

# Side-View Face Recognition

Pinar Santemiz

Luuk J. Spreeuwers

Raymond N.J. Veldhuis

University of Twente

Signals and Systems Group, Department of Electrical Engineering

Drienerlolaan 5 P.O.Box 217 7500AE

Enschede, The Netherlands

p.santemiz@utwente.nl L.J.Spreeuwers@utwente.nl R.N.J.Veldhuis@utwente.nl

## Abstract

As a widely used biometrics, face recognition has many advantages such as being non-intrusive, natural and passive. On the other hand, in real-life scenarios with uncontrolled environment, pose variation up to side-view positions makes face recognition a challenging work. In this paper we discuss the use of side-view face recognition in house safety applications. Our goal is to recognize people as they pass through doors in order to estimate their location in the house. In order to preserve privacy of the people we attach our cameras on the door posts to limit the view, and deal with side-view face images. Here, we review current methods for side-view face recognition, compare available databases to be used for our task, and introduce a new method, where we use manually labeled landmarks for registration and warp the image to obtain shape-free face images. We test our system on side-view face images from CMU-Multi PIE database, and analyze our initial results.

## 1 Introduction

Face recognition is a biometric method with many applications for its nature of being non-intrusive, natural, and passive. Especially, in applications such as surveillance systems, smart homes, or any application dealing with identifying people from videos, face recognition is the primary biometrics. However, it is a challenging task to recognize faces in real-life scenarios where the environment is uncontrolled due to occlusion, expression, or pose variations.

One possible implementation area for face recognition techniques are home safety applications. The accidents and injuries happening in the home environment are mostly caused by overlooked risks, busy schedule of the parents, or external threats. Therefore, face recognition can be used to increase the situational awareness, to prevent the factors that may cause further accidents, or to detect an emergency in time.

In this paper we introduce a novel method for side-view face recognition to be used in house safety applications. Our aim is to identify people as they walk through doors, and estimate their location in a house. In our system, we will use video recordings from cameras that are attached to door posts. Due to the location of the cameras, the range of sight will be limited and the privacy of the people will be preserved.

Side-view face recognition is a challenging problem due to the complex structure of human face. A literature survey on face recognition under pose variations can be found in [20]. The first attempts to compare side-view face images were based on comparing profile curves, where fiducial points or features describing the profile were used for recognition. One such method is proposed by Gao and Leung [7], where they match profile line segments, and apply Hausdorff distance to measure similarity. They achieve 96.7% recognition accuracy on the Bern Database [6], which contains side-view face images and silhouettes of 30 people. Bhanu and Zhou [4] propose a curvature-based matching approach, where they use the curvature values to find nasion and throat

point, and then compare the curvature values in-between using Dynamic Time Warping (DTW). They achieve a recognition accuracy of 90.00% on Bern database. In a later work [21], they propose a method to construct a high resolution face profile image from low resolution videos. They use an elastic registration algorithm for alignment of profiles, and apply recognition using DTW. They experiment on 28 video sequences of 14 people walking with a right angle to the camera, and recognize more than 70% of the people correctly.

In applications where identification of people from videos is aimed, the system should be able to handle pose variations. Therefore, in many approaches where videos are involved, people make use of the texture information in addition to profile curves. Tsalakanidou *et al.* [18] present a face recognition technique based on depth and color eigenfaces, where they use the depth map for exploiting the 3D information. They experiment on XM2VTS database using 40 subjects, and recognize 87.5% of them correctly. Gross *et al.* [10] investigate recognition of human faces in a meeting room, where they propose a method called Dynamic Space Warping (DSW). Here, they apply Principal Component Analysis (PCA) [19] to vectors of sub-images from a given face, and compare these sequences using dynamic programming. They evaluate their algorithm using recordings of six meetings with six people, and achieve an accuracy of 89.4% on images without occlusion.

Instead of handling the pose variation at feature level, in some applications the images are warped or synthesized using 2D or 3D-aided systems, so that images contain the same pose as the image that is compared to. An early approach was proposed by Beymer and Poggio [3], where they generate 14 virtual views of a given face, and use them together with the original example for enrollment. They compute the correlation between images using optical flow and template matching, and achieve a recognition accuracy of 70.20% in a database with 62 people using a cross-validation methodology. Blanz and Vetter [5] estimate 3D shape and texture of faces from single 2D images using a statistical, morphable 3D model. The results on CMU-PIE and FERET show that, the algorithm achieves 95.00% and 95.90% correct identifications, respectively. Kakadiaris *et al.* [13] presents a side-view face recognition system, where they make use of 3D face models for enrollment, and extract profiles under different poses. For recognition, they extract profiles from given images and use Vector Distance Function (VDF) to match the profiles to the gallery profiles. Their system achieves a 60.00% recognition accuracy on the database UHDB1.

In this study, we first review available databases that are appropriate for our task in Section 2. We use manually labeled landmarks for registering side-view face image, and removing the background. Then we apply warping to the image to obtain shape-free face images, and use Principal Component Analysis (PCA) [19], Linear Discriminant Analysis (LDA) [2], and Local Binary Pattern (LBP) [1] to describe the face images. The details of our registration, preprocessing, and feature extraction techniques are presented in Section 3. We test our system in a subset of side-view face images in CMU-Multi PIE database [9], and analyze our results in Section 4. Finally, we will give our conclusion in Section 5, and discuss our future work.

## 2 Database

There are a number of databases that contain face images with variant poses including side-view face images [8]. Most of them are collected in a controlled environment with uniform background, artificial illumination changes, or restricted pose variations. The largest available database for side-view face recognition is CMU-MultiPIE database [9] with 337 subjects and 15 poses. It is an extension of CMU-PIE database [17], which contains only 68 subjects and 13 poses. Another database that is commonly used in side-view face recognition studies is FERET database having 200 subjects and 9 poses.

Table 1: Databases with side-view face images

Name	No. of Subjects	No. of Pose	Illumination	Occlusion	Expression	Color/Gray/3D	Image/Video
CMU-PIE [17]	68	13	43	glasses	4	Color	Img
CMU-MultiPIE [9]	337	15	18	glasses	6	Color	Img
FERET [8]	200	9	–	–	uncontrolled	Color	Img
Bern [6]	30	5	–	–	–	Gray	Img
XM2VTSDB [14]	295	many	–	–	–	Color	Vid
M2VTS [12]	37	many	–	glasses	uncontrolled	Color	Vid
MMI [15]	19	1	–	–	79	Color	Vid
UHDB1 [11]	141	17	uncontrolled	–	2	Color	Vid
	141	5	–	–	–	3D&Color	Img
Bosphorus [16]	105	13	–	4	34	3D&Color	Img

There are also some databases that are collected in more uncontrolled settings. MMI database [15] is a web-based facial expression database including 1500 samples of 19 people. In addition to the video, in some databases 3D captures, or sound files are also provided. UHDB1 database contains sixteen captures of 141 subjects, where the subject is sitting in a car, and a camera is capturing the scene at a right angle to the subject. The database also includes five 3D captures of each subject in variant poses. XM2VTS database [14] contains four recordings of 295 subjects taken over a period of four months. Each recording contains a speaking head shot and a rotating head shot. The database includes high quality color images, sound files, video sequences and a 3D Model. It is an extension of the M2VTS database, which contains voice and motion sequences of 37 people, who have been asked to count from ‘0’ to ‘9’ in their native language, and rotate their head from left to right. The Bosphorus Database [16] is a recent 3D face database that includes a rich set of expressions, various poses and different types of occlusions.

In our work, we use a subset of side-view face images from CMU-MultiPIE database, where we discarded subjects wearing glasses. In the subset that we use there are 223 subjects with a total of 552 side-view face images. The images have a resolution of  $640 \times 480$  pixels, where the background and illumination is constant.

### 3 Feature Extraction

In our feature extraction approach, we first register images using manually labeled landmarks. Then, we apply some preprocessing to the registered images to remove the background, and apply warping to achieve shape-free texture images. To describe the face images, we implement Local Binary Pattern (LBP), and two baseline algorithms, Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA).

#### 3.1 Registration

For registration, we manually label 13 landmark points on each face image as seen in Figure 1. We use Procrustes analysis to find the transformation parameters between each image. First, we align the landmarks of the images in the training set to find the average landmarks. Then, we compute the transformation between each image and the average landmarks, and transform images, accordingly.

Finally, in order to have fixed sized images we place a bounding rectangle around the face, and crop the image. In order to find the parameters of the bounding rectangle, we compute the distance between the nasion and the nose tip using the average landmarks,



Figure 1: Manually labeled landmark points.

and use this distance multiplied with a constant to decide on the width and height of the rectangle. Moreover, we locate the rectangle using the average nose tip landmark, and crop all images accordingly. An example can be seen in Figure 2(a).

### 3.2 Preprocessing

After obtaining registered images, we apply some preprocessing to the images to eliminate the effect of background. Here we make use of the landmarks, and interpolate them to estimate the profile. Using this profile information, we remove the background, and get a cropped image that contains both the shape information and the texture information. In order to obtain a shape-free texture image, we warp the image by stretching the profile until the edge of the bounding box. After warping, we smooth the image using a low-pass filter. Finally, we draw a circle with its center in the middle of the right edge of the bounding rectangle and remove the area outside that circle. Sample images for our preprocessing method are given in Figure 2, where the first image shows the result of registration, the second image shows the image after removing the background, and the last image shows the output after warping.

### 3.3 Baseline Algorithms

As baseline algorithms for face recognition we used PCA (Eigenface approach) and LDA (Fisherface approach). PCA is an unsupervised method that is used to reduce the dimensionality of the feature space by projecting face images onto a space that spans the significant variations among known face images [19]. LDA is an extension of the Eigenface approach. It is a supervised method that is used for further dimensionality reduction and classification [2].

In our implementation, we used PCA and LDA for extracting features from face images. We used vectorized pixel values of the face images in a training set to learn PCA parameters. Then, we projected each face image into the PCA space, and using the projected values of the training samples we learned LDA parameters. For classification, we use one nearest neighbor method using cosine similarity measure.

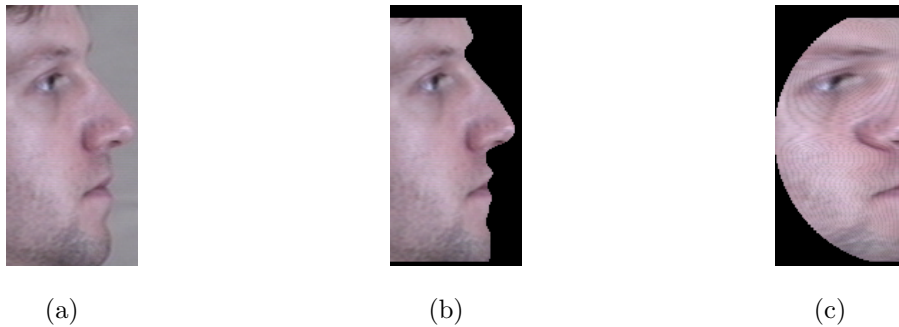


Figure 2: Preprocessing steps. (a) Registered image. (b) Cropped image. (c) Warped image.

### 3.4 Local Binary Pattern (LBP)

Local Binary Pattern is a feature description method which is used for describing local spatial structure of an image [1]. It is widely used in face recognition applications due to its invariance against illumination changes, and its computational simplicity. In our system, we divide the images into 75 subregions, and compute the LBP histograms for each region. Then, we concatenate these histograms to achieve the feature vector of the image. For classification, we use one nearest neighbor method using Chi square distance measure.

## 4 Experimental Results

In our experiments, we use a subset of side-view face images in CMU-Multi PIE database, where we eliminated the subjects wearing glasses. We use a total of 552 side-view face images from 223 subjects. The images were acquired in a controlled environment with constant background and illumination, and have a resolution of  $640 \times 480$  pixels. We divided this set into three subsets: a training set with 97 subjects and 280 images, an enrollment set with 126 subjects and one image for each subject, and a test set with 79 subjects who are included in the enrollment set, and a total of 146 images.

We use three types of images to extract features from: 1) registered image, 2) cropped image, and 3) warped image. An example can be seen in Figure 2. Using PCA, LDA, and LBP we perform identification experiments. We achieve 91.10% recognition accuracy when we use cropped images and LBP for describing face images. Our rank-one accuracies can be seen in Table 2, and the Cumulative Match Characteristic (CMC) curves can be seen in Figure 3. In order to better analyze these results, we also investigated the erroneous cases. Some examples can be seen in Figure 4.

Table 2: Rank 1 Identification Performances

	Registered Image	Cropped Image	Warped Image
<b>PCA</b>	59.59	44.52	50.00
<b>LDA</b>	77.40	63.01	69.18
<b>LBP</b>	89.04	91.10	84.93

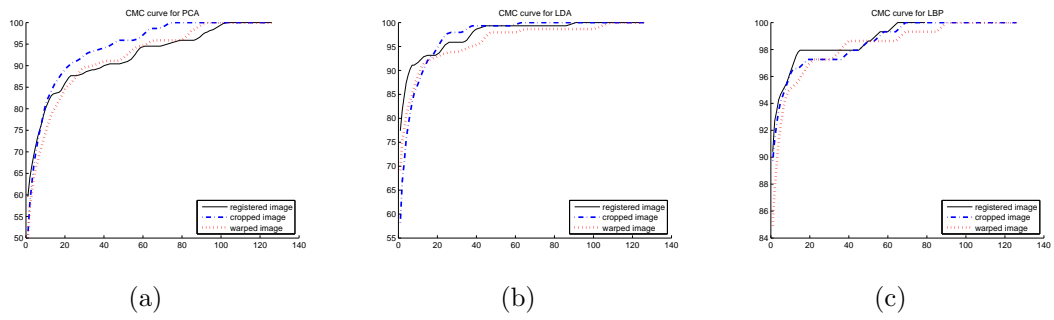


Figure 3: Cumulative Match Characteristic (CMC) curves. (a) CMC curve of PCA, (b) CMC curve of LDA, (c) CMC curve of LBP.



Figure 4: Misclassification examples. The three images from left to right are the test image, corresponding enrollment image, and nearest image found by the classifier.

When we analyze these results, we see that LBP consistently perform better than PCA and LDA. It has been shown that compared to holistic methods, LBP is less sensitive against variations that occur due to illumination, expression, or pose [1]. LBP describes the image by dividing it into local regions, extracting texture descriptors for each region independently, and then combining these descriptors to form a global description of the image. Consequently, LBP is not effected by small local changes as much as PCA or LDA.

In Figures 4, we see misclassification examples caused by hair, illumination, expression, and facial hair. In examples shown in Figures 4(a), 4(c), and 4(e), all methods failed to recognize the person correctly. However, in examples shown on the right side, LBP succeeded to identify the person. When we compare these examples, we see that when the variation is limited LBP can deal with these problems up to a certain point.

When we use cropped or warped image instead of registered image, the accuracies decrease dramatically, except for LBP. It must be noted that variations on pose or expression are highly effective on the profile. Moreover, using manually labeled landmarks to find the profiles may cause some inaccuracies. Since PCA and LDA are highly effected by these variations, and using cropped images emphasize the effect of the profile, we observe a drop on the accuracies. On the other hand, when we look at the CMC curves, we see that on higher ranks cropped images perform better than registered images. This shows that using more accurate shape descriptors will improve our results.

We also see that the accuracies using warped image is higher than using cropped image for PCA and LDA. This indicates that, our warping algorithm is effective in obtaining shape-free face images, and since we crop the image using a semi-circle, we eliminate the effect of hair up to some extent.

## 5 Conclusion and Future Work

In this work we investigate side-view face recognition to be used in house safety applications, where we aim to identify people as they walk through doors, and estimate their location in a house. Here, we present our initial results that we achieved using side-view face images from CMU-Multi PIE database. We use manually labeled landmarks to register images, and apply some preprocessing to obtain shape-free face images. We test our system using PCA, LDA and LBP, and achieve 91.10% recognition accuracy using LBP.

We see that, our warping algorithm is useful for eliminating the effects of hair, and to obtain a shape-free face image. However, it is much affected by the variations of pose or expression. In the future, we aim to find the profile and the landmarks automatically to improve our registration and warping algorithms. We also think of using an elastic registration algorithm to deal with pose variations, in order to obtain more accurate shape descriptors.

## 6 Acknowledgement

This work is supported by GUARANTEE (ITEA 2) 08018 project.

## References

- [1] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Transactions on PAMI*, vol.28, pp.2037–2041, 2006.

- [2] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transactions on PAMI*, vol.19, pp.711–720, 1997.
- [3] D. Beymer, and T. Poggio, "Face recognition from one example view," *Computer Vision, IEEE Int. Conf. on*, pp.500–507, 1995.
- [4] B. Bhanu, and X. Zhou, "Face Recognition from Face Profile Using Dynamic Time Warping," *Int. Conf. on Pattern Recognition (ICPR)*, vol.4, pp.499–502, 2004.
- [5] V. Blanz, and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Transactions on PAMI*, vol.25, pp.1063–1074, 2003.
- [6] <ftp://ftp.iam.unibe.ch/pub/Images/FaceImages> .
- [7] Y. Gao, and M. Leung, "Line segment Hausdorff distance on face matching," *Pattern Recognition*, vol.35, pp.361–371, 2002.
- [8] R. Gross, "Face Databases," *Handbook of Face Recognition*, pp.301–327, 2005.
- [9] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-PIE," *Image and Vision Computing*, vol.28, pp.807–813, 2010.
- [10] R. Gross, J. Yang, and A. Waibel, "Face Recognition in a Meeting Room," *IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.294, 2000.
- [11] <http://cbl.uh.edu/URxD/datasets/> .
- [12] <http://www.tele.ucl.ac.be/PROJECTS/M2VTS/m2fdb.html> .
- [13] I.A. Kakadiaris, H. Abdelmunim, W. Yang, and T. Theoharis, "Profile-based face recognition," *IEEE Int. Conf. on Automatic Face and Gesture Recognition (FG '08)*, pp.1–8, 2008.
- [14] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "XM2VTSDB: The Extended M2VTS Database," *Int. Conf. on Audio and Video-based Biometric Person Authentication*, pp.72–77, 1999.
- [15] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-Based Database for Facial Expression Analysis," *IEEE Int. Conf. on Multimedia and Expo*, pp.317–321, 2005.
- [16] A. Savran, N. Alyuz, H. Dibeklioglu, O. Celiktutan, B. Gokberk, B. Sankur, and L. Akarun, "Bosphorus Database for 3D Face Analysis," *Biometrics and Identity Management* , vol.5372, pp.47–56, 2008.
- [17] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *IEEE Transactions on PAMI*, vol.25, pp.1615–1618, 2003.
- [18] F. Tsalakanidou, "Use of depth and colour eigenfaces for face recognition," *Pattern Recognition Letters*, vol.24, pp.1427–1435, 2003.
- [19] M. Turk, and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol.3, pp.71–86, 1991.
- [20] X. Zhang, and Y. Gao, "Face recognition across pose: A review," *Pattern Recognition*, vol.42, pp.2876–2896, 2009.
- [21] X. Zhou, and B. Bhanu, "Human Recognition Based on Face Profiles in Video," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, pp.15, 2005.