Improved Face Recognition At A Distance Using Light Field Camera & Super Resolution Schemes

R. Raghavendra* Kiran B. Raja*† Bian Yang* Christoph Busch*†
{raghavendra.ramachandra, kiran.raja, bian.yang, christoph.busch}@hig.no

*Norwegian Biometrics Laboratory
Gjøvik University College
2802 Gjøvik, Norway

†Hochschule Darmstadt - CASED
Haardtring 100,
64295 Darmstadt, Germany

ABSTRACT

In this paper, we present an empirical study on exploring the Light Field Camera (LFC) for identifying multiple faces present at different distance. Since LFC can render multiple focus images in single exposure, one can combine these multiple images to obtain single all-in-focus image. Thus the constructed all-in-focus image will have all regions in focus and hence allows one to capture more information about the subject present even at a far distance. At the same time one can also construct the super resolution image to further improve the face recognition at a distance. Thus, in this work, we explore both all-in-focus and super resolution schemes to evaluate the multiple face recognition at a distance using LFC. We carry out extensive experiments on light field face dataset and present both qualitative and quantitative results

Categories and Subject Descriptors

D.4.6 [**Security and Protection**]: Access controls, Authentication

General Terms

Biometrics, Pattern recognition, Computational imaging

Keywords

Biometrics, face recognition, light field camera, surveillance, identification, super-resolution

1. INTRODUCTION

Increasing interest in face based biometrics for surveillance systems in commercial, security and defense systems can be attributed to the robustness and ease of face acquisition that can be captured with or without the co-operation of the subject. Inspired from these advances, non-intrusive face biometric research is gaining pace. Recognizing a subject at a distant range (about $10m - 20m$) without subjects' co-operation is termed as Face Recognition at a Distance (FRD) [17]. FRD presents number of challenges including varied illumination, non-cooperative pose, limited field of view, out of focus imaging among many others in recognizing the subject due to unconstrained capture. Some of the earlier works have used new sensors [17] [6] or proposed new algorithms to improve the acquired image quality [1]. However, proposed algorithms are closely coupled to the sensors employed for acquiring the samples. Using multiple cameras to cover wide and narrow field of view [17] or employing a pan, tilt and zoom (PTZ) camera [2] or a still camera with a telephoto lens [6] are some of the popular methods for FRD. Although, these sensors or techniques are reported to perform well in recognizing single subject at a time, they fail in accurately recognizing multiple subjects present at different distance. Limited depth-of-field and focus in conventional camera fails to capture the best focused images corresponding to multiple subjects present in different image planes.

This work aims at exploring the strength of Light Field camera (LFC) to address the problem of recognizing multiple faces with different poses present in a given scene at various distance using all-in-focus and super resolution schemes. The principle of LFC involves in capturing the image by sampling the 4D light-field on its sensor in a single photographic exposure by inserting a micro-lens array [7] or a pin-hole array [4] or masks [15] between the sensor and main lens. Light-field camera captures scene at different focus (or depth) in single exposure without any additional motion of lens to set the focus on objects in different focal planes. This allows to refocus a particular image or obtain an all-in-focus image after acquisition. We exploit the availability of depth images for biometric applications and especially for FRD.

Recent work [12] on light field images for face recognition has demonstrated the superiority over the conventional imaging system. Based on the results presented in [12], we are motivated to study the applicability of LFC for multiple face recognition in various distance range by exploring all-in-focus and super resolution schemes. To this extent, We have employed the first available light field face data-base [10] to evaluate the performance of well-know face recognition algorithm based on the Local Binary Pattern (LBP) and Sparse Representation Classifier (SRC) on both all-in-focus and super resolution samples. The main contributions of this paper are in: (1) Empirical analysis of multiple face recognition at a distance using all-in-focus and super resolution images constructed from the multiple depth images rendered by LFC; (2) Experimental evaluation of face identification at various distance using LFC.

SIN '13, Nov 26-28 2013, Aksaray, Turkey
2013 ACM 978-1-4503-2498-4/13/11.
http://dx.doi.org/10.1145/2523514.2523572.

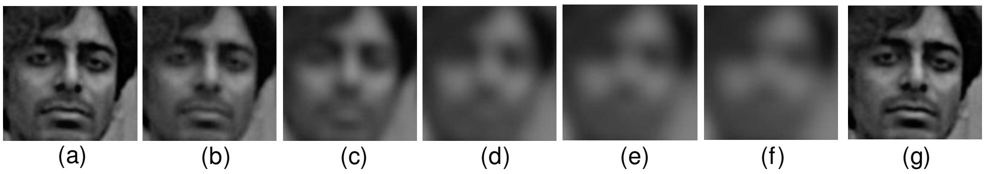

Figure 1: Depth images obtained using LFC (a)-(f) Different depth images; (g) All-in-focus image

The rest of the paper is organized as follows: Section 2 will discuss the proposed method, Section 3 discusses the super resolution techniques involved and Section 4 briefs about the feature extraction and identification. Section 5 presents the results and Section 6 draws the conclusion.

2. PROPOSED METHOD

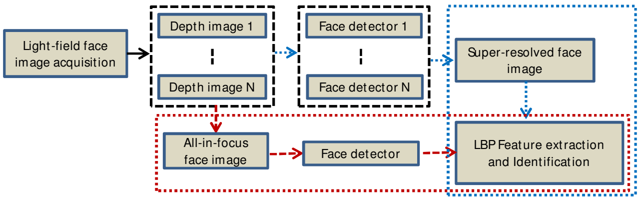

Figure 2: Block diagram of the proposed scheme

Figure 2 illustrates the proposed method that can be visualized in three steps: (1) Obtain all-in-focus image using different depth images rendered by the LFC. (2)(a) Perform face detection on the all-in-focus image to extract the face region present in the scene. (2)(b) Perform face detection on individual depth images and compute super-resolved face image. (3) Finally, carry out the feature extraction and identification independently on all-in-focus and super-resolved face image.

Each image acquired using LFC results in a set of depth images corresponding to different image plane. Each capture produces a raw file that consists of set of images that are focused at different depths in the scene. These multi-focus images will have only one particular region (or area) in focus based on the imaging plane. In addition to these multi-focus image, the raw file also hosts metadata with information about the regions in focus in each of the depth image. This meta information is used to select the best focused region from multiple focus (or depth) images to obtain an all-in-focus image. Figure 1 (a) - (f) illustrates the face image in different depth plane. Figure 1 (g) shows the all-in-focus image. It can be observed that the face loses focus if the imaging plane is different from the plane where the face is present.

After obtaining all-in-focus image, we perform the face detection using Viola-Jones face detector [16] on all-in-focus and different focus images individually. Viola-Jones face detector [16] is chosen due to its robustness and performance in real-time scenario. We have trained the face detector using 2429 face samples and 3000 non-face samples. We then segment each of the face region detected using face detector. This provides an all-in-focus face image and a set of face images corresponding to multiple depth which are further used to obtain super-resolved face image. Finally, the obtained all-in-focus face image and super-resolved face image is used to extract features and identify the subject.

3. SUPER RESOLUTION TECHNIQUES

To exploit the availability of depth images rendered by light field camera, we employ four different super-resolution schemes [5, 13, 9, 3, 19] to construct a high quality super-resolved face image. We present summary of these techniques in the following section.

Iterative Back Projection Method (BP): Iterative back-projection [5] method uses image registration in an iterative manner to achieve high resolution image with sub-pixel accuracy. With an initial guess of high resolution image, difference between various low resolution images are back-projected in iterative way to reduce the error between initial guess and low resolution images.

Projection Onto Convex Sets (POCS): In Projection Onto Convex Sets (POCS) method [13], high-resolution image is constructed from a set of low resolution images using the coarse detector data through a series of shifts and rotation along with a combination of apriori information.

Papoulis' and Gerchberg's (PG) Method: PG method [9, 3] can be considered as a special case of the POCS that uses the apriori information in an iterative approach to compute super-resolution image by reducing the energy of error. Each iteration consists of projecting high frequency and low frequency components followed by low pass filtering. The known pixel values in the image are reprojected onto a new high resolution image.

Robust Super-Resolution (RSR): Robust Super Resolution technique[19], is an iterative robust median estimator based approach to discard the outlier data in the imaging model. Motion blur, object motion, parallax etc, in image are taken into consideration to estimate the high resolution image neglecting the differences arising due to aliasing.

4. FEATURE EXTRACTION AND IDENTIFICATION

In this work, we have employed Local Binary Pattern (LBP) [8] for feature extraction and Sparse Reconstruction Classifier (SRC)[18] to achieve robust identification of the subject. Both of these techniques are well known for the performance and applicability in face biometrics [14].

Local Binary Pattern (LBP) : LBP algorithm is well known to extract face features. LBP thresholds the differences of the center value and the neighbourhood in a 3 x 3 grid for one pixel. Output value for a pixel is regarded as an 8-bit binary number representing the pixel. The descriptor of the given image is constructed by the histogram of these binary numbers in the whole image [8]. We have employed the LBP operator with a radius of 2 (chosen based on experimental trials).

Sparse Reconstruction Classifier (SRC): Sparse representation has been proven useful in analyzing face biometrics. Sparse representation can handle noisy data, illumination variance and occlusions. We have employed the sparse representation for our present work and we carry out L1 - minimization via SPGL1 solver based on spectral gradient projection [18]. We have obtained the comparison scores that directly correspond to the residual errors obtained using SRC.

We have used the combination of LBP feature extraction scheme and sparse representation classifier (LBP-SRC). We evaluate the probe samples using this algorithm and discuss the results in the following section.

5. EXPERIMENTS AND RESULTS

In this section, we present and discuss the results of LBP-SRC and super-resolution schemes on light-field database. This work is carried out on the newly acquired light field face database [10] consisting of a total of 986 face samples from 25 subjects. The samples are distributed in three different scenarios acquired in indoor (140), corridor (63) and outdoor (100) settings, totalling to 303 probe samples. Obtained all-in-focus image and super-resolved face image from depth images of light field camera are used as probe images. The reference images are obtained using DSLR camera which resulted in a total of 200 high quality samples. For detailed description of employed database, readers can refer [10]. Figure 3 shows the example probe images obtained using all-in-focus and super-resolution techniques.

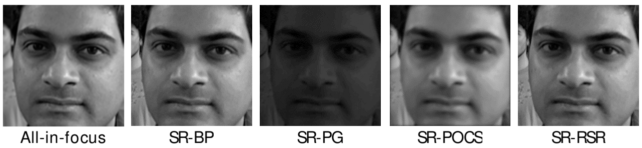

Figure 3: Probe samples for identification obtained from different techniques.

In this work, all the results are presented in terms of identification rate (rank 1) that is obtained by comparing $1:N$ subjects in the dataset, therefore a higher value of the identification rate correlates to better performance. In order to effectively evaluate the face recognition algorithms on light field dataset, we employ reference samples from enrollment dataset and probe samples correspond to those acquired using LFC. Figure 4 shows the qualitative results of the employed face detector on the LFC acquired samples (all-in-focus). It can be observed that, the face detector can even detect the face that are present at the farthest distance from the camera. This fact can be attributed to the all-in-focus functionality of the LFC. In order to obtain the quantitative value of the face detection rate, we carry out the face detection on all probe samples that showed the detection rate of 94.8%. In this work, we employed the likelihood based scheme mentioned in [11] to reduce the false positive of the employed face detector.

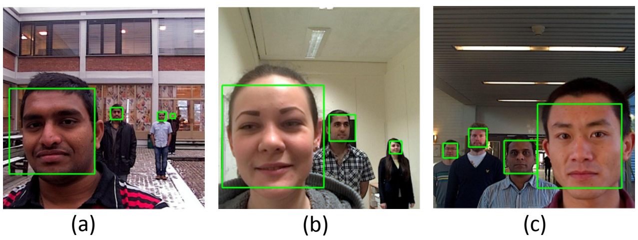

Figure 4: Face detection results (a) Outdoor (b) Indoor (c) Corridor (best viewed in color)

Method	Distance	Identification Rate (%)		
		Indoor	Corridor	Outdoor
All-in-focus	Near	**38.98**	**36.11**	**26.81**
	Middle	41.10	**50.08**	**53.62**
	Far	19.92	13.89	19.57
SR-BP	Near	20.66	17.41	23.45
	Middle	**43.80**	46.08	45.44
	Far	35.53	**36.52**	32.12
SR-PG	Near	20.10	19.50	19.58
	Middle	41.05	20.65	47.55
	Far	**39.02**	18.77	**32.86**
SR-POCS	Near	20.16	15.78	22.99
	Middle	41.93	49.12	49.94
	Far	37.91	35.08	31.08
SR-RSR	Near	20.83	21.42	22.64
	Middle	42.91	48.71	46.1
	Far	36.25	31.3	31.44

Table 1: Performance of various schemes

Since, each probe sample is acquired in the distance of $0.5m - 20m$, we divide this range into three parts to extensively analyze the performance of LFC in Near (or short) distance ($0.5m - 2m$), medium distance ($2m - 12m$) and far (or long) distance ($12m - 20m$). Table 1 shows the identification results achieved using LBP-SRC on all-in-focus face image and super-resolved face image in 3 different scenarios. It can be observed that identification achieved using all-in-focus images show the best result in all three scenarios at near and middle distance range. The obtained identification rate is 38.98% (indoor), 36.11% (corridor) and 26.81% (outdoor) in near distance range and 50.08% (corridor), 53.62% (outdoor) in mid-distance range. However, BP super-resolution scheme shows better performance in indoor conditions at mid-distance with 43.80% identification accuracy and 36.52% identification in corridor scenario at far distance. Similarly, PG super-resolution yields better identification compared to other methods in far distance scenario with 39.02% (indoor) and 32.86% (outdoor) accuracy.

It can be observed from Table 1 that LFC is capable of identifying large number of subjects at mid-distance in corridor and outdoor scenario. Alternatively, super-resolution techniques show the best performance at far distance. Irrespective of the super-resolution technique used, all-in-focus based identification stands out with high performance in the near distance range. Another aspect of varied pose of subjects in the near range can be attributed to lower identification rate. From our observation, the light field database consists of samples acquired with varying poses. The overall identification accuracy is expected to boost provided there is increase in spatial resolution of LFC. Apart from this, LFC offers a wide spectrum of advantages like extended depth of field without reducing the aperture, small in size and low cost. A combination of all-in-focus and super-resolved face image will boost face based biometric applications to the next level.

6. CONCLUSION

This paper has explored all-in-focus property of LFC and four well known super-resolution techniques to achieve multiple subject identification present at different distance. Extensive experiments carried out on GUCLF[10] database indicates the strength of LFC in identifying multiple subjects present at various distance. From our experiments, it can be argued that LFC is promising for face biometrics. Best identification rate of 53.62% is noted for outdoor scenario at mid-distance range. On a more general note, all-in-focus image based face recognition stands out in near and mid distance range in indoor, corridor and outdoor scenario. Thus, LFC can perform well in near and mid-distance range and light field based super resolved images perform well in far distance. A combination of all-in-focus and super-resolved image can be employed to perform robust face recognition at a distance even under the presence of multiple subjects in the scene. Critical observations from the experiments suggest that distance is a constraint for face recognition. Near distance $(0.5 - 2m)$ and far distance $(12 - 20m)$ do not provide high recognition rate. Future works on this database could include the decision level fusion using all-in-focus and super-resolved image to achieve higher identification in all three scenarios and different distance range.

Acknowledgment

This work was funded by the EU 7^{th} Framework Program (FP7) under grant agreement n^{o} 284862 for the large-scale integrated project FIDELITY.

7. REFERENCES

[1] N. AL-Hassan, S. A. Jassim, and H. Sellahewa. Enhancing face recognition at a distance using super resolution. In *Multimedia and security*, pages 123–132, 2012.

[2] H.-C. Choi, U. Park, and A. Jain. PTZ camera assisted face acquisition, tracking and recognition. In *IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1 –6, 2010.

[3] R. Gerchberg. Super-resolution through error energy reduction. *Journal of Modern Optics*, 21(9):709–720, 1974.

[4] R. Horstmeyer, G. Euliss, R. Athale, and M. Levoy. Flexible multimodal camera using a light field architecture. In *International Conference on Computational Photography (ICCP)*, pages 1–8, 2009.

[5] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graphical models and image processing*, 53(3):231–239, 1991.

[6] H. Maeng, S. Liao, D. Kang, S.-W. Lee, and A. K. Jain. Night-time face recognition at long distance: Cross-distance and cross-spectral matching. In *Asian Conference on Computer Vision (ACCV)*, pages 1 –14, 2012.

[7] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand- held plenoptic camera. In *Stanford University Computer Science Technical Report CSTR 2005-02*, pages 1–11, 2005.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971 –987, 2002.

[9] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and Systems*, 22(9):735–742, 1975.

[10] R. Raghavendra, Kiran B. Raja, B. Yang, and C. Busch. GUCLF: A new light field face database. In *RIAO/OPTILAS 2013, VIII Iberoamerican Optics Meeting and XI Latinamerican Meeting on Optics, Lasers and Applications*, pages 1–5, 2013.

[11] R. Raghavendra, B. Yang, and C. Busch. Accurate face detection in video based on likelihood assessment. *Norsk informasjonssikkerhetskonferanse (NISK)*, 2012, 2012.

[12] R. Raghavendra, B. Yang, Kiran B. Raja, and C. Busch. A new perspective - face recognition with light-field camera. In *6th IAPR International Conference on Biometrics (ICB)*, pages 1–8, 2013.

[13] H. Stark and P. Oskoui. High-resolution image recovery from image-plane arrays, using convex projections. *Journal of Optical Society of America A*, 6(11):1715–1726, 1989.

[14] S.Z.Li and A. Jain. *Handbook of Face Recognition*. Springer-verlag edition, 2004.

[15] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.*, 26(3):1–12, July 2007.

[16] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:137–154, 2004.

[17] F. Wheeler, R. Weiss, and P. Tu. Face recognition at a distance system for surveillance applications. In *IEEE International Conference on Biometrics: Theory, Applications and Systems*, pages 1 –8, 2010.

[18] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210 – 227, 2009.

[19] A. Zomet, A. Rav-Acha, and S. Peleg. Robust super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–645, 2001.