

Local relation map : a novel illumination invariant face recognition approach

Lian, Zhichao; Er, Meng Joo

2012

Lian, Z., & Er, M. J. (2012). Local relation map : a novel illumination invariant face recognition approach. *International journal of advanced robotic systems*, 9, 128-.

<https://hdl.handle.net/10356/106237>

<https://doi.org/10.5772/51667>

© 2012 Zhichao et al.; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded on 17 Apr 2021 08:25:07 SGT

Local Relation Map: A Novel Illumination Invariant Face Recognition Approach

Regular Paper

Lian Zhichao^{1,*} and Er Meng Joo¹¹ School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

* Corresponding author E-mail: lian0069@e.ntu.edu.sg

Received 31 May 2012; Accepted 17 Jul 2012

DOI: 10.5772/51667

© 2012 Zhichao et al.; licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract In this paper, a novel illumination invariant face recognition approach is proposed. Different from most existing methods, an additive term as noise is considered in the face model under varying illuminations in addition to a multiplicative illumination term. High frequency coefficients of Discrete Cosine Transform (DCT) are discarded to eliminate the effect caused by noise. Based on the local characteristics of the human face, a simple but effective illumination invariant feature local relation map is proposed. Experimental results on the Yale B, Extended Yale B and CMU PIE demonstrate the outperformance and lower computational burden of the proposed method compared to other existing methods. The results also demonstrate the validity of the proposed face model and the assumption on noise.

Keywords Pattern recognition, Face recognition, Illumination variation

1. Introduction

Face recognition has been an active research area over the last 20 years. Numerous approaches have been proposed and they have made considerable progress in controlled environments. However, it is still a difficult task for a face

recognition system to obtain good performance in real-time, especially under variations in illumination, pose, expression, aging, etc. Illumination variation is one of the most challenging issues in face recognition. Differences caused by illumination variations have proven to be much more significant than differences between individuals [1]. To address the issue, lots of approaches have been proposed in recent years which can be classified into three categories: illumination modelling, illumination invariant feature extraction and illumination normalization [4].

Illumination modelling uses the assumption of certain surface reflectance properties, such as Lambertian surface, to construct images under different illumination conditions. Basri and Jacobs in [2] proved that the images of a convex Lambertian object obtained under different illumination conditions can be well approximated by a 9D linear subspace. High computational load and the requirement of several training images limit its applications in practical problems.

Illumination invariant feature extraction methods attempt to extract facial features that are robust against illumination variations. The common representations include edge map, image intensity derivatives and

Gabor-like filtering image [1]. Recently, a local binary pattern (LBP) [3] was proposed as another effective illumination invariant feature and this has gained much attention. The LBP operator is one of the best local texture descriptors. Besides the robustness against pose and expression variations as common texture features, the LBP is also robust to monotonic grey level variations caused by illumination variations. However, none of above mentioned features is robust enough against illumination variations, especially when larger changes of illumination direction exist.

Illumination normalization methods preprocess face images so that the images present stable features under varying illumination conditions. Further recognition will be performed based on the normalized images. Compared to other two categories, the methods of normalization usually take less computational load and require less training samples. The Discrete Cosine Transformation (DCT) method proposed in [4] is one of the most representative normalization methods. The original image is processed by logarithm transformation followed by the DCT. After that, low frequency coefficients are zeroed to eliminate the effects caused by illumination variations, since illumination variations are considered to mainly lie in the low frequency band [4]. Lian and Er [5] presented a similar work employing a Discrete Walsh-Hardmard transform instead of the DCT because of the advantages of easy implementation and high speed. Xie and Lam proposed a local normalization (LN) method in [6] based on the local statistical properties of facial images to deal with illumination variations. They extended the assumption that illumination is related to low frequency bands so that illumination can be considered as a constant (only related to the DC component) in a small local area. Different from most previous work, the LN involves the noise in the illumination model as an additive term besides the multiplicative term illumination. In the LN, the noise in a local area is considered as a constant. In [7], a logarithmic total variation (LTV) model is proposed which decomposes an input image into large-scale output u and small-scale output v . The small-scale output is regarded as an illumination invariant feature which can be used in further recognition tasks. The idea behind the model is similar to the above mentioned DCT method, but the LTV has a better edge-preserving ability. However, the performance of the LTV is still not satisfactory, especially in cases where larger illumination variations exist, as shown in the following experiments. Furthermore, the LTV is an iterative method and it takes a very high computational burden.

Recently, we investigated a novel illumination invariant feature local relation map (LRM), yielding good results that were published in the ISNN 2011 conference [8]. The main

idea of the LRM is based on the local properties of facial images and it has a much lower computational load compared to the similar LN method [6]. Furthermore, to reduce the effects caused by noise, an additive term as noise is considered in the illumination model. Different to previous researchers in [6], we propose applying high frequency DCT coefficients obtained in the entire image to estimate the noise. After that, a logarithm transform is used to change the model into an additive model. With the simplified model, the LRM is extracted as facial features for further recognition tasks. In this paper, we provide a more detailed analysis of the proposed approach, present additional results and discuss further extensions.

The rest of this paper is organized as follows. In Section 2, we introduce our illumination model and novel illumination invariant feature LRM in detail. Furthermore, we demonstrate that the LRM is illumination invariant through theoretical analysis in this section. Experimental results and discussions are presented in Section 3. Finally, conclusions are drawn in Section 4.

2. Illumination Invariant Face Recognition Technique

2.1 Face and Illumination Model

In most existing approaches, the image under different illuminations is simply modelled as

$$f(x, y) = r(x, y) \cdot e(x, y) \quad (1)$$

where $f(x, y)$ is the image grey level, $r(x, y)$ is the reflectance and $e(x, y)$ is the illumination. Based on the model, illumination variations are proposed to mainly lie in the low frequency band [4]. Therefore, low frequency DCT coefficients in the logarithm domain are discarded in [4] to compensate for illumination variations.

In this paper, the face model under varying illuminations is modelled as (2)

$$f(x, y) = r(x, y) \cdot e(x, y) + n(x, y). \quad (2)$$

In this model, an additive term $n(x, y)$ as noise is involved in addition to the reflectance $r(x, y)$ and illumination term $e(x, y)$. In [6], the noise $n(x, y)$ in a local area is modelled as a constant based on the local properties of the human face. In this paper, observing the noise in an entire image, we propose applying the high frequency DCT coefficients obtained in the entire image to estimate the noise. To remove the effects of noise, the DCT is firstly applied in face images. The values of k dimensions of high frequency DCT coefficients are set to zeros in a zigzag mode. Experimental results and more discussions in Section 3 will be presented to demonstrate the validity of the model and our assumption.

After the denoising described in the above section, the model is simplified into

$$f'(x, y) = r(x, y) \cdot e(x, y) \quad (3)$$

Taking the logarithm transform on Eq. (3), we have

$$\log f'(x, y) = \log r(x, y) + \log e(x, y). \quad (4)$$

From the numerical view, the logarithm transform can compress light pixel values and expand the dark ones [1]. Moreover, the logarithmic mapping from an input $f'(x, y)$ to the perceived sensation $\log f'(x, y)$ is justified by Weber's law, proposed by Ernst Weber in 1834. From another point of view, the logarithm transform can convert the product into a sum which makes the normalization task mentioned before easier.

2.2 Local Relation Map

A human face can be treated as a combination of lots of small and flat facets [6]. In such a small facet W , the illumination $e(x, y)$ can be considered as a constant as in [6]. Compared with the model (4), the $\log e(x, y)$ can be taken as a constant A . Therefore, for a special illumination condition, a small facet W can be modelled as

$$I(x, y) = p(x, y) + A, (x, y) \in W \quad (5)$$

where

$$I(x, y) = \log f'(x, y) \text{ and } p(x, y) = \log r(x, y)$$

Based on Eq. (5), we propose a simple illumination invariant feature local relation map, which eliminates the effect of A by comparing the relation between the grey level of target point with those of the points in the boundary of W . The details of the LRM are described as follows:

- 1) Given a point (x, y) , determine the local facet W . In this paper, we mainly focus on a square local facet because it is easier to implement.
- 2) Determine the boundary points U in the facet. For a square facet with size n , there is only $4(n-1)$ boundary points.
- 3) Compare the grey level of point (x, y) with those of points in U as

$$I'(x, y) = I(x, y) - \sum_{(a,b) \in U} I(a, b) / 4(n-1) \quad (6)$$

- 4) After all the points on a given image are processed, the illumination invariant feature, named local relation map, is obtained.

In addition to the square facet employed by this work, a circular or diamond facet is an appropriate choice. Mathematically, any point within the small facet W can be taken as U . In this paper, we only consider boundary points in the facet as U . One of the reasons for this is that it reduces the computational complexity to $O(n)$, compared to $O(n^2)$ when all the points in the square facet are involved in the process. The effect on real computational time will be discussed in the experimental section. Another reason is that boundary points can summarize the local properties of the facet, similar to the idea in local binary patterns. Our ongoing study suggests boundary points can be used in the adaptive size selection of the facet.

2.3 Properties of Local Relation Map

In this section we will demonstrate that the LRM is illumination invariant through theoretical analysis. Given two images I_1 and I_2 of the same person taken under different illumination conditions, for the same point (x, y) , we have

$$I_1(x, y) = p(x, y) + A_1 \quad (7)$$

$$I_2(x, y) = p(x, y) + A_2 \quad (8)$$

After the process of the LRM, we have

$$\begin{aligned} I'_1(x, y) &= I_1(x, y) - \sum_{(a,b) \in U} I_1(a, b) / 4(n-1) \\ &= p(x, y) - \sum_{(a,b) \in U} p(a, b) / 4(n-1) \end{aligned} \quad (9)$$

and

$$\begin{aligned} I'_2(x, y) &= I_2(x, y) - \sum_{(a,b) \in U} I_2(a, b) / 4(n-1) \\ &= p(x, y) - \sum_{(a,b) \in U} p(a, b) / 4(n-1) \end{aligned} \quad (10)$$

By comparing (9) and (10), we have

$$I'_1(x, y) = I'_2(x, y). \quad (11)$$

This means that the LRM is unrelated to illumination conditions. Therefore, we can use the LRM for further face recognition.

3. Experimental Results and Discussions

In this section, we present the evaluation of our algorithm and other existing methods on the Yale Face database B [9], the Extended Yale Face database B [10] and CMU PIE [11]. In [7], the LTV has been proven to achieve the best

performances in the Yale B and CMU PIE among several representative methods. However, the authors in [7] have not provided the comparison between the LTV and the methods proposed in [4, 6]. To make our evaluation convincing, we implement the approaches in [3, 4, 6, 7] besides our proposed method and name them as the LBP, DCT, LN and LTV, respectively, for comparison. In the following experiments, the nearest neighbourhood classifier is used with the Euclidean distance.

3.1 Experiments on CMU PIE Database

There are altogether 68 subjects with pose, illumination and expression variations in the CMU PIE [11]. Because we are concerned with the illumination variation problem, only frontal face images (Pose 27) under different illumination variations (without expression variations) of CMU PIE are used, and there are 1428 images (21 images per person) in total. Only the images with the frontal flash (Flash 08) are chosen as gallery set (one image per subject) and the remaining images are used as probe set (20 images per subject). All the images are manually cropped, aligned and resized as in [4]. To compare the illumination model (2) with noise with the model (1) without noise, we simply apply the LRM in the logarithm domain of the images without the step of denoising, and name the method "LRM without denoising". All the results are shown in Table 1.

From the table it is easy to see that all the methods obtain good performances and our methods obtains the best performance, the same as the LN. Compared to the performances in the second experiment on the Yale B and Extended Yale B, the performances of all the methods in the CMU are better. One of the probable reasons for this is that the CMU database contains limited illumination variations, which are similar to Subsets 1, 2 and 3 in the Yale B and Extended Yale B. Hence, it is easier to deal with illumination variations in the CMU database. Furthermore, there is no difference between the performances of our proposed two methods.

Method	Error rate (%)
DCT	0.37
LTV	1.03
LN	0.22
LRM	0.22
LRM without denoising	0.22

Table 1. Performance comparisons of different methods on CMU

3.2 Experiments on Yale B and Extended Yale B Database

In the experiments we use the Yale Face database B and Extended Yale Face database B as the test database. In the Yale Face database B, there are 10 persons with 64

different illumination conditions for nine poses per person [9]. In the Extended Yale Face database B, there are 16128 images of 28 persons with the same conditions as Yale B [10]. Because the main concern in this paper is illumination variation, only 64 frontal face images per person under different illumination conditions are chosen. After combining the Extended Yale B with the Yale B except 18 corrupted images, there are 2414 images of 38 subjects named as the Completed Yale B. The images are divided into five Subsets based on the angle between the light direction and the camera axis as with the other methods shown in Table 2. All the images are manually cropped, aligned and resized as in [10].

	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5
Light angle	0~12	13~25	26~50	51~77	>77
Number of images	263	456	455	526	714

Table 2. Subsets divided based on light source direction

In the experiments, only one frontal image per person with normal illumination (0°light angle) is applied as the gallery and the other images are used as probe sets. All the results are shown in Table 3. From the table it is clear that the proposed method achieves the best total performance compared with the other methods. The results demonstrate the validity of our assumption that the noise can be modelled based on high frequency components. For small illumination variations, such as Subset 3, the DCT and LRM without denoising (they both only involve the illumination in the face model as Eq. (1)) obtain better performances. For large illumination variations, such as Subset 4 and 5, the LN and LRM (they consider the face model as Eq. (2)) outperform the other two methods. The comparison demonstrates that the noise does not need to be considered when only small illumination variations exist and the noise needs to be modelled as an additive term when larger illumination variations happen.

Method	Error rate (%)			
	Subset 3	Subset 4	Subset 5	total
DCT	10.5	10.8	12.6	8.1
LBP	2.0	22.4	64.4	24.3
LTV	14.1	14.1	14.6	10.0
LN	12.3	6.3	8.4	6.2
LRM	11.2	7.6	7.6	6.0
LRM without denoising	10.5	8.2	10.9	7.0

Table 3. Performance comparisons of different methods on Complete Yale B

3.3 Noise Modelling

As discussed in the last section, noise can be ignored when illumination variation is small and noise needs to be considered more seriously when larger illumination variation exists. In [6], the noise is modelled as a constant in a small neighbourhood. In this paper, the noise is proposed to be related to high frequency components obtained in an entire image. In this section, we will compare these two different assumptions through experiments in Subsets 4 and 5 where larger illumination variations exist.

No matter which assumption is more accurate, the purposes of the above two assumptions are the same, that is to obtain a face model as Eq. (3) after processing. To test these two assumptions, we define a denoising performance criterion (DPC) based on the following equations

$$DPC = \sum_{i=1}^{test\ person\ number} \sum_{j=1}^{different\ lighting\ conditions} \sum_{(x,y) \in I} VR(x,y,i,j) \quad (12)$$

$$VR(x,y,i,j) = \text{var}(PDR(x,y,i,j,a,b)) \quad (a,b) \in W(x,y) \quad (13)$$

$$PDR(x,y,i,j,a,b) = (I_{i,j}(a,b) - I_{i,j}(x,y)) / (I_{i,1}(a,b) - I_{i,1}(x,y)) \quad (14)$$

Where $I_{i,j}$ is the image of the individual i under the lighting condition j (when $j=1$, it is the normal illumination condition, so $I_{i,1}$ is the gallery image of the individual i in the database), $I_{i,j}(x,y)$ is the grey value of point (x,y) in the image $I_{i,j}$, $W(x,y)$ is a small neighbourhood of point (x,y) and $VR(x,y,i,j)$ is the variance of $PDR(x,y,i,j,a,b)$ within W .

Based on (3) and (12), if the illumination $e(x,y)$ in a facet is a constant, we have

$$\begin{aligned} &PDR(x,y,i,j,a,b) \\ &= e_{i,j}(a,b) \cdot (r_{i,j}(a,b) - r_{i,j}(x,y)) / \\ &e_{i,1}(a,b) \cdot (r_{i,1}(a,b) - r_{i,1}(x,y)) \end{aligned} \quad (15)$$

where e is the illumination term unrelated to the person i , and r is the reflectance unrelated to illumination condition j . Therefore, we have

$$\begin{aligned} &PDR(x,y,i,j,a,b) \\ &= e_j(a,b) \cdot (r_i(a,b) - r_i(x,y)) / e_1(a,b) \cdot (r_i(a,b) - r_i(x,y)) \\ &= e_j(a,b) / e_1(a,b) \end{aligned} \quad (16)$$

Because the illumination is considered to be a constant in W , $PDR(x,y,i,j,a,b)$ is given by

$$PDR(x,y,i,j,a,b) = e_j(a,b) / e_1(a,b) \quad \forall (a,b) \in W \quad (17)$$

From (17) we can see that $PDR(x,y,i,j,a,b)$ should be a constant for each (a,b) in a small neighbourhood W near the point (x,y) . Based on the definitions of (12) and (13), the DPC should be zero.

From the definition of the DPC and above analysis, it is easy to find that if the DPC is smaller, the corresponding assumption on noise and the related solution is more convincing. In this section, we calculate the real DPCs of each method in Subsets 4 and 5 of Extended Yale B as shown in Fig.1 and 2.

From the figures, we can see that for the same size neighbourhood, each DPC of the LRM in each Subset is smaller than the corresponding DPC of the LN, except that in Subset 5 the W size equals 10. Therefore, our assumption that the noise should be taken as a high frequency component in the entire image is more convincing.

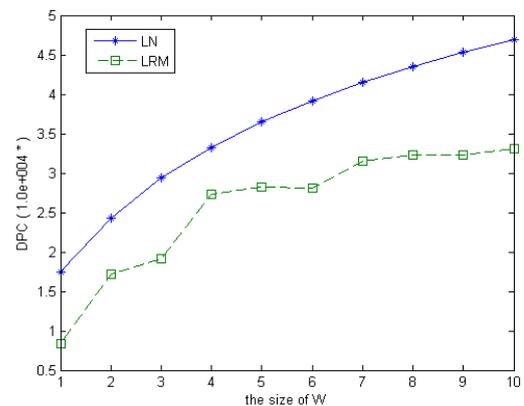


Figure 1. DPCs of different methods in Subset 4

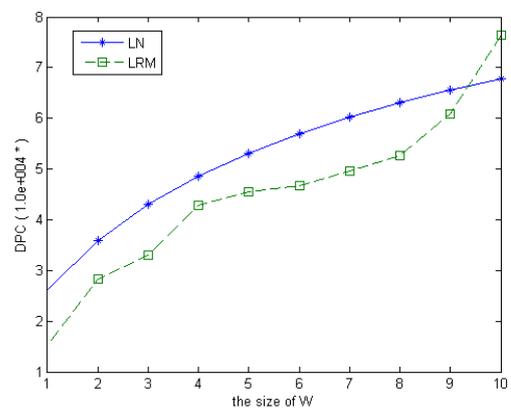


Figure 2. DPCs of different methods in Subset 5

3.4 Computational Complexity

Furthermore, we compare the computational time of the LN and that of our proposed method because they achieve a better total recognition performance and apply similar local properties to that of the human face. Suppose that the image size is $m \times m$ and the size of the local area is $n \times n$. The real computational time is calculated using Matlab in a personal computer with a 2.66GHz CPU. The comparison is shown in Table 4. For comparison, the LTV in [7] takes 17163.9 ms per image in the same condition. From the table we can see that our method significantly reduces computational burden and is 64% faster.

Methods	Computational complexity	Real computational time (per image)
The LN	$O(n^2 m^2)$	2.51s
The LRM	$O((\log m + n)m^2)$	0.91s

Table 4. Comparison of computational complexities

4. Conclusions

In this paper a low computation complexity face recognition approach is proposed to address the problem of illumination variations. Different from most existing methods, an additive term is involved in the face model under varying illuminations as noise. As an assumption, we propose to zero an appropriate number of high frequency DCT coefficients to eliminate the effect caused by the noise. The experimental results demonstrate the validity of the proposed face model and assumption on noise. Based on the local characteristics of the human face, a simple but effective illumination invariant feature LRM is proposed. Theoretical analysis and the experimental results show that the LRM is robust against illumination variations and achieves superior performances on the CMU PIE, Yale B and Extended Yale B. Furthermore, the proposed method shows a good computational efficiency which is important in real time applications. Further research on adaptive size selection of the local area will be carried out in the future.

5. Acknowledgments

The author would like to thank Yale University for the use of the Yale Face Database B and the Extended Yale Face Database B. The author would also like to thank Carnegie Mellon University for the use of the CMU Pose, Illumination and Expression (PIE) database.

6. References

- [1] Adini Y, Moses Y, Ullman S (1997) Face recognition: the problem of compensating for changes in illumination direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19: 721-732.
- [2] Barsi R, Jacobs D.W (2003) Lambertian reflectance and linear subspaces. *IEEE Trans. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25: 218-233.
- [3] Ahonen T, Hadid A, Pietikainen M, (2006) Face description with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28: 2037-2040.
- [4] Chen W, Er M.J, Wu S (2006) Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 36: 458-466.
- [5] Lian Z.C., Er M.J. (2010) Illumination normalisation for face recognition in transformed domain, *Electronics Letters*, 15: 1060-1061.
- [6] Xie X, Lam K.M (2006) An efficient illumination normalization method for face recognition. *Pattern Recognition Letters*, 27: 609-617.
- [7] Chen T, Yin W, Zhou X.S, Comaniciu D, Huang T.S (2006) Total variation models for variable lighting face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28: 1519-1524.
- [8] Lian Z.C., Er M.J., Li J.K. (2011) A Novel Local Illumination Normalization Approach for Face Recognition. *The Eighth International Symposium on Neural Networks (ISNN 2011), Lecture Notes in Computer Science 6676, LNCS (PART 2)*, 350-355.
- [9] Georghiades A.S, Belhumeur P.N, Kriegman D.J (2001) From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23: 643-660.
- [10] Lee K.C, Ho J, Kriegman D (2005) Acquiring Linear Subspaces for Face Recognition under Variable Lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27: 684-698.
- [11] Sim T, Baker S, Bsat M, (2002) The CMU Pose, Illumination, and Expression (PIE) database. *IEEE AFGR 2002*, Washington DC.