Face Identification from Unconstrained Settings and Occlusions

D. Silambarasan, Anju Mary Thomas

Department of EEE, Anna University Regional Campus, Coimbatore, India

Abstract— *The images taken in the ordinary mobile phone cameras are often taken in unconstrained settings. Because of this the identification of face will sometime become very difficult. The paper presented below try to overcome the above conditions of identifying the face from non-uniform illumination, pose, blur and also partial occlusions. Blurred face is modelled as convex combination of transformed focused gallery images .This is extended to those images with illumination variations, which are forming a bi-convex set and finally the variations in pose are also included. For handling the occlusions occurring in the images, the basic characteristics of the human face is taken into consideration. The robustness is increased by the transformation spread function and local binary pattern is added to it that can identify the person even with 70% occlusion in his photograph.*

Keywords—Blur, Different Pose, Non Uniform Illumination, Occlusion, Transformation Spread Function,

I. INTRODUCTION

Identification of face will become very hard when the image is captured under unconstrained settings [1][12].These settings can be blur, changes in the illumination source, pose or expression and the partial occlusions occurring in the images. These kind of problems occur mainly in the images taken in ordinary mobile phone cameras. The paper proposes methods to recover the images even from the above environments too. The process of face identification work with focused gallery images which are already stored in the memory.

Normally, the blur in the image occurring due to the camera shake is modelled as a convolution with a single blur kernal and the blur is assumed to be uniform all over the image [2] [3].But the space variant blur that is found in hand-held cameras can't be assumed like the above. In most of the recent methods of image restoration model the motion- blurred image as an average of transformed image [4] [8]. There are four major methods which are used to identify the image from blurring .In the first method the input blurred image is deblurred and then used for recognition. The major challenge to be faced in this is blind image deconvolution that might be infeasible while considering the complete space of blur kernals. For the second method, blur invariant features are extracted from the blurred image and then taken for recognition. The local phase quantization is the method used for the above purpose. It is applicable to small blurs and not effective for large blurs. The third method used is called the direct recognition approach. In this blurred version of the gallery image is created artificially and the blurred probe image is matched to them. But the approach lags in identifying the whole space of blur kernals.

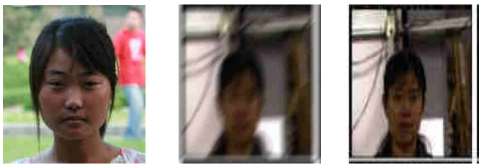
The fourth method is the joint deblur and recognize the face image. It involves solving for the original sharp image, blur kernel and then identify the face image and thus the method is computationally intensive. As far as the illumination aspect is concerned two approaches are used one based on the 9D subspace model for face and other is the extraction and matching of illumination insensitive facial features. The techniques used for face recognition handling pose is classified into 2D and 3D techniques. A dictionary based approach which can handle illumination and pose is proposed in [9].

It is assumed that only a single focused gallery image is available. The face is assumed to possess a uniform planar structure and the blurred face is modelled as the weighted average of the available focused gallery image. Each one of these is assigned a weight which is the exposure duration of the transformation .These weights are referred to as the Transformation Spread Function (TSF).

The face recognition is based on the TSF model. All the transformations that exist in the 6D space is applied to the focused gallery image and the resulting images that's obtained represented as a column matrix. It can be showed that the set of all such images forms a convex set. In order to identify the original person from the blurred image, the distance between the probe and the convex combination is found. The focused image which has the minimum distance to that of probe image is identified as the match. The camera motion while taking the photograph is modelled as l_1-norm constraint on the TSF weights. By minimizing the cost function an estimate of the blurred probe image corresponding to the gallery image is found out. The gallery image blurred with the

ISSN: 2349-6495

corresponding TSF is compared with the probe whose LBP values are also extracted.



(a) (b) (c)

(d)

Fig.1:(a) Focused image (b) Blurred image(c) Different pose (d) Occluded image

The 9D subspace model and the bi-convexity property is used to identify the face from blur and illumination. An initial estimate of pose for the images in the gallery and comparing it with the available input pose, the match is found out. For handling the occlusions in the image the distance between the eyes, nose and lips is considered which will be unique to a person.

II. CONVOLUTION MODEL

In the convolutional model a pixel which is blurred is assumed as the average of the pixels in the neighborhood of the original image. Hence the blur in the images captured by the camera is assumed as a convolution of a single blur kernal on all over the image. Let the original image be I and Y be the blur kernal of size (2k+1)*(2k+1), then the blurred image, I_b :

$$I_b(r,c) = I * Y(r,c) = \sum_{i=-k}^{k}\sum_{j=-k}^{k} Y(i,j)I(r-i,c-j) \qquad (1)$$

where * represents the convolution operator and r,c are the row and column indices of the image. The convolution model is enough for explaining the in-plane camera translations but it is not capable to identify images from out-of-plane and in-plane rotation.

III. MODELLING OF MOTION BLUR

General camera motion will result in blurring of the images captured. An optimization algorithm is used to recover the images from these non-uniform motion blur. As mentioned earlier the blurred face can be modelled as the averaging of the set of all possible transformations that can occur by the relative motion of the camera and the object which is being photographed [10].

Fig.2: Image showing motion blur while taking photograph

For each of the focused gallery image, it can be assumed that there are N classes of possible blurred images. The reconstruction error t_m of the blurred image, by making use of the focused gallery image can be obtained by solving the equation below:

$$t_m = \min_{h_T} \|k - A_m h_T\|_2^2 + \beta\|h_T\|_1 \qquad (2)$$

subject to $h_T \geq 0$

where k is the focused gallery image , m=1,2,3…..N ,A_m is the transformation matrix of the image and h_T is the transformation spread function (TSF)

Not all the regions in face convey the same amount of information, the equation (1) can be modified by introducing a weighting matrix W, when computing the reconstruction error as:

$$t_m = \min_{h_T} \|W(k - A_m h_T)\|_2^2 + \beta\|h_T\|_1 \qquad (3)$$

subject to $h_T \geq 0$

where W is a diagonal matrix. As the t_m metric is sensitive to very small pixel misalignments, the Local Binary Patterns (LBP) is used. Then the optimal TSF is given by the equation.

$$h_{T_m} = \operatorname{argmin}_{h_T} \|W(k - A_m h_T)\|_2^2 + \beta\|h_T\|_1 \qquad (4)$$

subject to $h_T \geq 0$

For each of the focused gallery image, using the optimal TSF, blurring is performed. Now for each of the focused gallery image and the blurred image, division into non-overlapping rectangular patches is done. The LBP [11] histogram is extracted for each of the patches independently. As the image is divided into blocks the face can be viewed now as the combination of different patches. The algorithm described below can be now used to recover the original image.

Algorithm 1:

The input for the algorithm is blurred image and a set of focused gallery images are available

1. For each of the gallery image f_m the optimal TSF value is found out

ISSN: 2349-6495

2. Blur each of the gallery image with its respective TSF and extract the LBP features.
3. The LBP features of the probe image and those of the transformed gallery images are compared and a closest match is found.

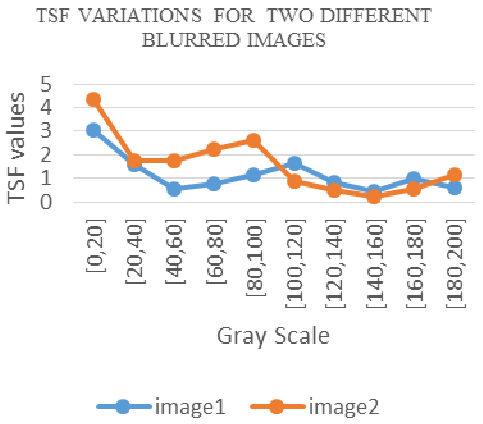


Fig.3: The graph showing the variations in TSF values of two different individuals having blurring in their photographs

IV. MODELLING OF ILLUMINATION VARIATIONS, POSE AND OCCLUSIONS

4.1 Illumination Variations

Lighting or illumination is the usage of light to achieve a practical or aesthetic effect. Lighting could be done by the use of both artificial light sources namely lamps and light fixtures, as well as natural illumination like capturing daylight. The light coming through windows, skylights, or light shelves is used as the main source of light during daytime in buildings. This can save energy rather than using artificial lighting, which is the major component of energy consumption in buildings. Proper lighting makes task performance easier, also improve the appearance of an area, and also give positive psychological effects on occupants. An object is observed by the human eye when a part of the incident light striking its surface is reflected back to the human eye. The face is modelled as a convex Lambertian surface, and it is assumed that there exists a configuration of nine light source directions so that the subspace formed by the images taken under these nine sources is very effective in identifying face in different ranges of lighting conditions.

Fig.4: Example of improperly illuminated image

An image of a person, j under any illumination can be written as:

$$j = \sum_{i=1}^{9} \alpha_i j_i \qquad (5)$$

The nine basis images i.e. f_i s for the 9D subspace is modelled as a Lambertian reflectance model as:

$$j_i(r,c)=\mu(r,c)\max(n(r,c)^T s_i,0) \qquad (6)$$

where μ and n are the albedo and the surface normal respectively at the pixel location (r,c) and s is the illumination direction and n is the face normal.

For modelling space-invariant blur, the set of all images under varying illumination and blur forms a bi-convex set, i.e., if either the blur or the illumination is fixed, the resulting subset is convex. As mentioned, according to the motion blur model for faces, the set of all motion-blurred images that is obtained by blurring a focused gallery image using the TSF model also forms a convex set. Therefore, the set of all images under varying illumination and non-uniform motion blur context= also forms a bi-convex set. The solution for this is the minimization of the following cost function given by

$$[h_{T_m}, \alpha_{m,i}] = \operatorname{argmin}_{h_T,\alpha_i} \left\| W\left(k - \sum_{i=1}^{9} \alpha_i A_{m,i} h_T\right) \right\|_2^2 + \beta \|h_T\|_1$$

$$\text{subject to } h_T \geq 0 \qquad (7)$$

First obtain the nine basis images $j_{m,i}$, i = 1, 2, . .,9 for each gallery image j_m ,m = 1, 2, . . . , N. Next, for each gallery image j_m, the optimal TSF h_{Tm} and illumination coefficients $\alpha_{m,i}$ are found. To determine the identity of the probe, reblur and re-illuminate each one of the gallery images j_m using the estimated TSF value h_{Tm} and the illumination coefficients namely $\alpha_{m,i}$, compute the LBP features from these transformed gallery images and compare them with those from the probe g to find the closest match.

Algorithm 2:

ISSN: 2349-6495

The input now is blurred and non-uniformly illuminated image

1. Obtain the nine basis images for each gallery image
2. Find the optimal TSF and the illumination coefficients for each gallery image.
3. Transform the gallery images using the computed TSF value and extract the LBP features.
4. The LBP features of the probe image g and those of the transformed gallery images are compared and the closest match is found.

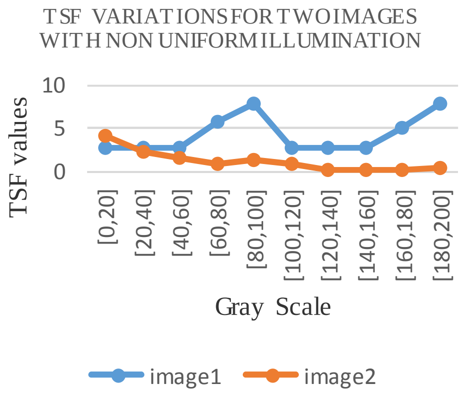


Fig.5: Graph showing the TSF variations for two different images which are non-uniformly illuminated

4.2 Handling Pose Variations

The third algorithm is using an estimate of the pose, matches the incoming probe with a synthesized non-frontal gallery image. The nine illumination basis images are estimated as before using, but now the new synthesized pose is found out.

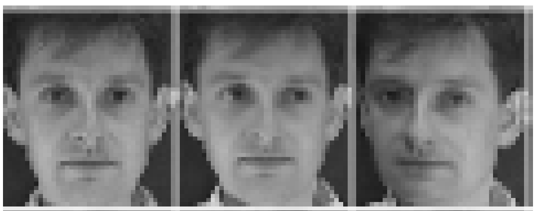


Fig.6: Different poses of a person

Algorithm 3:

The input is blurred, non-uniformly illuminated and differently posed image of a person.

1. After identifying the probe image, the estimate of the pose of the image is found.
2. Synthesize the new pose for each of the gallery image.
3. Obtain the nine basis images for the above image.
4.Obtain the optimal TSF and the illumination co-efficient.
5.Blur and re-illuminate the synthesized gallery image using the above feature and extract the LBP features.
6. Compare the LBP features of the above image and that of the transformed image, get the closet match.

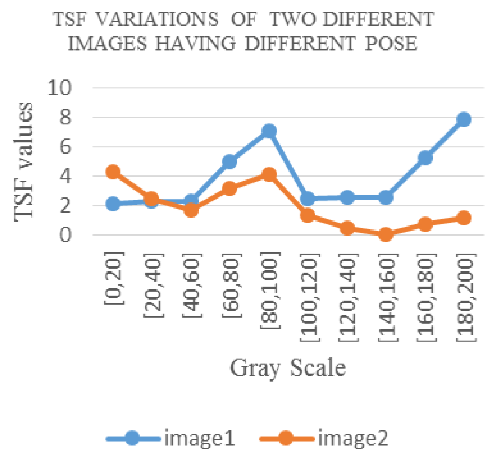
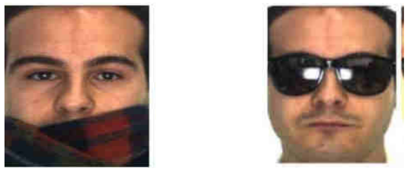


Fig.7: Graph showing TSF variations in two images having different pose

4.3 Partial Occlusions

Occlusion is defined as the blockage or obstruction occurring in the object while it is being photographed. Identifying the object from the occlusions is a challenging task. Here the occlusions are found while taking the image of individuals. It can be the sunglasses or may be due to some other blockages.

In order to tackle with this problem, the common feature of the human face is being utilized. They are the distance between the retina of the eyes, the distance between the nose tip and the upper lip, the distance between the centre of eyes and nose tip.



(a) (b)

Fig.8: Example of occluded image

Algorithm 4:

The input for the algorithm is occluded image and the focused image of the human is already available in the database.

1. Identify the probe image which is occluded
2. Give the image as input for algorithm for separating the face from the surrounding objects.
3. Using optimum TSF value obtain the TSF model
4. Using the vision cascade object detector tool identify the pair of eyes, nose and lips.
5. Measure the Euclidean distance between the centre of the eyes, between eyes and nose and the distance between the tip of the nose and lips.
6. Find the optimal TSF and the illumination coefficients for each gallery image.

7. Transform the gallery images using the computed TSF value and extract the LBP features.
8. Compare these features with the image from the focused gallery, a match is found with the image which shows the closest value for the above parameters.

For handling the occlusion the algorithm first find out the length of the lips, distance between the centre of eyes, length of nose. The method [15] of discriminating the face parts is done using a threshold, in which a certain measure is set for the face parts. Using that metric eyes, nose, lips etc are identified.

V. RESULT ANALYSIS

It is known that not all portion of an image contain the same amount of information. As far as the real data sets are concerned it can be filled with large amount of unwanted units which will not relevant to our matter of concern. Hence the image is first divided into a number of patches and weights are assigned according to the amount of needed information.

In order to make the algorithm in [13] more efficient Local Binary Pattern is utilized, but it is still not sensitive to slight misalignments. In the new approach Transformation Spread Function (TSF) is used. For better efficiency optimum TSF value is used.

The image which is first reduced in size so as to obtain exactly ten TSF values for the image. As mentioned earlier the unconstrained settings in the images can be non-uniform illumination, pose, blur .The comparison between the two methods is shown in the following table. The following values are obtained by evaluating the equation (7) and that with the equation used in [13].

It is clear from the following table that the new algorithm is more sensitive to minor variation (illumination, blur and pose) in the image. When the real datasets are concerned these slight changes are very much relevant.

Table 1 Comparison table of [13] and β added new algorithm for non-uniformly illuminated image

Algorithm without β	β added Algorithm
2.0910	2.0935
2.3460	2.3485
2.3460	2.3485
4.9980	5.0005
7.0890	7.0915
2.5200	2.5225
2.5760	2.5785
2.5760	2.5785
5.2080	5.2105
7.8400	7.8425

Table 2 Comparison table of [13] and β added new algorithm for different pose

Algorithm without β [13]	β added Algorithm
2.6790	2.6815
2.6790	2.6815
2.6790	2.6815
5.8140	5.8165
7.8090	7.8115
2.6790	2.6815
2.6790	2.6815
2.6220	2.6245
4.9590	4.9615
7.8660	7.8685

Table 3 Comparison table of [13] and β added new algorithm for blurred image

Algorithm without β [13]	β added Algorithm
3.0251	3.0254
1.5752	1.5755
0.5549	0.5552
0.7876	0.7878
1.1277	1.1280
1.6224	1.6226
0.8256	0.8258
0.4608	0.4611
0.9696	0.9698
0.6048	0.6050

Along with the TSF value LBP feature is added to increase the robustness of the algorithms. As mentioned earlier the image is first divided into local regions and the LBP descriptors are extracted for each of the regions independently .Further these features are concatenated to form a global descriptor of the face.

VI. CONCLUSION

The methodology mentioned in this paper can be used to identify the individual from unconstrained settings such as motion blur, non-uniform illumination, varying pose as well as the occlusions. The images with occlusions are identified using features of the face which is unique to an individual.

ACKNOWLEDGEMENT

The results put forward in this paper is the simplified version of the algorithm which can handle the problems in images taken. The third author is thankful to the first and second authors for their support and guidance and also to Anna University Regional Campus Coimbatore.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition:A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4,pp. 399–458, Dec. 2003.

[2] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman,"Removing camera shake from a single photograph," *ACM Trans.Graph.*, vol. 25, no. 3, pp. 787–794, Jul. 2006.

[3] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 73:1–73:10,Aug. 2008.

[4] Y.-W. Tai, P. Tan, and M. S. Brown, "Richardson-Lucy deblurring for scenes under a projective motion path," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 33, no. 8, pp. 1603–1618, Aug. 2011.

[5] O.Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186,2012.

[6] A. Gupta, N. Joshi, L. Zitnick, M. Cohen, and B. Curless, "Single image deblurring using motion density functions," in *Proc. Eur. Conf. Comput.Vis.*, 2010, pp. 171–184.

[7] Z. Hu and M.-H. Yang, "Fast non-uniform deblurring using constrained camera pose subspace," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 1–11.

[8] C. Paramanand and A. N. Rajagopalan, "Non-uniform motion deblurring for bilayer scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*,Jun. 2013, pp. 1115–1122.

[9] V. M. Patel, T. Wu, S. Biswas, P. J. Phillips, and R. Chellappa,"Dictionary-based face recognition under variable lighting and pose," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 954–965, Jun. 2012.

[10] Abhijith Punnappurath, Ambasamudram Narayanan Rajagopalan, Sima Taheri , Rama Chellappa, Guna Seetharaman," Face Recognition Across Non-Uniform Motion, Blur, Illumination, and Pose" *IEEE Trans.on Image Processing*, vol. 24, no. 7,, pp 2067-2082 July 2015.

[11] S.S.Ghatge , V.V.Dixit," Face Recognition under varying illumination with Local binary pattern" *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* Vol. 2, Issue 2, February 2013.

[12] Michal ˇ Sorel and Filip ˇ Sroubek," Space-Variant Deblurring using one Blurred and one Underexposed Image" 2009.

[13] P. Vageeswaran, K. Mitra, and R. Chellappa, "Blur and illumination robust face recognition via set-theoretic characterization," *IEEE Trans.Image Process.*, vol. 22, no. 4, pp. 1362–1372, Apr. 2013.

[14] Yen-Yu Lin1,Tyng-Luh Liu, Chiou-Shann Fuh, "Face Detection with Occlusions" Vol.13 No.1 2007.