-spectral images is that although invisible spectral images can be invariant to visible illumination change, there can be variation in the invisible spectra of ambient illumination.

In this paper we present a completely different approach to address the illumination variation problem. Rather than studying passively the variation of illumination itself or attempting to extract illumination invariant feature, we actively create an active and invariant illumination condition for both gallery images and probe images. Two face images are captured for every subject. The first capture is done when the LED lamp is on, and the other capture is done when LED is off. The difference of these two images is an image of the face illuminated only by the Near-IR illumination provided by the LED lamp, and is independent of environmental illumination. Meanwhile, the invisibility of Near-IR illumination ensures that the capture is non-intrusive.

The rest of the paper is organized as follows: A brief review of the previous applications of active Near-IR illumination in computer vision is presented in Section 2. Section 3 describes the hardware of capture system and the acquisition of a face database. We give the details and results of the recognition experiments performed on this face database in Section 4, and conclusions in Section 5.

## 2 Active Near-IR Illumination

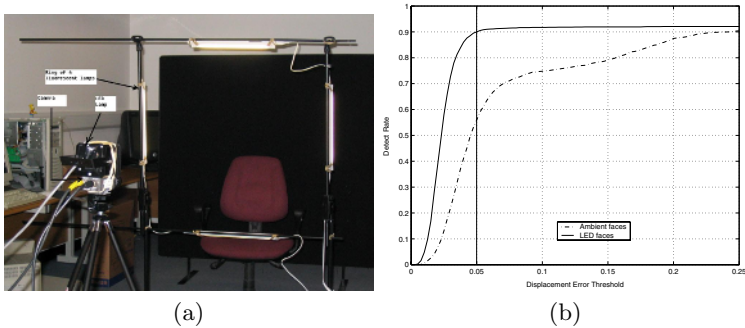
Active vision is not new in the computer vision area. In structure/coded light approaches, light patterns are projected onto object surfaces to facilitate 3D surface reconstruction. Active illumination is often used for shadow removal.

The Near-IR band falls into the reflective portion of the infrared spectrum, between the visible light band ( $0.3\mu\text{m}$ - $0.6\mu\text{m}$ ) and the thermal infrared band ( $2.4\mu\text{m}$ - $100\mu\text{m}$ ). Thus it has advantages over both visible light and thermal infrared. Firstly, since it can be reflected by objects, it can serve as active illumination source, in contrast to thermal infrared. Secondly, it is invisible, making active Near-IR illumination unobtrusive. In [9] IR patterns are projected to the human face to solve the correspondence problem in multi-camera 3D face reconstruction. Dowdall and et. al performed face detection on Near-IR face images [2]. Skin region is detected based on the fact that skin has different responses to the upper band and the lower band of Near-IR illumination. Morimoto and Flickner [6] proposed a multiple face detector which deployed a robust eye detector, exploiting the retro-reflectivity of the eyes. One Near-IR light set is used to provide bright pupil image, whilst another setting is used to generate dark pupil image, while keeping similar brightness in the rest of the scene. The pupils are very prominent and easy to detect in the difference image. Similar eye detectors using active illumination are used in [4] for 3D face pose estimation and tracking.

Although active Near-IR illumination has been widely used in face processing as detailed above, the novel idea advocated in this paper is to use it to provide constant and non-intrusive illumination for face recognition.

## 3 Face Database Acquisition

A database of face images of 40 subjects has been captured indoor. This database contains two subsets: ambient faces (faces under only ambient illumination) and



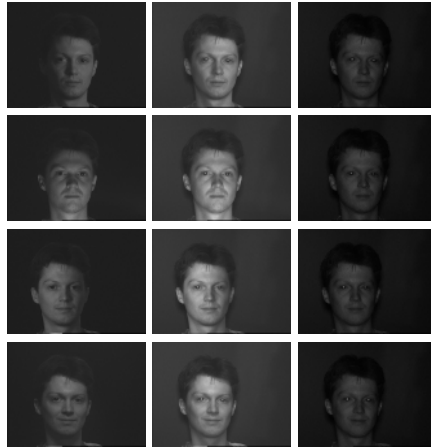
**Fig. 1.** (a) A picture of face capture system. (b) The automatic eye center detection results for LED faces and ambient faces.

LED faces (faces under only LED illumination). Two capture sessions have been conducted with a time interval of several weeks. For each session, 4 different illumination configurations are used with light sources directed individually from left, bottom, right and top. 6 recordings were acquired for each illumination configuration. LED illumination is provided by a LED lamp with peak out wavelength at 850nm. This lamp is attached close to the Near-IR sensor so that the reflective component of the Near-IR light from the eyes will be projected straight into the camera. See fig1(a). This allows us to obtain face images with prominent bright pupils. For each recording a face image under ambient illumination only and one image under combined ambient and LED illumination are captured. A LED face image is obtained by taking the difference of these two images. Therefore, we have  $40 \times 2 \times 4 \times 6 = 1920$  ambient faces and the same amount of LED faces. See [11] for more details about face capture and system setup.

## 4 Experiments and Results

### 4.1 Face Localisation

For all face images, we manually marked the eye centres as ground truth positions, and also used two different automatic localization algorithms for ambient faces and LED faces respectively. For ambient faces, we used the algorithm based on Gaussian Mixture Model (GMM) face feature detector and enhanced appearance model [3], which has been trained on 1000 images from BANCA face database. For LED faces we used a simple correlation-based localization algorithm has been applied to LED faces. We used a different approach for LED faces because usually bright pupils can be found in LED faces and they can serve as strong features for eye localization. General face detectors which have not been trained on faces with bright pupils do not work on LED faces. From the localisation errors shown on fig 1(b), it is evident that the illumination variations directly lead to the poor performance on ambient faces. With the help of the bright pupils and the consistency in LED illumination, the simple correlation-based approach gives much better results on LED faces.



**Fig. 2.** Ambient faces (the left column), combined illumination faces (the middle column) and LED illuminated faces (the right column) under 4 different illumination configurations. The ambient illumination change caused significant differences in the appearance of the whole face. All important facial features look very different in different illumination conditions. Ambient faces and LED faces are relatively dark because the aperture for the camera is adjusted to avoid the saturation of the combined illuminated faces.



**Fig. 3.** Resulting images after the histogram equalization is performed for manually and automatically registered ambient faces (top 2 rows) and for corresponding LED faces (bottom 2 rows). It is obvious that data from LED faces exhibits much less variation as compared to the data from ambient faces. Bright pupils are prominent in LED faces. There are localisation errors in some automatically registered faces.

Face images are registered according to the manually marked or automatically detected eye centre positions, then cropped and sampled to the same size ( $55 \times 50$ ). Histogram equalization is applied subsequently. Fig 3 shows some samples of faces after the histogram equalization has been performed. The resulting images are then projected to an LDA subspace obtained from XM2VTS face database.

This LDA subspace is constructed from the PCA projections of all the 2360 face images of 295 subjects in XM2VTS face database and is supposed to be a subspace focusing on discriminative information among subjects.

## 4.2 Recognition Experiments and Results

In the above LDA subspace, several different face recognition tests have been carried out on manually registered and automatically registered subsets of LED faces and ambient faces. A machine learning toolbox named WEKA [8] developed by University of Waikato has been used to perform experiments on the above data set. We applied Support Vector Machine (SVM) as the classifier, because it performed well in our previous experiments [11].

The whole dataset was divided into different subsets to serve as training sets and test sets in different test protocols. The rules of naming a subset are listed as below: 1)  $S_i$  for data in Session  $i$ ,  $i = 1, 2$ ; 2)  $C_i$  for data in Illumination Condition  $i$ ,  $i = 1..4$ ; 3)  $X_i$  for data in all illumination conditions except condition  $i$ ,  $i = 1..4$ ; 4)  $M$  for manually registered data,  $A$  for automatically registered data. For instance,  $MC_2S_1$  stands for the manually registered data in Session 1 with Illumination Condition 2,  $AX_1S_2$  stands for the automatically registered data in Session 2 with Illumination Conditions 2,3 and 4.

In the first experiment we measured the face recognition error across different sessions, and/or across different illumination conditions within each manually marked subset and within each automatically registered subset, respectively. Table 1 shows the error rates obtained under each test protocol. Each row corresponds to one test protocol. In a *Cross Session* test training set and test set are from different sessions. In a *Cross Illum.* test training set contains data with one illumination and test set contains data with the other illumination conditions. The error rate shown in the table under a specific test protocol is the average error among all tests under this protocol. For example, the test error under the first protocol is the average of the errors of 2 subtests. In one of these two subtests, data from Session 1 is used for training and Session 2 data for testing, while in the other, Session 2 data is used for training and Session 1 data for testing. It is shown for all tests that the test results on LED faces are consistently much better than on ambient faces, regardless of the way the faces were registered. The advantage that LED faces offer over ambient faces is significant. The tests on manually registered data of LED faces achieved error rates close to zero.

**Table 1.** Error in face recognition experiment 1 (in percentage)

Ambient Faces		LED Faces		Test Protocol		
Manu. Reg.	Auto. Reg.	Manu. Reg.	Auto. Reg.	Training Set	Test Set	Description
1.61	13.70	0.05	5.16	$S_i$	$S_{(3-i)}$	Cross Session
42.57	67.22	0.07	3.26	$C_i$	$X_i$	Cross Illum.
52.95	72.74	1.75	8.87	$C_iS_j$	$X_iS_{(3-j)}$	Cross Both

**Table 2.** Error in face recognition experiment 2 (in percentage)

Ambient Faces	LED Faces	Test Protocol		
		Training Set	Test Set	Description
24.95	7.92	$MS_i$	$AS_{(3-j)}$	Cross Session
60.07	7.81	$MC_i$	$AX_i$	Cross Illum.
68.14	9.53	$C_iS_j$	$X_iS_{(3-j)}$	Cross Both

It can also be shown that cross-illumination tests on ambient faces gave very poor results. Among the tests on manually registered ambient faces (see the first column), if the training data contains data with all illumination conditions, the error rate is as low as 1.61%. However if training data does not contain any illumination condition appearing in test data, the test error rate increased to 42.57%. If the training data and test data are from two different sessions, the result is even worse with an error rate of 52.95%. In sharp contrast, the test results on LED faces are consistently good for cross-session tests, cross-illumination tests and the tests involving their combination. Even in the combination test, which is the most difficult one, the test error rate for manually registered LED faces is as low as 1.75%.

Due to errors in automatic eye localization, each test on automatically registered data obtained poorer results than the same test on manually registered data. However, the increases of errors on ambient faces are much larger than those on LED faces. This is the outcome of the relatively good performance of automatic eye localization on LED faces.

The second experiment reports the results of face recognition tests across manually registered data and automatically registered data. The test protocols are the same as those in the first experiment except that manually registered data serves as training set and automatically registered data as test set. Table 2 shows the test errors. Again, test errors on LED faces are much smaller than on ambient faces. Moreover, compared to corresponding tests in the previous experiment, the test errors are similarly poor on ambient faces, and slightly worse on LED faces. The combined cross-session and cross-illumination test in this experiment represents a practical application scenario of automatic face recognition. Usually the gallery images are manually registered, while the probe images are captured at a different time, under a different illumination condition and faces are automatically registered. The error on this test for LED face is 9.53%, but for ambient faces it is 68.14%, which is extremely poor.

## 5 Conclusion and Future Work

We proposed in this paper a novel way to overcome the illumination problem in face recognition by using active Near-IR illumination. Active Near-IR illumination provides a constant invisible illumination condition and facilitates the automatic eye detection by introducing bright pupils. Significantly better results have been obtained on LED faces than on ambient faces in cross-illumination

test, cross-session and combined tests. The proposed active Near-IR illumination approach to face recognition is promising for face recognition. Further work will be the development of a more specific eye detection algorithm for Near-IR illuminated faces to improve the performance of automatic system.

## References

1. K. W. Bowyer, K. Chang, and P. Flynn. A survey of approaches to three-dimensional face recognition. In *Proceedings of International Conference on Pattern Recognition*, 2004.
2. J. Dowdall, I. Pavlidis, and G. Bebis. Face detection in the near-ir spectrum. *Image Vis. Comput.*, 21:565–578, 2003.
3. M. Hamouz, J. Kittler, J. K. Kamarainen, P. Paalanen, and H. Kalaviainen. Affine-invariant face detection and localization using gmm-based feature detector and enhanced appearance model. In *Proceedings of the Sixth International Conference on Automatic Face and Gesture Recognition*, pages 67–72, May 2004.
4. Qiang Ji. 3d face pose estimation and tracking from a monocular camera. *Image and Vision Computing*, 20:499–511, 2002.
5. S.G. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. Recent advances in visual and infrared face recognition - a review. *Computer Vision and Image Understanding*, 2004.
6. C. H. Morimoto and M. Flickner. Real-time multiple face detection using active illumination. In *Proceedings of the Fourth International Conference on Automatic Face and Gesture Recognition*, 2000.
7. J. Short, J. Kittler, and K. Messer. A comparison of photometric normalisation algorithm for face verification. In *Proceedings of the Sixth International Conference on Automatic Face and Gesture Recognition*, pages 254–259, May 2004.
8. I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with JAVA Implementations*. Morgan Kaufmann, 1999.
9. I. A. Ypsilos, A. Hilton, and S. Rowe. Video-rate capture of dynamic face shape and appearance. In *Proceedings of the Sixth International Conference on Automatic Face and Gesture Recognition*, pages 117–122, May 2004.
10. W. Zhao, R. Chellappa, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35:399–458, December 2003.
11. X. Zou, J. Kittler, and K. Messer. Face recognition using active near-ir illumination. In *Proceedings of British Machine Vision Conference*, 2005.