MVA2013 IAPR International Conference on Machine Vision Applications, May 20-23, 2013, Kyoto, JAPAN

Choosing Multi-illumination training Images based on the degree of linearindependency

Cong Xia, Jiansheng Chen, Chang Yang, Jing Wang, Jing Liu, Guangda Su, Gang Zhang

Department of Electrical Engineering, Tsinghua University, Beijing, China

Abstract

Multi-illumination training images are usually used in robust face recognition against light variation.Usually a large number of training images are taken under sophisticated imaging system. There definitely exists a large amount of redundancy in these images. Can we pick out several representative samples from them as training images instead of using the whole set? Or even, can we find a better scheme for putting lights when taking or synthesizing these training images? In this paper we propose a method wherewe study the degree of linear independency of face images under different illuminations, and prove that images with different linear independency has different contribution in spanning the illumination subspace.A relatively good combination of images with different linear independency is proposed after analysis and experiments. It also brings forward guidance for light control in obtaining training images.

1. Introduction

Illumination model is widely used to handle the problem of lighting variation in face recognition. The strongest theoretical results so far are due to the Basri and Jacobs [1].The result in the paper is that, for convex Lambertian objects, distant illuminations and fixed pose, all images of the object can be well approximated by linear combinations of nine basis images. But the nine basic images are not real images as some of the pixel values are negative [2], because as specified by the spherical harmonic functions, the nine "harmonic lights" are not real lighting conditions, as for some directions, the intensity is negative. Also, the direct application of this result in most practical systems is misguided for several reasons [5]. Specularities, self-shadowing, and inter-reflections all dramatically affect the appearance of face images, and they all do so in a way that violates the modeling assumptions of the Basri analysis.

Fortunately, even with these effects, for most materials, the relationship between illumination and image is still linear [5] (provided the sensor has a linear response curve), so only positive weights are allowed. As in [7] the space of all images of an object with fixed pose and varying illumination is a convex cone lying in the positive orthant. So, what kind of images does it take to do a good job of representing images sampled from this cone, and when a set of images taken under arbitrary lights are given, how can we choose a most typical subset from them that can represent the illuminations well?

Inmuch previous work, sophisticated imaging systems are designed [3,5,6], and sometimesonly the images illuminated from directions above horizontal are tested in face recognition [2-4]. In [8] the sparse representation and classification (SRC) algorithm achieved impressive results even when the test images may haveilluminations from the back. Such as its experiments on Yale B database, in which each individual has 32 images (selected at random) as training and the other 32 for testing. Obviously there may exists redundancy in the 32 training images. The work is extended in [5], in which 38 training images are taken under lights at evenly spaced spots in a certain angle range, but it doesn't mean every image makes equal contribution for the recognition.

To eliminate redundancy and to find the most representative light directions of training images, we proposed to investigate the linear independency of images under various illuminations. And we research different ways of choosing training images according to their degree of linear independency. Through analysis and experiments, we conclude a guidance advice for effectively picking or obtaining training images.

2. Linear independency of images under various illuminations

Suppose $S = \{I_1, I_2, ..., I_n\}$ is a database of multiple registered training images per subject,taken under varying illuminations. x_i is the vector form of I_i ($i = 1, 2, ..., n$). We let R_i donate the set obtained by deleting I_i from S .and D_i denote the degree of independency of I_i

$$D_i = dist(x_i, R_i) \quad (1)$$

Where the distance function $dist$ between a vector x and a linear subspace R is defined as

$$dist(x, R) = \left\| x - P \bullet \alpha^T \right\|_1 \quad (2)$$

In which $P = [r_1, r_2, ...r_s]$ is the basis of R, which is got from PCA decomposition, and $\alpha = [\alpha_1, \alpha_2, ...\alpha_s]$ is the projection coefficient of x in R_i.

So we can sort the images in S according to D_i, to get their relative degree of linear independency. That will be performed in detail in Section 4.

3. Choosing training image in SRCalgorithm

SRC algorithm has got amazingly high recognition rate in multi-illumination test images [7]. It assumes access to a group of training sets $\{S_1, S_2, ..., S_K\}$of K subjects. The images of subject j ($j = 1, 2, ...K$), stacked as vectors, form a matrix $A_j \in R^{m \times n_j}$. Taken together, all the images form a large matrix $A = [A_1 \mid A_2 \mid ... \mid A_K] \in R^{m \times N}$ ($N = n_1 + n_2 + ... + n_K$), a well aligned test image y_0 can be represented as sparse linear combination Ax_0 of all images in the databases. So when A is huge, the computation is heavy and the algorithm takes a large amount of time.
Suppose for each subject, we are reducing the number of training images from N to T ($T < N$). We sort the n_j images of subject j in ascending order according to their degree of linear independency

$$S_j = \{I_{j1}^{sort}, I_{j2}^{sort}, ...I_{jn_j}^{sort}\} \quad (3)$$

Then we divide them into k groups, each with p images ($k * p = n_j$).

$$\begin{aligned} g_{j1} &= \{I_{j1}^{sort}, I_{j2}^{sort}, ...I_{j,p}^{sort}\} \\ g_{j2} &= \{I_{j,p+1}^{sort}, I_{j,p+2}^{sort}, ...I_{j,2*p}^{sort}\} \\ &... \\ g_{jk} &= \{I_{j,(k-1)p+1}^{sort}, I_{j,(k-1)p+2}^{sort}, ...I_{j,k*p}^{sort}\} \end{aligned} \quad (4)$$

So $S_j = \{g_{j1}, g_{j2}, ..., g_{jk}\}$ .When we pick t_j images from n_j , we distribute the k groups different proportions $R = \{r_1, r_2, ...r_k\}$ ($r_1 + r_2 + ... + r_k = 1$, $r_1, r_2, ..., r_k > 0$).In next chapter we can learn that, by adjusting the distribution of R , the images with different degree of linear independency have different importance.

4. Experiments

We test our method on Yale B face database. Images in Yale B are obtained from 38 individuals, captured under 64 different lighting conditions.

4.1. Linear independency of images

We calculate the degree of independency of the images as in equation (1), andsort the images of each individual in ascending order according to the calculated linear independency. Fig. 1 is an example of the sorting of the whole 64 images of one individual. We can see,theorder is roughly from "good" lights (frontal and uniform) to "bad" lights (with shadows or specularities).

It can be explained that, when all the images are taken from the same person with the same pose, which means the shape and albedo of the surface are fixed, according to the lambertian model [9], the linear relationship is decided by the illuminations of the images. The rays around the frontal direction lie in a small volume in the middle of the illumination cone, they constitute good linear combination for each other, so the reconstruction error is relatively small. While, the collections of images that are produced by extreme lighting conditions (lighting from the sides, up/down, or behind) spread to the lateral area of the illumination, and more extremely, as the last ten images in Figure.1, the sets of pixels illuminated in each image are mutually disjoint. Therefore, they will produce the maximal possible value of $dist(x_i, R_i)$.

Figure 1. 64 illuminations sorted in ascending order according to the calculated degree of linear independency.

4.2. Choosing training images for recognition

After calculating the degree of linear independency of images, we explore the different ways of choosing training images for good performance in recognition.

We also use the Yale B database.And as the same in [8],for each individual, 32 images are randomly picked as testing, and for the other 32 images, the linear independency is calculated, and the images are sorted as in Figure. 1. We divide them equally into 4 groups{ $g_{j1}, g_{j2}, g_{j3}, g_{j4}$ } ($j = 1, 2, ..., 38$) (the group number is randomly picked), each has 8 images in it. Then we pick $T(T < 32)$ images from the 4 groups for training. The various distributions of picking ratios {r1, r2, r3, r4}, which correspond to { $g_{j1}, g_{j2}, g_{j3}, g_{j4}$ }respectively, are tested for recognition. Figure.2 showed 8 kind of distributions, from scheme 1~scheme 7, the barycenter of picking images shifts from g_{j1} to g_{j4} .

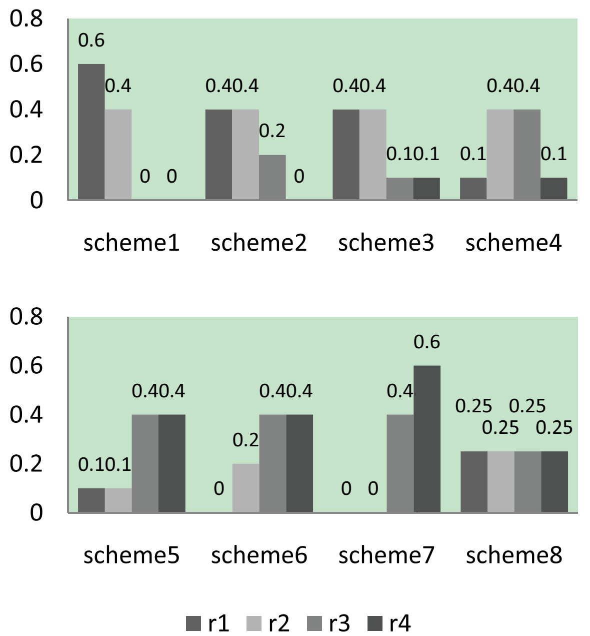

Figure.2. 8distributions of {r1,r2,r3,r4}. From scheme 1~7, the barycenter shift from group1 to group 4, and scheme 8 is an equal distribution.

In recognition test, we let T =8, 12, 16 and 20. And the images are randomly picked in every subgroup { $g_{j1}, g_{j2}, g_{j3}, g_{j4}$ } ($j = 1, 2, ..., 38$). Results of these experiments are shown in Figure. 3.1216 images of 38 individuals are used for testing. Each experiment is repeated for twenty times and the average proportion of rank one is recorded as the recognition rate.For memory reasons, we respectively down-sample all the images to 12*10, 16*12, 20*16 and 32*16, corresponding to T =8, 12, 16 and 20,. We also randomly picked T images from the 32 images as comparison with the 8 distributions.As there are no more than 8 images in each group, the distribution of ratios vary slightly around these schemes when T changes. And the sum of images in two subsections are not enough for T =20, so we only do the experiments on portion 2~6 and portion 8.

From the result we can see that, scheme5~7 have much better results that those of scheme 1~3.Scheme 4 and scheme 8 have similar recognition rates as random picking. The distribution of scheme 5 seems to have the best recognition effect.While what they have in common is that, if the training images are mainly composed of images with small degree of independency, which correspond with the images under “good” lights, the recognition rate is the lowest. That’s the same when the training images are all “bad” lights. The best composition seems to the combination of more “bad” lights and a few “good” lights. Its effect seems better than that of uniform distribution and random picking.

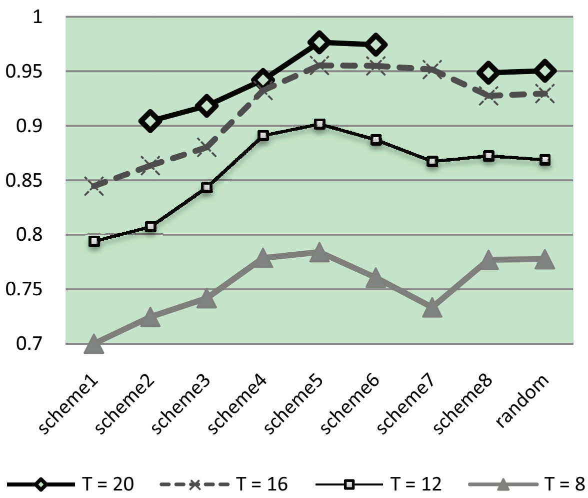

Figure.3. the recognition rates of 8 distributions of images, compared with the randomly picked images.Let the number of picked images Tbe 20, 16, 12, 8, four curves of recognition rates in shown.

It can be explained that, when only rays under nearly frontal directions are selected, they can span only a subspace with a very small volume, it’s a poor approximation of the illumination cone. While the extreme lights (left/right, up/down, or behind) produce a subspace with large volume, they are only the boundary of the illumination cone and haven’t contained the interior of the cone. So the optimal combination is a few directions concentrated in the frontal area and more directions spread to the lateral area. While as the images are picked randomly in each subgroup, we can hardly control the exact directions of images we choose, it’s possible that all the images we choose are under the same light direction (such as lights from left or from right).So when the number of chosen images T is too small (such as $T < 10$), the algorithm turns a little instable (the recognition rates are very different in repeat tests). That also exists in random picking.

We continue the idea of the scheme 5 (more side-lighted images and a few frontal ones) and let T varies from 16 to 32 (for detail, the distribution in four subgroups are: T=16(2,2,6,6); T=20(2,2,8,8); T=24(2,6,8,8); T=28(4,8,8,8); T=32(the whole set)).We compare the results with that of randomly picking T images from the original 32 ones. All the images are down-sampled to 32*16. We also repeat each experiment for twenty times, and in table. 1, we present the maximum and minimum values of recognition rates of every scheme. And in Figure. 4, the average value is shown as curves.

We can see that when T decreases from 32 to 16, the recognition goes down much lower than that of random picking. So if certain error is allowed, it's possible to use smaller scale of training image set with our method. Furthermore, if we have the condition to take images in fixed direction, or generate images from 3D face, we don't have to take or generate a large scale of training images, or let the light direction be equally spaced. We just need a few directions concentrated in the frontal area and the next directions spread quasi-uniformly to the lateral area. According to the discussion in the previous paragraphs, that's the most effective distribution for robust recognition.

Table 1.Maximum and minimum of recognition rates

T	Our method		Random picking	
	max	min	max	min
16	96.10%	94.94%	93.49%	92.66%
20	98.08%	97.33%	96.10%	95.50%
24	98.30%	97.96%	96.74%	96.51%
28	98.75%	98.32%	97.83%	98.00%
32			98.75%	98.01%

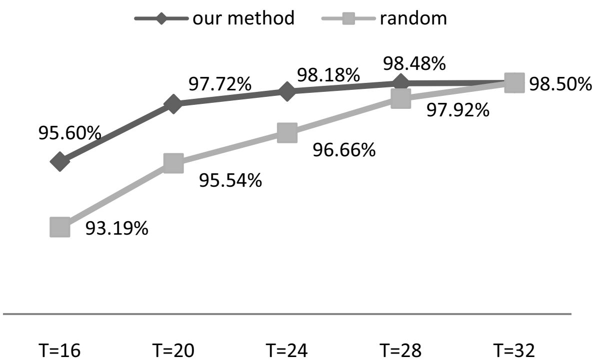

Figure.4. Comparison with randomly picked training set.The size of training set T varies from 16 to 32. The upper curve is the result of our method, and the lower on is that of random picking.

5. Conclusion

In this paper, we proposed a method to learn the linear independency of images under various illuminations, and discussed the relationship between the degree of linear independency and the light condition of images.And we also learned different distributions of images with various degree of linear independency have different recognition effects.

From the bad performance of choosing rays clustered around the direct frontal direction, we are reminded that the traditional methods where only a single frontal gallery image is available per individual, are sensitive for light varying. Multi-illumination gallery images are robust for light changing, while it brings increase in computation. For this situation, we demonstrate that the combination of a few frontal lighted images with more side-lighted ones is an effective way that can get equally good recognition performance with fewer training images. And it's also a valuable advice for taking training image with particular system or generating from 3D face.

While one limitation of our experiments is that, we didn't control thedirection of images when picking them from original training set. That may cause instability in small scale of chosen training set. If we roughly estimate the light direction of each image when picking them, or take/generate training images from fixed directions, we can distribute the ratios of rays' angle sizes, so it's possible that very small scale of training images are enough for good performance in face recognition.

This work is supported by the National Natural Science Foundation of China (grant no. 61101152) and the Chuanxin Foundation from Tsinghua University (grant no. 110107001).

References

[1] R. Basri and D. Jacobs. "Lambertian Reflectance and Linear Subspaces,"IEEE Trans. Pattern Analysis and Machine Intelligence, vol.25, no.2, pp.218-233, 2003

[2] K.C. Lee, J. Ho, D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting",IEEE Trans. Pattern Analysis and Machine Intelligence, vol.34, no.2, Feb 2012.

[3] A. Georghiades, D.Kriegman, and P. Belhumeur."From Few to Many: Generative Models for Recognition under Variable Pose and Illumination," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 40, pp.640-660, 2001

[4] H. Wang, S.Z. Li, Y. Wang, "Face Recognition under Varying Lighting Conditions Using Self Quotient Image". FGR' 04.

[5] A. Wagner, J. Wright, A. etc."Toward a Practical Face Recognition System: Robust Alignment and Illumination by Sparse Representation", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.34, no.2, Feb.2012.

[6] Hideichi Sato, Shree K. Nayar, Katsushi Ikeuchi, "Extracting Shape and Reflectance of Glossy Surfaces by Using 3D photometric Sampling Method" . Proceeding of the IAPR Conference on Machine Vision Applications (IAPR MVA 1990), pp. 133-136.

[7] P. Belhumeur and D. Kriegman, "What Is the Set of Images of an Object under All Possible Illumination Conditions?" Int'l J. Computer Vision, vol. 28, no. 3, pp. 245-260, 1998

[8] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, "Robust Face Recognition via Sparse Representation." IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no.2, pp.210-227, Feb. 2009.

[9] J. Lambert, PhotometriaSive de Mansura etGradibusLuminus, Colurum et Umbrae. EberhardKlett, 1760.