11

Illumination Processing in Face Recognition

Yongping Li, Chao Wang and Xinyu Ao
Shanghai Institute of Applied Physics, Chinese Academy of Sciences
China

1. Introduction

Driven by the demanding of public security, face recognition has emerged as a viable solution and achieved comparable accuracies to fingerprint system under controlled lightning environment. In recent years, with wide installing of camera in open area, the automatic face recognition in watch-list application is facing a serious problem. Under the open environment, lightning changes is unpredictable, and the performance of face recognition degrades seriously.

Illumination processing is a necessary step for face recognition to be useful in the uncontrolled environment. NIST has started a test called FRGC to boost the research in improving the performance under changing illumination. In this chapter, we will focus on the research effort made in this direction and the influence on face recognition caused by illumination.

First of all, we will discuss the quest on the image formation mechanism under various illumination situations, and the corresponding mathematical modelling. The Lambertian lighting model, bilinear illuminating model and some recent model are reviewed. Secondly, under different state of face, like various head pose and different facial expression, how illumination influences the recognition result, where the different pose and illuminating will be examined carefully. Thirdly, the current methods researcher employ to counter the change of illumination to maintain good performance on face recognition are assessed briefly. The processing technique in video and how it will improve face recognition on video, where Wang's (Wang & Li, 2009) work will be discussed to give an example on the related advancement in the fourth part. And finally, the current state-of-art of illumination processing and its future trends will be discussed.

2. The formation of camera imaging and its difference from the human visual system

With the camera invented in 1814 by Joseph N, recording of human face began its new era. Since we do not need to hire a painter to draw our figures, as the nobles did in the middle age. And the machine recorded our image as it is, if the camera is in good condition.

Currently, the imaging system is mostly to be digital format. The central part is CCD (charge-coupled device) or CMOS (complimentary metal-oxide semiconductor). The CCD/CMOS operates just like the human eyes. Both CCD and CMOS image sensors operate in the same manner -- they have to convert light into electrons. One simplified way to think about the sensor used in a digital camera is to think of it as having a 2-D array of thousands or millions of tiny solar cells, each of which transforms the light from one small portion of the image into electrons. The next step is to read the value (accumulated charge) of each cell in the image. In a CCD device, the charge is actually transported across the chip and read at one corner of the array. An analog-to-digital converter turns each pixel's value into a digital value. And the value is mapping to the pixel value in the memory, thus forming the given object image. Although they shared lots of similarity as human eyes， however, the impression is different. One of the advantage of human visual system is the human eye could view color constantly regardless of the luminance value in the surrounding. People with normal visual capabilities could recall the leave of tree is always green either in the morning, at the noon, or in the dust of sunset. Color constancy is subjective constancy, it remains relatively constant under normal variation situation. This phenomena was explained by N. Daw (Conway & Livingstone, 2006) using the Double-opponent cells, later E. land developed retinex theory to explain it (Am. Sci., 1963). However, for the CCD/CMOS, the formed color of the leave of the tree is related to the surrounding luminance value greatly. Thus, the difference between them is the reason that there should be some difference in the face recognition between human and machine. Machine could not take it for granted the appearance has some ignorance of its surrounding luminance value.

Human gets the perception of objects from the radiance reflected by the objects. Usually, the reflection from most objects is scattered reflection. Unlike reflected by smooth surface, the ray is deflected in random directions by irregularities in the propagation medium.

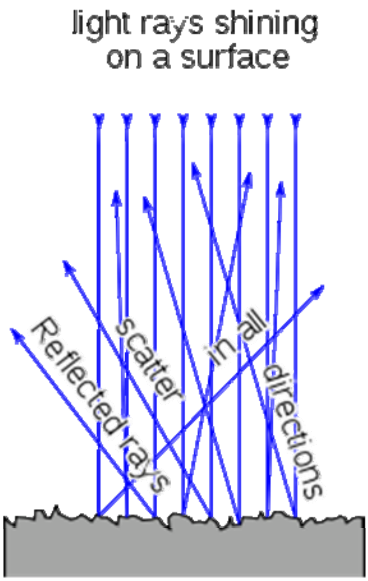


Fig. 1. Diagram of diffuse reflection (taken from the wikipedia.org)

If it is captured by eyes of human being, then perception could be fulfilled. Illumination independent image representation is an important research topic for face recognition. The face images were recorded under tightly controlled condition, where different pose, various distance, and different facial expression were presented. The edge maps, Gabor-filtered images, and the derivative of gray image were tried, but none of them could achieve the goal to be illumination independent, and none of these works provided a good enough framework to overcome the influence of various lighting condition.

3. Models for illumination

To overcome the problem, some mathematical models describing the reflectance of object in computer graphics were utilized to recover the facial image under various lighting condition. Image of a human face is the projection of its three-dimensional head on a plane, the important factors influencing the image representation is the irradiance. In computer graphics, Lambertian surface (Angel, 2003) is used to model the object surface's irradiance. The surface is called Lambertian surface if light falling on it is scattered such that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. It could be modeled mathematically as in the following equation (1), where $I(x, y)$ is the image irradiance, ρ is the surface reflectance of the object, $n(x, y)$ is the surface normal vector of object surface, and s is the incidence ray.

$$I(x, y) = \rho(x, y)n(x, y)^T \cdot s \tag{1}$$

The Lambertian surface luminance could be called to be isotropic technically. Recently, Shashua and Riklin-Raviv (Shashua & Riklin-Raviv, 2001) proposed a method to extract the object's surface reflectance as an illumination invariant description. The method is called quotient image, which is extracted from several sample image of the object. The quotient image is defined as shown in equation 2, using the quotient image, it could recover image under some different lighting condition. It is reported outperformed the PCA. However, it works in very limited situation.

$$Q_y = \frac{I_y}{I_a} = \frac{\rho_y(x,y)n^T s}{\rho_a(x,y)n^T s} = \frac{\rho_y(x,y)}{\rho_a(x,y)} \tag{2}$$

Basri and Jacobs (Basri & Jacobs, 2003) illustrated that the illumination cone of a convex Lambertian surface could be represented by a nine-dimensional linear subspaces. In some limited environment, it could achieve some good performance. Further, Gross et al. (Gross et al., 2002) proposed a similar method called Eigen light-fields. This method claimed to only have one gallery and one probe image to estimate the light-field of the subject head, there is none further requirement on the subject pose and illumination value. And the authors declared that the performance of the proposed method on the CMU PIE database (Sim et al., 2002) is much better than that of other related algorithm.

The assumption of Lambertian model requires perfect situation, E. Nicodemus (Nicodemus, 1965) put forward a theory called BRDF (bidirectional reflectance distribution function) later. The BRDF is a four-dimensional function that defines how light is reflected at an opaque surface. The function takes an incoming light direction ω_i, and outgoing direction ω_o, both defined with respect to the surface normal n, and returns the ratio of reflected radiance exiting along ω_o to the irradiance incident on the surface from direction ω_i, Note that each direction ω is itself parameterized by azimuth angle φ and zenith angle θ, therefore the BRDF as a whole is 4-dimensional. BRDF is used in the field of modelling the reflectance on an opaque surface. These parameters could be illustrated in Fig. 2.

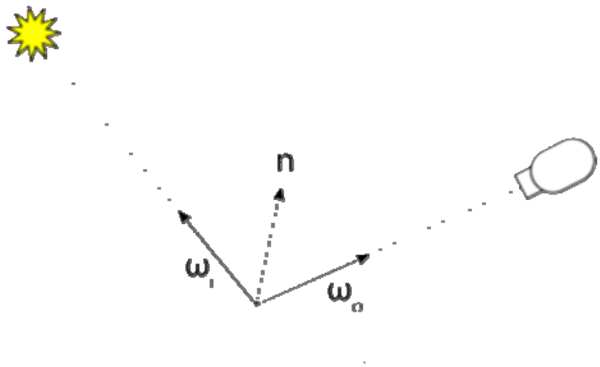


Fig. 2. Diagram showing BRDF, ω_i points toward the light source. ω_o points toward the viewer (camera). n is the surface normal

BRDF model is extensively used in the rendering artificial illuminating effects in computer graphics. To counter the effect of illumination variation, we could artificial render the different lighting situation by using this model. Comparing with Lambertain model, BRDF is of 4 dimensions, the complexity of related computation process is very large. Also, inverting the rendering situation is an ill posed problem, the equation must try some assumptions in serial to solve this problem. Thus, the efforts to employ BRDF model to attack the illumination is not successful currently.

The above models are general approaches to illumination invariant presentation; they have no requirement on the content of the image. However, recently years, there is lots of work towards to make human face image independent of illuminance, and it will be discussed thoroughly in the next section.

4. Current Approaches of Illumination Processing in Face Recognition

Many papers have been published to study on illumination processing in face recognition in the recent years. By now, these approaches can be divided into two categories: passive approaches and active approaches (Zou et al., 2007, a).

4.1 Passive Approaches

The idea of passive approaches: attempt to overcome illumination variation problem from images or video sequences in which face appearance has been altered due to environmental illumination change. Furthermore, this category can be subdivided into three classes at least, described as follows.

4.1.1 Photometric Normalization

Illumination variation can be removed: the input face images can be normalized to some state where comparisons are more reliable.

Mauricio and Roberto (Villegas & Paredes, 2005) divided photometric normalization algorithms into two types: global normalization methods and local normalization methods. The former type includes gamma intensity correction, histogram equalization, histogram matching and normal distribution. The latter includes local histogram equalization, local histogram matching and local normal distribution. Each method was tested on the same face databases: the Yale B (Georghiades et al., 2000) and the extended Yale B (Georghiades et al., 2001) face database. The results showed that local normal distribution achieves the most consistent result. Short. et al. (Short et al., 2004) compared five classic photometric normalization methods: a method based on principal component analysis, multiscale retinex (Rahman et al., 1997), homomorphic filtering, a method using isotropic smoothing to estimate the luminance function and one using anisotropic smoothing (Gross & Brajovic, 2003). The methods were tested extensively across the Yale B, XM2VTS (Messer et al., 1999) and BANCA (Kittler et al., 2000) face databases using numerous protocols. The results showed that the anisotropic method yields the best performance across all three databases. Some of photometric normalization algorithms are illuminated in detail as follows.

4.1.1.1 Histogram Equalization

Histogram equalization (HE) is a classic method. It is commonly used to make an image with a uniform histogram, which is considered to produce an optimal global contrast in the image. However, HE may make an image under uneven illumination turn to be more uneven.

S.M. Pizer and E.P. Amburn (Pizer & Amburn, 1987) proposed adaptive histogram equalization (AHE). It computes the histogram of a local image region centered at a given pixel to determine the mapped value for that pixel; this can achieve a local contrast enhancement. However, the enhancement often leads to noise amplification in “flat” regions, and “ring” artifacts at strong edges. In addition, this technique is computationally intensive.

Xudong Xie and Kin-Man Lam (Xie & Lam, 2005) proposed another local histogram equalization method, which is called block-based histogram equalization (BHE). The face image can be divided into several small blocks according to the positions of eyebrows, eyes, nose and mouth. Each block is processed by HE. In order to avoid the discontinuity between adjacent blocks, they are overlapped by half with each other. BHE is simple so that the computation required of BHE is much lower than that of AHE. The noise produced by BHE is also very little.

4.1.1.2 Gamma Intensity Correction

Shan et al. (Shan et al., 2003) proposed Gamma Intensity Correction (GIC) for illumination normalisation. The gamma transform of an image is a pixel transform by:

$$G(x, y) = I(x, y)^{1/\gamma} \tag{3}$$

where $G(x, y)$ is the output image; $I(x, y)$ is the input image; γ is the Gamma coefficient. With the value γ varying, the output image is darker or brighter. In GIC, the image $G(x, y)$ is transformed as to best match a canonically illuminated image $I_C(x, y)$. To find the best optimal γ, the value should be subject to:

$$\gamma = \arg\min_{\gamma^*} \sum_{x,y} \left[I(x, y)^{1/\gamma^*} - I_C(x, y)\right]^2 \tag{4}$$

4.1.1.3 LogAbout

To solve illumination problem, Liu et al. (Liu et al., 2001) proposed the LogAbout method which is an improved logarithmic transformations as the following equation:

$$g(x, y) = a + \frac{\ln(f(x,y)+1)}{b \ln c} \tag{5}$$

where $g(x, y)$ is the output image; $f(x, y)$ is the input image; a, b and c are parameters which control the location and shape of the logarithmic distribution.

Logarithmic transformations enhance low gray levels and compress the high ones. They are useful for non-uniform illumination distribution and shadowed images. However, they are not effective for high bright images.

4.1.1.4 Sub-Image Homomorphic Filtering

In Sub-Image Homomorphic filtering method (Delac et al., 2006), the original image is split vertically in two halves, generating two sub-images from the original one (see the upper part of Fig. 3). Afterwards, a Homomorphic Filtering is applied in each sub-image and the resultant sub-images are combined to form the whole image. The filtering is subject to the illumination reflectance model as follows:

$$I(x, y) = R(x, y) \cdot L(x, y) \tag{6}$$

where $I(x, y)$ is the intensity of the image; $R(x, y)$ is the reflectance function, which is the intrinsic property of the face; $L(x, y)$ is the luminance function.

Based on the assumption that the illumination varies slowly across different locations of the image and the local reflectance changes quickly across different locations, a high-pass filtering can be performed on the logarithm of the image $I(x, y)$ to reduce the luminance part, which is the low frequency component of the image, and amplify the reflectance part, which corresponds to the high frequency component.

Similarly, the original image can also be divided horizontally (see the lower part of Fig. 3), and the same procedure is applied. But the high pass filter can be different. At last, the two resultant images are grouped together in order to obtain the output image.

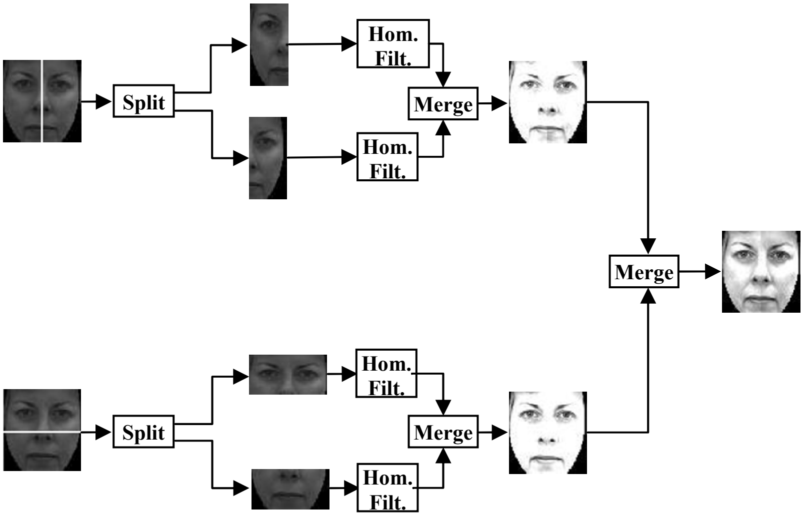


Fig. 3. Sub-image Homomorphic filtering

4.1.2 Illumination Variation Modeling

Some papers attempt to model the variation caused by changes in illumination, so as to generate a template that encompasses all possible environmental changes. The modeling of faces under varying illumination can be based on a statistical model or a physical model. For statistical model, no assumption concerning the surface property is needed. Statistical analysis techniques, such as PCA and LDA, are applied to the training set which contains faces under different illuminations to achieve a subspace which covers the variation of possible illumination. For physical model, the model of the process of image formation is based on the assumption of certain object surface reflectance properties, such as Lambertian reflectance (Basri & Jacobs, 2003). Here we also introduce some classic algorithms on both aspects.

4.1.2.1 Illumination Cone

Belhumeur and Kriegman (Belhumeur & Kriegman, 1998) proposed a property of images called the illumination cone. This cone (a convex polyhedral cone in IRn and with a dimension equal to the number of surface normals) can be used to generate and recognize images with novel illumination conditions.

This illumination cone can be constructed from as few as three images of the surface, each under illumination from an unknown point source. The original concept of the illumination cone is based on two major assumptions: a) the surface of objects has Lambertian reflectance functions; b) the object's surface is convex in shape.

Every object has its own illumination cone, the entirety of which is a set of images of the object under all possible lighting conditions, and each point on the cone is an image with a unique configuration of illumination conditions. The set of n-pixel images of any object seen under all possible lighting conditions is a convex cone in IRn.
Georghiades et al. (Georghiades et al., 1998; Georghiades et al., 2001) have used the illumination cone to further show that, using a small number of training images, the shape and albedo of an object can be reconstructed and that this reconstruction can serve as a model for recognition or generation of novel images in various illuminations. The illumination cone models the complete set of images of an object with Lambertian reflectance under an arbitrary combination of point light sources at infinity. So for a fixed pose, an image can be generated at any position on the cone which is a superposition of the training data (see Fig. 4).

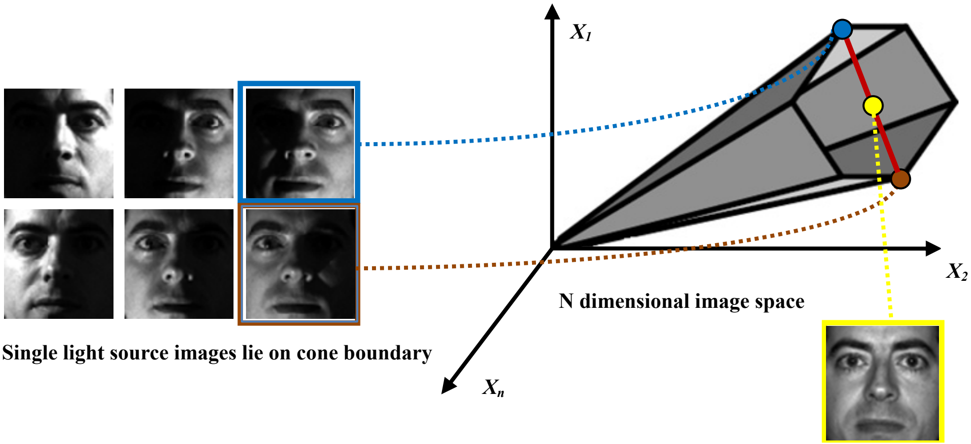


Fig. 4. An example of the generation of novel data from an illumination curve

4.1.2.2 3D Linear Subspace

Belhumeur et al. (Belhumeur et al., 1997) presented 3D linear subspace method for illumination invariant face recognition, which is a variant of the photometric alignment method. In this linear subspace method, three or more images of the same face under different lighting are used to construct a 3D basis for the linear subspace. The recognition proceeds by comparing the distance between the test image and each linear subspace of the faces belonging to each identity.
Batur and Hayes (Batur & Hayes, 2001) proposed a segmented linear subspace model to generalize the 3D linear subspace model so that it is robust to shadows. Each image in the training set is segmented into regions that have similar surface normals by K-Mean clustering, then for each region a linear subspace is estimated. Any estimation only relies on a specific region, so it is not influenced by the regions in shadow.
Due to the complexity of illumination cone, Batur and Hayes (Batur & Hayes, 2004) proposed a segmented linear subspace model to approximate the cone. The segmentation is based on the fact that the success of low dimensional linear subspace approximations of the illumination cone increases if the directions of the surface normals get close to each other.

The face image pixels are clustered according to the angles between their normals and apply the linear subspace approximation to each of these clusters separately. They also presented a way of finding the segmentation by running a simple K-means algorithm on a few training images, without ever requiring to obtain a 3D model for the face.

4.1.2.3 Spherical Harmonics

Ravi Ramamoorthi and Pat Hanrahan (Ramamoorthi & Hanrahan, 2001) presented spherical harmonics method. Basri and Jacobs (Basri & Jacobs, 2003) showed that, a low-dimensional linear subspace can approximate the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions which can be represented by Spherical Harmonics.

Zhang and Samaras (Zhang & Samaras, 2004) combined the strengths of Morphable models to capture the variability of 3D face shape and a spherical harmonic representation for the illumination. The 3D face is reconstructed from one training sample under arbitrary illumination conditions. With the spherical harmonics illumination representation, the illumination coefficients and texture information can be estimated. Furthermore, in another paper (Zhang & Samaras, 2006), 3D shape information is neglected.

4.1.3 Illumination Invariant Features

Many papers attempt to find some face feature which is insensitive to the change in illumination. With the feature, the varying illumination on face cannot influence the recognition result. In other words, we can eliminate the illumination factor from the face image. The best way is to separate the illumination information from the identity information clearly. Here some algorithms are listed as follows.

4.1.3.1 Edge-based Image

Gao and Leung (Gao & Leung, 2002) proposed the line edge map to represent the face image. The edge pixels are grouped into line segments, and a revised Hausdorff Distance is designed to measure the similarity between two line segments. In the HMM-based face recognition algorithms, 2D discrete cosine transform (DCT) is often used for generating feature vectors. For eliminating the varying illumination influence, Suzuki and Shibata (Suzuki & Shibata, 2006) presented a directional edge-based feature called averaged principal-edge distribution (APED) to replace the DCT feature. APED feature is generated from the spatial distributions of the four directional edges (horizontal, $+45^{o}$, vertical, and -45^{o}).

4.1.3.2 Gradient-based Image

Given two images I and J of some plannar Lambertian object taken under the same viewpoint, their gradient-based image ∇I and ∇J must be parallel at every pixel where they are difined. Probabilistically, the distribution of pixel values under varying illumination may be random, but the distribution of image gradients is not.

Chen et al. (Chen et al., Chen) showed that the probability distribution of the image gradient is a function of the surface geometry and reflectance, which are the intrinsic properties of the face. The direction of image gradient is revealed to be insensitive to illumination change. S.

Samsung (Samsung, 2005) presented integral normalized gradient image for face recognition. The gradient is normalized with a smoothed version of input image and then the result is integrated into a new greyscale image. To avoid unwanted smoothing effects on step edge region, anisotropic diffusion method is applied.

4.1.3.3 Wavelet-based Image

Gomez-Moreno et al. (Gomez-Moreno et al., 2001) presented an efficient way to extract the illumination from the images by exploring only the low frequencies into them jointly with the use of the illumination model from the homomorphic filter. The low frequencies where the illumination information exists can be gained by the discrete wavelet transform. In another point of view, Du and Ward (Du & Ward, 2005) performed illumination normalization in the wavelet domain. Histogram equalization is applied to low-low sub-band image of the wavelet decomposition, and simple amplification is performed for each element in the other 3 sub-band images to accentuate high frequency components. Uneven illumination is removed in the reconstructed image obtained by employing inverse wavelet transform on the modified 4 sub-band images.
Gudur and Asari (Gudur & Asari, 2006) proposed a Gabor wavelet based Modular PCA approach for illumination robust face recognition. In this algorithm, the face image is divided into smaller sub-images called modules and a series of Gabor wavelets at different scales and orientations. They are applied on these localized modules for feature extraction. A modified PCA approach is then applied for dimensionality reduction.

4.1.3.4 Quotient Image

Due to the varying illumination on facial appearance, the appearances can be classified into four components: diffuse reflection, specular reflection, attached shadow and cast shadow. Shashua et al. (Shashua & Riklin-Raviv, 2001) proposed quotient image (QI), which is the ratio of albedo between a face image and linear combination of basis images for each pixel. This ratio of albedo is illumination invariant. However, the QI assumes that a facial appearance includes only diffuse reflection. Wang et al. (Wang et al., 2004) proposed self quotient image (SQI) by using only single image. The SQI was obtained by using the Gaussian function as a smoothing kernel function. The SQI however is neither synthesized at the boundary between a diffuse reflection region and a shadow region, nor at the boundary between a diffuse reflection region and a specular reflection region. Determining the reflectance type of an appearance from a single image is an ill-posed problem.
Chen et al. (Chen et al., 2005) proposed total variation based quotient image (TVQI), in which light estimated by solving an optimal problem so-called total variation function. But TVQI requires complex calculation. Zhang et al. (Zhang et al., 2007) presented morphological quotient image (MQI) based on mathematical morphological theory. It uses close operation, which is a kind of morphological approach, for light estimation.

4.1.3.5 Local Binary Pattern

Local Binary Pattern (LBP) (Ojala et al., 2002) is a local feature which characterizes the intensity relationship between a pixel and its neighbors. The face image can be divided into some small facets from which LBP features can be extracted. These features are concatenated into a single feature histogram efficiently representing the face image. LBP is unaffected by any monotonic grayscale transformation in that the pixel intensity order is not changed after such a transformation. For example, Li et al. (Li et al., 2007) used LBP features to compensate for the monotonic transform, which can generate an illumination invariant face representation.

4.1.3.6 3D Morphable Model

The 3D Morphable model is based on a vector space representation of faces. In this vector space, any convex combination of shape and texture vectors of a set of examples describes a realistic human face. The shape and texture parameters of the model can be separated from the illumination information.

Blanz and Vetter (Blanz & Vetter, 2003) proposed a method based on fitting a 3D Morphable model, which can handle illumination and viewpoint variations, but they rely on manually defined landmark points to fit the 3D model to 2D intensity images.

Weyrauch et al. (Weyrauch et al., 2004) used a 3D Morphable model to generate 3D face models from three input images of each person. The 3D models are rendered under varying illumination conditions to build a large set of synthetic images. These images are then used to train a component-based face recognition system.

4.2 Active Approaches

The idea of active approaches: apply active sensing techniques to capture images or video sequences of face modalities which are invariant to environmental illumination.

Here we introduce two main classes as follows.

4.2.1 3D Information

3D face information can be acquired by active sensing devices like 3D laser scanners or stereo vision systems. It constitutes a solid basis for face recognition, which is invariant to illumination change. Illumination is extrinsic to 3D face intrinsic property. Humans are capable to recognize some person in the uncontrolled environment (including the varying illumination), precisely because they learn to deal with these variations in the real 3D world.

3D information can be represented in different ways, such as range image, curvature features, surface mesh, point set, and etc. The range image representation is the most attractive. Hesher et al. (Hesher et al., 2003) proposed range image to represent 3D face information. Range images have the advantage of capturing shape variation irrespective of illumination variability. Because the value on each point represents the depth value which does not depend on illumination.

Many surveys (Kittler et al., 2005; Bowyer et al., 2006; Abate et al., 2007) on 3D face recognition have been published. However, the challenges of 3D face recognition still exist (Kakadiaris et al., 2007): ⑴ 3D capture creates larger data files per subject which implies significant storage requirements and slower processing. The conversion of raw 3D data to efficient meta-data must thus be addressed. ⑵ A field-deployable system must be able to function fully automatically. It is therefore not acceptable to assume user intervention for locating key landmarks in a 3D facial scan. ⑶ Actual 3D capture devices have a number of drawbacks when applied to face recognition, such as artifacts, small depth of field, long acquisition time, multiple types of output, high price, and etc.

4.2.2 Infrared Spectra Information

Infrared (IR) image represents a viable alternative to visible imaging in the search for a robust and practical face recognition system.

According to astronomy division scheme, the infrared portion of the electromagnetic spectrum can be divided into three regions: near-infrared (Near-IR), mid-infrared (Mid-IR) and far-infrared (Far-IR), named for their relation to the visible spectrum. Mid-IR and Far-IR belong to Thermal-IR (see Fig. 5). These divisions are not precise. There is another more detailed division (James, 2009).

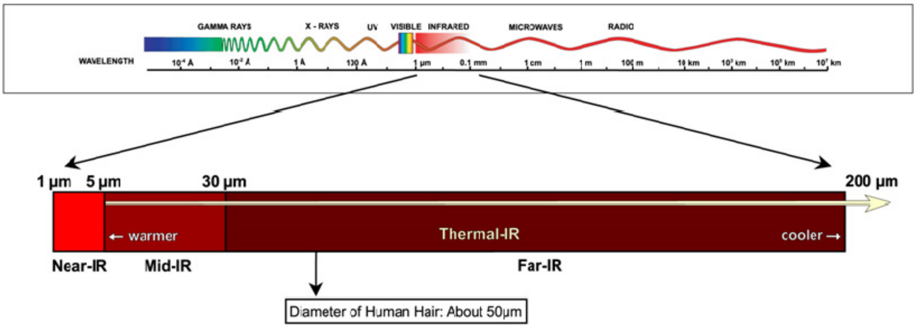


Fig. 5. Infrared as Part of the Electromagnetic Spectrum

Thermal-IR directly relates to the thermal radiation from object, which depends on the temperature of the object and emissivity of the material. For Near-IR, the image intensifiers are sensitive.

4.2.2.1 Thermal-IR

Thermal IR imagery has been suggested as an alternative source of information for detection and recognition of faces. Thermal-IR cameras can sense temperature variations in the face at a distance, and produce thermograms in the form of 2D images. The light in the thermal IR range is emitted rather than reflected. Thermal emissions from skin are an intrinsic property, independent of illumination. Therefore, the face images captured using Thermal-IR sensors will be nearly invariant to changes in ambient illumination (Kong et al., 2005).

Socolinsky and Selinger (Socolinsky & Selinger, 2004, a) presented a comparative study of face recognition performance with visible and thermal infrared imagery, emphasizing the influence of time-lapse between enrollment and probe images. They showed that the performance difference between visible and thermal face recognition in a time-lapse scenario is small. In addition, they affirmed that the fusion of visible and thermal face recognition can perform better than that using either alone. Gyaourova et al. (Gyaourova et al., 2004) proposed a method to fuse the both modalities of face recognition. Thermal face recognition is not perfect enough. For example, it is opaque to glass which can lead to facial occlusion caused by eyeglasses. Their fusion rule is based on the fact that the visible-based recognition is less sensitive to the presence or absence of eyeglasses. Socolinsky and Selinger (Socolinsky & Selinger, 2004, b) presented visible and thermal face recognition results in an operational scenario including both indoor and outdoor settings. For indoor settings under controlled illumination, visible face recognition performs better than that of thermal modality. However, Outdoor recognition performance is worse for both modalities, with a sharper degradation for visible imagery regardless of algorithm. But they showed that fused of both modalities performance outdoors is nearing the levels of indoor visible face recognition, making it an attractive option for human identification in unconstrained environments.

4.2.2.2 Near-IR

Near-IR has advantages over both visible light and Thermal-IR (Zou et al., 2005). Firstly, since it can be reflected by objects, it can serve as active illumination source, in contrast to Thermal-IR. Secondly, it is invisible, making active Near-IR illumination friendly to client. Thirdly, unlike Thermal-IR, Near-IR can easily penetrate glasses.

However, even though we use the Near-IR camera to capture face image, the environmental illumination and Near-IR illumination all exist in the face image. Hizem et al. (Hizem et al., 2006) proposed to maximize the ratio between the active Near-IR and the environmental illumination is to apply synchronized flashing imaging. But in outdoor settings, the Near-IR energy in environmental illumination is strong. Zou et al. (Zou et al., 2005) employed a light emitting diode (LED) to project Near-IR illumination, and then capture two images when the LED on and off respectively. The difference between the two images can be independent of the environment illumination. But when the face is moving, the effect is not good. To solve this problem, Zou et al. (Zou et al., 2007, b) proposed an approach based on motion compensation to remove the motion effect in the difference face images.

Li et al. (Li et al., 2007) presented a novel solution for illumination invariant face recognition based on active Near-IR for indoor, cooperative-user applications. They showed that the Near-IR face images encode intrinsic information of the face, which is subject to a monotonic transform in the gray tone. Then LBP (Ojala et al., 2002) features can be used to compensate for the monotonic transform so as to derive an illumination invariant face representation.

Above active Near-IR face recognition algorithms need that both the enrollment and probe samples are captured under Near-IR conditions. However, it is difficult to realize in some actual applications, such as passport and driver license photos. In addition, due to the distance limitation of Near-IR, many face images can only be captured only under visible lights. Chen et al. (Chen et al., 2009) proposed a novel approach, in which the enrollment samples are visual light images and probe samples are Near-IR images. Based on learning the mappings between images of the both modalities, they synthesis visual light images from Near-IR images effectively.

5. Illumination Processing in Video-based Face Recognition

Video-based face recognition is being increasingly discussed and occasionally deployed, largely as a means for combating terrorism. Unlike face recognition in still, it has its own unique features, such as temporal continuity and dependence between two neighboring frames (Zhou et al., 2003). In addition, it requires high real time in contrast to face recognition in still. Their differences are compared in Table 1.

Face Recognition in Video	Face Recognition in Still
Low resolution faces	High resolution faces
Varying illumination	Even illumination
Varying pose	Frontal pose
Varying expression	Neutral expression
Video sequences	Still image
Continuous motion	Single motion

Table 1. The comparison between face recognition in video and in still

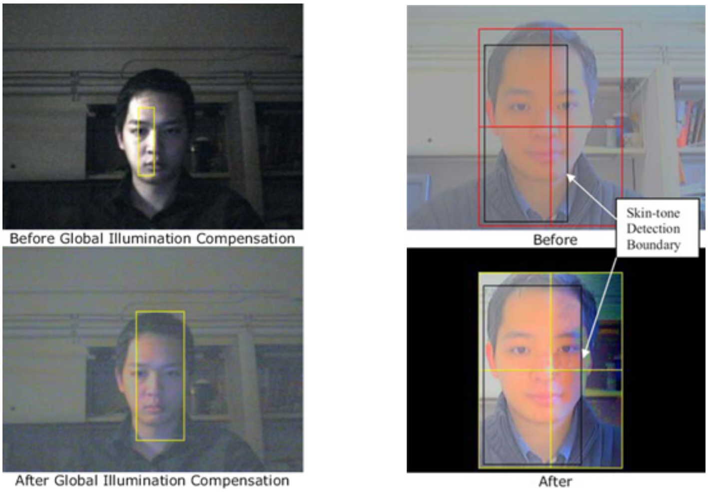
Most existing video-based face recognition systems (Gorodnichy, 2005) are realized in the following scheme: the face is first detected and then tracked over time. Only when a frame satisfying certain criteria (frontal pose, neutral expression and even illumination on face) is acquired, recognition is performed using the technique of face recognition in still. However, maybe the uneven illumination on face always exists, which lead that we cannot find a suitable time to recognize the face.

Using the same algorithms, the recognition result of video-based face recognition is not satisfying like face recognition in still. For example, the video-based face recognition systems were set up in several airports around the United States, including Logan Airport in Boston, Massachusetts; T. F. Green Airport in Providence, Rhode Island; San Francisco International Airport and Fresno Airport in California; and Palm Beach International Airport in Florida. However, the systems have never correctly identified a single face in its database of suspects, let alone resulted in any arrests (Boston Globe, 2002). Some illumination processing algorithms mentioned in Section 3 can be applied for video-based face recognition, but we encounter three main problems at least: (1) Video-based face recognition systems require higher real-time performance. Many illumination processing algorithms can achieve a very high recognition rate, but some of them take much more computational time. 3D face modeling is a classic one. Building a 3D face model is a very difficult and complicated task in the literature even though structure from motion has been studied for several decades. (2) In video sequences, the direction of illumination on face is not single. Due to the face moving or the environmental illumination changing, the illumination on face is in dynamic change. Unlike illumination processing for face recognition in still, the algorithms need more flexible. If the light source direction cannot change suddenly, the illumination condition on face only depend on the face motion. The motion and illumination are correlative. (3) In contrast to general high resolution still image, video sequences often have low resolution (less than 80 pixels between two eyes). For illumination processing, it would be more difficulty. According to the three problems, we introduced some effective algorithms for video-based face recognition.

5.1 Real-time Illumination Processing

Unlike the still image, the video sequences are displayed at a very high frequency (about 10 – 30 frames/second). So it’s important to improve the real-time performance of illumination processing for video-based face recognition.

Chen and Wolf (Chen & Wolf, 2005) proposed a real-time pre-processing system to compensate illumination for face processing by using scene lighting modeling. Their system can be divided into two parts: global illumination compensation and local illumination compensation (see Fig. 6). For global illumination compensation, firstly, the input video image is divided into four areas so as to save the processing power and memory. And then the image histogram is modified to a pre-defined luminance level by a non-linear function. After that, the skin-tone detection is performed to determine the region of interest (ROI) and the lighting update information for the following local illumination compensation. The detection is a watershed between global illumination compensation and local illumination compensation. For local illumination compensation, firstly, the local lighting is estimated within the ROI determined from the previous stage. After obtaining the lighting information, a 3D face model is applied to adjust the luminance of the face candidate. The lighting information is not changed if there is no update request sent from the previous steps.



(a) Global illumination (b) Local illumination

Fig. 6. Global and local illumination compensation

Arandjelović and Cipolla (Arandjelović & Cipolla, 2009) presented a novel and general face recognition framework for efficient matching of individual face video sequences. The framework is based on simple image processing filters that compete with unprocessed greyscale input to yield a single matching score between individuals. It is shown how the discrepancy between illumination conditions between novel input and the training data set can be estimated and used to weigh the contribution of two competing representations. They found that not all the probe video sequences should be processed by the complex algorithms, such as a high-pass (HP) filter and SQI (Wang et al., 2004). If the illumination difference between training and test samples is small, the recognition rate would decrease with HP or SQI in contrast to non-normalization processing. In other words, if the illumination difference is large, normalization processing is the dominant factor and recognition performance is improved. If this notation is adopted, a dramatic performance improvement would be offered to a wide range of filters and different baseline matching algorithms, without sacrificing their online efficiency. Based on that, the goal is to implicitly learn how similar the probe and training samples illumination conditions are, to appropriately emphasize either the raw input guided face comparisons or of its filtered output.

5.2 Illumination change relating to face motion and light source

Due to the motion of faces or light sources, the illumination conditions on faces can vary over time. The single and changeless illumination processing algorithms can be unmeaning. The best way is to design an illumination compensation or normalization for the specific illumination situation. There is an implicit problem in this work: how to estimate the illumination direction. If the accuracy of the illumination estimation is low, the same to the poor face detection, the latter work would be useless. Here we will introduce several illumination estimation schemes as follows.

Huang et al. (Huang et al., 2008) presented a new method to estimate the illumination direction on face from one single image. The basic idea is to compare the reconstruction residuals between the input image and a small set of reference images under different illumination directions. In other words, the illumination orientation is regard as label information for training and recognition. The illumination estimation is to find the nearest illumination condition in the training samples for the probe. The way to estimate illumination of an input image adopted by the authors is to compute residuals for all the possible combinations of illumination conditions and the location of the minimal residual is the expectation of illumination.

Wang and Li (Wang & Li, 2008) proposed an illumination estimation approach based on plane-fit, in which environmental illumination is classified according to the illumination direction. Illumination classification can help to compensate uneven illumination with pertinence. Here the face illumination space is expressed well by nine face illumination images, as this number of images results in the lowest error rate for face recognition (Lee et al., 2005). For more accurate classification, illumination direction map, which abides by Lambert's illumination model, is generated. BHE (Xie & Lam, 2005) can weaken the light contrast in the face image, whereas HE can enhance the contrast. The difference between the face image processed by HE and the same one processed by BHE, which can reflect the light variance efficiently, generates the illumination direction map (see Fig. 7).

In order to make the direction clearer in the map, the Laplace filter and Gaussian low pass filter are also applied. In order to estimate the illumination orientation, a partial least square plane-fit is carried out on the current pixel of the illumination direction map. In actual, $I(x,y)$ is the fitted value. Suppose $f(x,y)$ is the observed value at (x,y). Then the least square between $I(x,y)$ $(I(x,y) = ax + by + c)$ and $f(x,y)$ is shown in Eq. (7).

$$S = \sum_{x,y}[f(x,y) - (ax + by + c)]^2 \tag{7}$$

note: x, y, $f(x,y)$ are known, so S is the function of a, b and c.

The illumination orientation can be defined as the value as follows:

$$\beta = \frac{\partial f(x,y)}{\partial y} \Big/ \frac{\partial f(x,y)}{\partial x}\bigg|_{x=y=0} = \frac{b}{a} = \frac{\sum_{x,y} f(x,y)y}{\sum_{x,y} f(x,y)x} \tag{8}$$

where β denotes the illumination orientation on the illumination direction map.

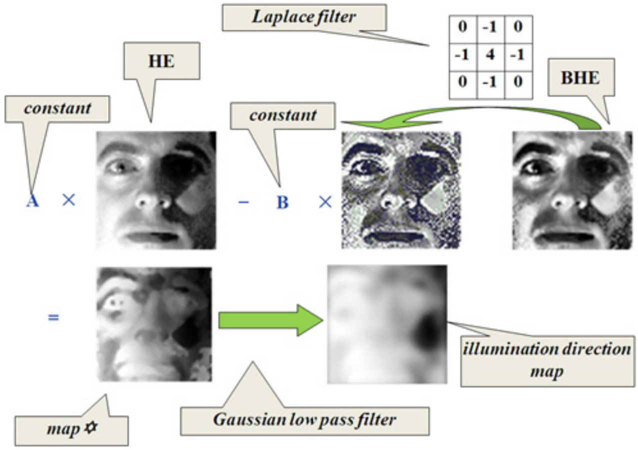


Fig. 7. Generation of illumination direction map

For the same person, the value of β is greatly different with illumination orientation variations; for different persons, the value of β is similar with the same illumination orientation. β can be calculated to make the lighting category determined.
Supposing that the light source direction is fixed, the surface of a moving face cannot change suddenly over a short time period. So the illumination varying on face can be regarded as a continuous motion. The face motion and illumination are correlative.
Basri and Jacobs (Basri & Jacobs, 2003) analytically derived a 9D spherical harmonics based on linear representation of the images produced by a Lambertian object with attached shadows. Their work can be extended from the still image to video sequences, where the video sequences can be only regarded as some separate frames, but it is inefficient. Xu and Roy-Chowdhury (Xu & Roy-Chowdhury, 2005; Xu & Roy-Chowdhury, 2007) presented a theory to characterize the interaction of face motion and illumination in generating video sequences of a 3D face. The authors showed that the set of all Lambertian reflectance functions of a moving face, illuminated by arbitrarily distant light sources, lies "close" to a bilinear subspace consisting of 9 illumination variables and 6 motion variables. The bilinear subspace formulation can be used to simultaneously estimate the motion, illumination and structure from a video sequence. The problem, how to deal with both motion and illumination, can be divided into two stages: □ the face motion is considered, and the change in its position from one time instance to the other is calculated. The change of position can be referenced as the coordinate change of the object. □ the effect of the incident illumination ray, which is projected onto face, and reflected conform to the Lambert's cosine law. For the second stage, incorporating the effect of the motion, Basri and Jacob's work is used.
However, the idea, supposing that the illumination condition is related to the face motion, has a certain limitation. If the environment illumination varies suddenly (such as a flash) or illumination source occultation, the relation between motion and illumination is not credible. All approaches conforming to the supposition would not work.

5.3 Illumination Processing for Low Resolution faces

As a novel input, it is difficult to capture a high resolution face in an arbitrary position of the video. But we can obtain a single high quality video of a person of interest, for the purpose of database enrolment. This problem is of interest in many applications, such as law enforcement. For low resolution faces, it is harder to adopt illumination processing, especially pixel-by-pixel algorithms.

However, it clearly motivates the use of super-resolution techniques in the preprocessing stages of recognition. Super-resolution concerns the problem of reconstructing high-resolution data from a single or multiple low resolution observations. Formally, the process of making a single observation can be written as the following generative model:

$$\mathrm{x} = \downarrow [\mathrm{t}(\hat{\mathrm{x}}) + \mathrm{n}] \tag{9}$$

where $\hat{x}$ is the high-resolution image; $t(\cdot)$ is an appearance transformation (e.g. due to illumination change, in the case of face images); n is additive noise; $\downarrow$ is the downsampling operator.

Arandjelović and Cipolla (Arandjelović & Cipolla, 2006) proposed the Generic Shape-Illumination (gSIM) algorithm. The authors showed how a photometric model of image formation can be combined with a statistical model of generic face appearance variation, learnt offline, to generalize in the presence of extreme illumination changes. gSIM performs face recognition by extracting and matching sequences of faces from unconstrained head motion videos and is robust to changes in illumination, head pose and user motion pattern. For the form of gSIM, a learnt prior is applied. The prior takes on the form of an estimate of the distribution of non-discriminative, generic, appearance changes caused by varying illumination. It means that unnecessary smoothing of person-specific, discriminative information is avoided. In the work, they make a very weak assumption on the process of image formation: the intensity of each pixel is a linear function of the albedo $a(j)$ of the corresponding 3D point:

$$\mathrm{X}(\mathrm{j}) = \mathrm{a}(\mathrm{j}) \cdot \mathrm{s}(\mathrm{j}) \tag{10}$$

where s is a function of illumination parameters , which is not modeled explicitly. Lambertian reflectance model is a special case.

Given two images X_1 and X_2, which are both the same person under the same pose, are of different illuminations.

$$\Delta \log \mathrm{X}(\mathrm{j}) = \log \mathrm{s}_2(\mathrm{j}) - \log \mathrm{s}_1(\mathrm{j}) \equiv \mathrm{d}_\mathrm{s}(\mathrm{j}) \tag{11}$$

So the difference between these logarithm-transformed images is not relative to the face albedo. Under the very general assumption that the mean energy of light incident on the camera is proportional to the face albedo at the corresponding point, d_s is approximately generic i.e. not dependent on the person's identity.

However, this is not the case when dealing with real images, as spatial discretization differently affects the appearance of a face at different scales. In another paper (Arandjelović & Cipolla, 2007) of the authors, they proposed not to explicitly compute super-resolution face images from low resolution input; rather, they formulated the image formation model in such a way that the effects of illumination and spatial discretization are approximately mutually separable. Thus, they showed how the two can be learnt in two stages: (1) a generic illumination model is estimated from a small training corpus of different individuals in varying illumination. (2) a low-resolution artifact model is estimated on a person-specific basis, from an appearance manifold corresponding to a single sequence compounded with synthetically generated samples.

6. Recent State-of-art Methods of Illumination Processing in Face Recognition

How to compensate or normalize the uneven illumination on faces is still a puzzle and hot topic for face recognition researchers. There are about 50 IEEE papers on illumination processing for face recognition within past 12 months. Here we illuminated some excellent papers published on the important conferences (e.g. CVPR and BTAS) or journals (such as IEEE Transactions on Pattern Analysis and Machine Intelligence) since 2008. Many papers, which have been introduced in the former sections, are not restated.

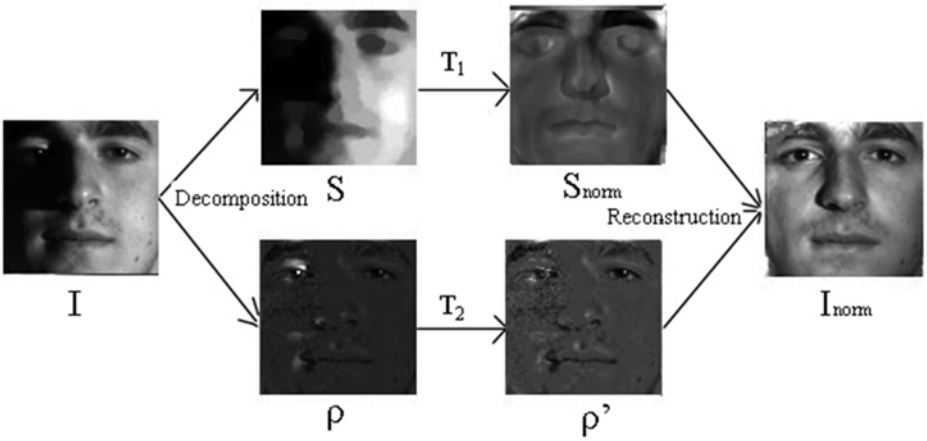


Fig. 8. Illumination Normalization Framework for Large-scale Features

Xie et al. (Xie et al., 2008) proposed a novel illumination normalization approach shown in Fig. 8. In the framework, illumination normalization whereas small-scale features (high frequency component) are only smoothed. Their framework can be divided into 3 stages: (1) Adopt an appropriate algorithm to decompose the face image into 2 parts: large-scale features and small-scale features. Methods in this category include logarithmic total variation (LTV) model (Chen et al., 2006), SQI (Wang et al., 2004) and wavelet transform (Gomez-Moreno et al., 2001) based method. However, some of the methods discard the large-scale features of face images. In this framework, the authors
use LTV. (2) Eliminate the illumination information from the large-scale features by some algorithms, such as HE, BHE (Xie & Lam, 2005) and QI (Shashua & Riklin-Raviv, 2001) etc. In addition, these methods also distort the small-scale features simultaneously during the normalization process. (3) a normalized face image is generated by combination of the normalized large-scale feature image and smoothed small-scale feature image.

Holappa et al. (Holappa et al., 2008) presented an illumination processing chain and optimization method for setting its parameters so that the processing chain explicitly tailors for the specific feature extractor. This is done by stochastic optimization of the processing parameters using a simple probability value derived from intra- and inter-class differences of the extracted features as the cost function. Moreover, due to the general 3D structure of faces, illumination changes tend to cause different effects at different parts of the face image (e.g., strong shadows on either side of the nose, etc.). This is taken into account in the processing chain by making the parameters spatially variant. The processing chain and optimization method can be general, not for any specific face descriptor. To illuminate the chain and optimization method, the authors take LBP (Ojala et al., 2002) for example. LBP descriptor is relatively robust to different illumination conditions but severe changes in lighting still pose a problem. To order to solve this problem, they strive for a processing method that explicitly reduces such intra-class variations that the LBP description is sensitive to. Unlike other slowly processed interactive methods, the authors use only logarithmic transformation of pixel values and convolution of the input image region with small sized filter kernels, which makes the method very fast. The complete preprocessing and feature extraction chain is presented in Fig. 9. For the optimization method, the scheme adopted by the authors is to maximize the probability that the features calculated from an image region, that the filter to be optimized is applied to, are closer to each other in the intra class case than in the extra class case.

Face recognition in uncontrolled illumination experiences significant degradation in performance due to changes in illumination directions and skin colors. The conventional color CCD cameras are not able to distinguish changes of surface color from color shifts caused by varying illumination. However, multispectral imaging in the visible and near infrared spectra can help reduce color variations in the face due to changes in illumination source types and directions. Chang et al. (Chang et al., 2008) introduced the use of multispectral imaging and thermal infrared imaging as alternative means to conventional broadband monochrome or color imaging sensors in order to enhance the performance of face recognition in uncontrolled illumination conditions. Multispectral imaging collects reflectance information at each pixel over contiguous narrow wavelength intervals over a wide spectral range, often in the visible and Near-IR spectra. In multispectral imaging, narrowband images provide spectral signatures unique to facial skin tissue that may not be detected using broadband CCD cameras. Thermal-IR imagery is less sensitive to the variations in face appearance caused by illumination changes. Because the Thermal-IR sensors only measure the heat energy radiation, which is independent of ambient lighting. Fusion techniques have been exploited to improve face recognition performance.

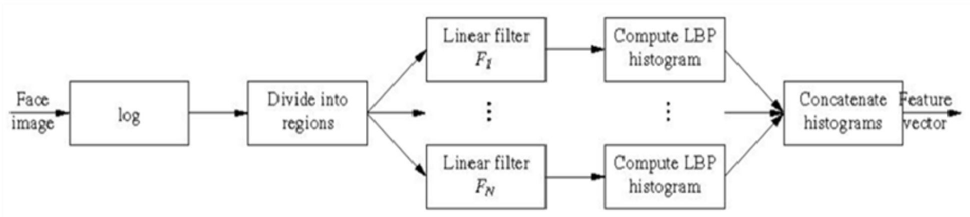
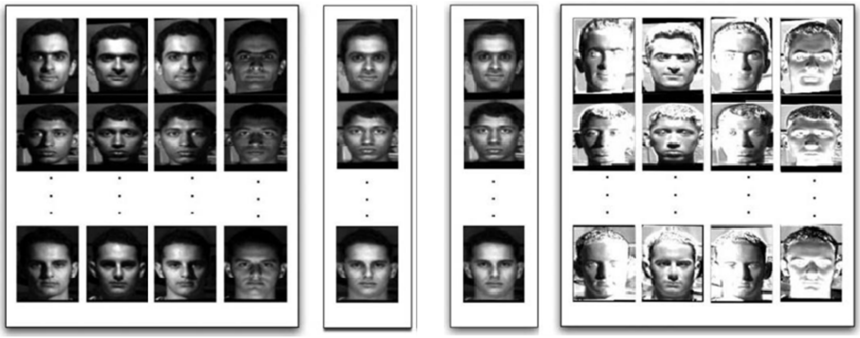


Fig. 9. Illumination Normalization Framework for Large-scale Features

The fusion of Thermal-IR and visible sensors is a popular solution to illumination-invariant face recognition (Kong et al., 2005). However, face recognition based on multispectral image fusion is relatively unexplored. The image based fusion rule can be divided into two kinds: pixel-based and feature-based fusion. The former is easy to implement but more sensitive to registration errors than the latter. Feature based fusion methods are computationally more complex but robust to registration errors.



(a) (b)

Fig. 10. (a) Example de-illumination training data for the Small Faces. Each column represents a source training set for a particular illumination model. In this case: illumination from the right; illumination from the top; illumination from the left; illumination from the bottom. The far right column is the uniformly illuminated target training data from which the derivatives are generated. (b) Example re-illumination training data for the Small Faces. The far left column is the uniformly illuminated source training data. Each remaining column represents the quotient image source training set for a particular illumination model. In this case: illumination from the right; illumination from the top; illumination from the left; illumination from the bottom.

Moore et al. (Moore et al., 2008) proposed a machine learning approach for estimating intrinsic faces and hence de-illuminating and re-illuminating faces directly in the image domain. For estimation of an intrinsic component, the local linear constraints on images are estimated in terms of derivatives using multi-scale patches of the observed images, comprising from a three-level Laplacian Pyramid. The problem of decomposing an observed face image into its intrinsic components (i.e. reflectance and albedo) is formulated as a nonlinear regression problem. For de-illuminating faces (see Fig. 10(a)), with the non-linear regression, the derivatives of the face image are estimated from a given class as it would appear with a uniform illumination. The uniformly illuminated image can then be reconstructed from these derivatives. So the de-illumination step can be regarded as an estimation problem. For re-illuminating faces (see Fig. 10(b)), it is just like an adverse stage of de-illuminating faces. The goal has changed from calculating the de-illuminated face to calculating new illuminations and the input images are de-illuminated faces. Besides these differences, the illumination estimation involves the same basic steps of estimating derivative values and integrating them to form re-illuminated images.

Most public face databases lack images with a component of rear (more than 90 degrees from frontal) illumination, either for training or testing. Wagner et al. (Wagner et al., 2009) made an experiment (see Fig. 11) which showed that training faces with the rear illumination can help to improve the face recognition. The experiment is that the girl should be identified among 20 subjects, by computing the sparse representation (Wright et al., 2009) of her input face with respect to the entire training set. The absolute sum of the coefficients associated with each subject is plotted on the right. The figure also show the faces reconstructed with each subject's training images weighted by the associated sparse coefficients. The red line corresponds to her true identity, subject 12. For the upper row of the figure, the input face is well-aligned (the white box) but only 24 frontal illuminations are used in the training for recognition. For the lower row of the figure, informative representation is obtained by using both well-aligned input face and sufficient (all 38) illuminations in the training. A conclusion can be drawn that illuminations from behind the face are also needed to sufficiently interpolate the illumination of a typical indoor (or outdoor) environment in the training. If not have, the representation will not necessarily be sparse or informative.

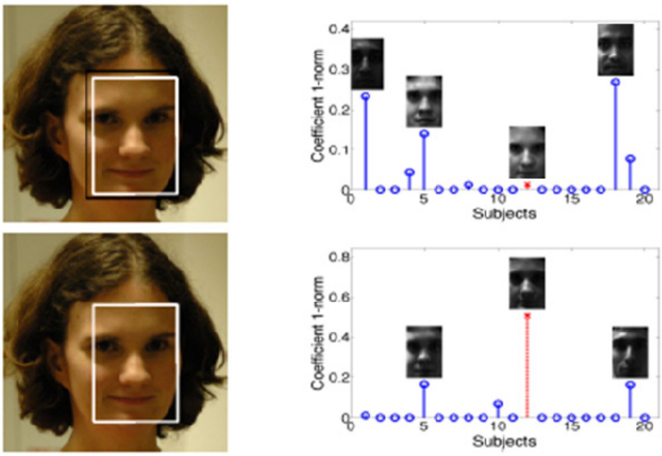


Fig. 11. Recognition Performance with and without rear illumination on faces for training

In order to solve the problem, the authors designed a training acquisition system that can illuminate the face from all directions above horizontal. The illumination system consists of four projectors that display various bright patterns onto the three white walls in the corner of a dark room. The light reflects off of the walls and illuminates the user's head indirectly. After taking the frontal illuminations, the chair is rotated by 180 degrees and then pictures are taken from the opposite direction. Having two cameras speeds the process since only the chair needs to be moved in between frontal and rear illuminations. The experiment results are satisfying. However, it is impossible to obtain
samples of all target persons using the training acquisition system, such as law enforcement for terrorists.

Wang et al. (Wang et al., 2008) proposed a new method to modify the appearance of a face image by manipulating the illumination condition, even though the face geometry and albedo information is unknown. Besides, the known information is only a single face image under arbitrary illumination condition, which makes face relighting more difficult.

(a)

(b)

(c)

Fig. 12. (a) Examples of Yale B face database. (b) SHBMM method. (c) MRF-based method

According to the illumination condition on face, the authors divided their methods into two parts: face relighting under slightly uneven illumination and face relighting under extreme illumination. For the former one, they integrate spherical harmonics (Zhang & Samaras, 2004; Zhang & Samaras, 2006) into the morphable model (Blanz & Vetter, 2003; Weyrauch et al., 2004) framework by proposing a 3D spherical harmonic basis morphable model (SHBMM), which modulates the texture component with the spherical harmonic bases. So any face under arbitrary unknown lighting and pose can be simply represented by three low-dimensional vectors, i.e., shape parameters, spherical harmonic basis parameters, and illumination coefficients, which are called the SHBMM parameters. As shown in Fig. 12 (b), SHBMM can perform well for face image under slightly uneven illumination. However, the performance decreases significantly in extreme illumination. The approximation error can be large, thus making it difficult to recover albedo information. This is because the representation power of SHBMM model is inherently limited by the coupling of texture and illumination bases. In order to solve this problem, the authors presented the other sub-method – a subregion-based framework, which uses a Markov random field (MRF) to model the statistical distribution and spatial coherence of face texture. So it can be called MRF-based framework. Due to MRF, an energy minimization framework was proposed to jointly recover the lighting, the geometry (including the surface normal), and the albedo of the target face (see Fig. 12 (c)).

Gradient-based image has been proved to insensitive to illumination. Based on that, Zhang et al. (Zhang et al., 2009) proposed an illumination insensitive feature called Gradientfaces for face recognition. Gradientfaces is derived from the image gradient domain such that it can discover underlying inherent structure of face images since the gradient domain explicitly considers the relationships between neighboring pixel points. Therefore, Gradientfaces has more discriminating power than the illumination insensitive measure extracted from the pixel domain.
Given an arbitrary image $I(x, y)$ under variable illumination conditions, the ratio of y-gradient of $I(x, y)$ $(\frac{\partial I(x,y)}{\partial y})$ to $I(x, y)$ $(\frac{\partial I(x,y)}{\partial x})$ is an illumination insensitive measure. Then Gradientfaces (G) of image I can be defined as

$$G = \arctan\left(\frac{I_{y-\mathrm{gradient}}}{I_{x-\mathrm{gradient}}}\right),\ G \in [0, 2\pi). \quad (12)$$

where $I_{x-\mathrm{gradient}}$ and $I_{y-\mathrm{gradient}}$ are the gradient of image I in the x, y direction, spectively.

7. Acknowledgement

This research work is partially funded by the National High Technology Research and Development Program (863) of China under Grant # 2008AA01Z124.

8. References

Abate, A.F.; Nappi, M.; Riccio, D. & Sabatino, G. (2007). 2D and 3D Face Recognition: A Survey. Pattern Recognition Letters, 28(14), pp. 1885-1906.
American Science, (1963). The Retinex, 3, pp. 53-60, Based on William Proctor Prize address, Cleveland, Ohio.
Arandjelović, O. & Cipolla, R. (2006). Face Recognition from Video Using the Generic Shape-Illumination Manifold, In: Proc. European Conference on Computer Vision (ECCV), 4, pp. 27–40.
Arandjelović, O. & Cipolla, R. (2007). A Manifold Approach to Face Recognition from Low Quality Video across Illumination and Pose using Implicit Super-Resolution, ICCV.
Arandjelović, O. & Cipolla, R. (2009). A Methodology for Rapid Illumination-invariant Face Recognition Using Image Processing Filters, Computer Vision and Image Understanding, 113(2), pp.159-171.
Basri, R. & Jacobs, D.W. (2003). Lambertian Reflectance and Linear Subspaces, IEEE Trans. PAMI, 25(2), pp. 218-233.
Batur, A.U. & Hayes, M.H. (2001). Linear Subspaces for Illumination Robust Face Recognition, CVPR.
Batur, A.U. & Hayes, M.H. (2004). Segmented Linear Subspaces for Illumination-Robust Face Recognition, International Journal of Computer Vision, 57(1), pp. 49-66.
Belhumeur, P.N. & Kriegman, D.J. (1998). What is the Set of Images of an Object under All Possible Illumination Conditions? Int. J. Computer Vision, 28(3), pp. 245–260.

Belhumeur, P.N.; Hespanha, J.P. & Kriegman, D.J. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pp. 711-720.

Blanz, V. & Vetter, T. (2003). Face Recognition Based on Fitting a 3D Morphable Model, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(9), pp. 1063-1074.

Boston Globe, (2002). Face Recognition Fails in Boston Airport.

Bowyer, K.W.; Chang, K. & Flynn, P. (2006). A Survey of Approaches and Challenges in 3D and Multi-Modal 3D+2D Face Recognition, In CVIU.

Chang, H.; Koschan, A.; Abidi, M.; Kong, S.G. & Won, C.H. (2008). Multispectral Visible and Infrared Imaging for Face Recognition, CVPRW, pp.1-6.

Chen, C.Y. & Wolf, W. (2005). Real-time Illumination Compensation for Face Processing in Video Surveillance, AVSS, pp.293-298.

Chen, H.F.; Belhumeur, P.N. & Jacobs, D.W. (2000). In Search of Illumination Invariants, CVPR.

Chen, J.; Yi, D.; Yang, J; Zhao, G.Y.; Li, S.Z. & Pietikainen, M. (2009). Learning Mappings for Face Synthesis from Near Infrared to Visual Light Images, IEEE Computer Society Conference, pp. 156 – 163.

Chen, T.; Yin, W.T.; Zhou, X.S.; Comaniciu, D. & Huang, T.S. (2005). Illumination Normalization for Face Recognition and Uneven Background Correction Using Total Variation Based Image Models, CVPR, 2, pp.532-539.

Chen, T.; Yin, W.T.; 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(9), pp. 1519-1524.

Conway B.R. & Livingstone M.S. (2006). Spatial and Temporal Properties of Cone Signals in Alert Macaque Primary Visual Cortex (V1), Journal of Neuroscience.

Delac, K.; Grgic, M. & Kos, T. (2006). Sub-Image Homomorphic Filtering Techniques for Improving Facial Identification under Difficult Illumination Conditions. International Conference on System, Signals and Image Processing, Budapest.

Du, S. & Ward, R. (2005). Wavelet-Based Illumination Normalization for Face Recognition, Proc. of IEEE International Conference on Image Processing, 2, pp. 954-957.

Edward, A. (2003). Interactive Computer Graphics: A Top-Down Approach Using OpenGL (Third Ed.), Addison-Wesley.

Fred, N. (1965). Directional reflectance and emissivity of an opaque surface, Applied Optics, 4 (7), pp. 767–775.

Gao, Y. & Leung, M. (2002). Face Recognition Using Line Edge Map, IEEE Trans. on PAMI.

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 Trans. Pattern Anal. Mach. Intelligence, 23(6), pp. 643-660.

Georghiades, A.S.; Belhumeur, P.N. & Kriegman, D.J. (2000). From Few to Many: Generative Models for Recognition under Variable Pose and Illumination, In IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 277-284.

Georghiades, A.S.; Kriegman, D.J. & Belhumeur, P.N. (1998). Illumination Cones for Recognition under Variable Lighting: Faces, CVPR, pp.52.

Gomez-Moreno, H.; Gómez-moreno, H.; Maldonado-Bascon, S.; Lopez-Ferreras, F. & Acevedo-Rodriguez, F.J. (2001). Extracting Illumination from Images by Using the Wavelet Transform. International Conference on Image Processing, 2, pp. 265-268.

Gorodnichy, D.O. (2005). Video-based Framework for Face Recognition in Video, Second Workshop on Face Processing in Video, pp. 330-338, Victoria, BC, Canada.
Gross, R. & Brajovic, V. (2003). An Image Preprocessing Algorithm, AVBPA, pp. 10-18.
Gross, R.; Matthews, I. & Baker, S. (2002). Eigen Light-Fields and Face Recognition Across Pose, International Conference on Automatic Face and Gesture Recognition.
Gudur, N. & Asari, V. (2006). Gabor Wavelet based Modular PCA Approach for Expression and Illumination Invariant Face Recognition, AIPR.
Gyaourova, A.; Bebis, G. & Pavlidis, I. (2004). Fusion of Infrared and Visible Images for Face Recognition, European Conference on Computer Vision, 4, pp.456-468.
Hesher, C.; Srivastava, A. & Erlebacher, G. (2003). A Novel Technique for Face Recognition Using Range Imaging. Seventh International Symposium on Signal Processing and Its Applications, pp. 201- 204.
Hizem, W.; Krichen, E.; Ni, Y.; Dorizzi, B. & Garcia-Salicetti, S. (2006). Specific Sensors for Face Recognition, International Conference on Biometrics, 3832, pp. 47-54.
Holappa, J.; Ahonen, T. & Pietikainen, M. (2008). An Optimized Illumination Normalization Method for Face Recognition, 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1-6.
Huang, X.Y.; Wang, X.W.; Gao, J.Z. & Yang, R.G. (2008). Estimating Pose and Illumination Direction for Frontal Face Synthesis, CVPRW, pp.1-6.
James, B. (2009). Unexploded Ordnance Detection and Mitigation, Springer, pp. 21-22, ISBN 9781402092527.
Kakadiaris, I.A.; Passalis, G.; Toderici, G.; Murtuza, M.N.; Lu, Y.; Karampatziakis, N. & Theoharis, T. (2007). Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4), pp. 640-649.
Kittler, J.; Hilton, A.; Hamouz, M. & Illingworth, J. (2005). 3D Assisted Face Recognition: A Survey of 3D Imaging, Modelling and Recognition Approaches. CVPRW.
Kittler, J.; Li, Y.P. & Matas, J. (2000). Face Verification Using Client Specific Fisher Faces, The Statistics of Directions, Shapes and Images, pp. 63-66.
Kong, S.G.; Heo, J.; Abidi, B.R.; Paik, J. & Abidi, M.A. (2005). Recent Advances in Visual and Infrared Face Recognition: A Review, Computer Vision and Image Understanding, 97(1), pp. 103-135.
Lee, K.; Jeffrey, H. & Kriegman, D. (2005). Acquiring Linear Subspaces for Face Recognition under Variable Lighting, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(5), pp. 1–15.
Li, S.Z.; Chu, R.F.; Liao, S.C. & Zhang, L. (2007). Illumination Invariant Face Recognition Using Near-Infrared Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4), pp. 627-639.
Liu, H.; Gao, W.; Miao, J.; Zhao, D.; Deng, G. & Li, J. (2001). Illumination Compensation and Feedback of Illumination Feature in Face Detection, Proc. IEEE International Conferences on Info-tech and Info-net, 23, pp. 444-449.
Messer, M.; Matas, J. & Kittler, J. (1999). XM2VTSDB: The Extended M2VTS Database, AVBPA.
Moore, B.; Tappen, M. & Foroosh, H. (2008). Learning Face Appearance under Different Lighting Conditions, 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1-8.

Ojala, T.; Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), pp. 971-987.

Pizer, S.M. & Amburn, E.P. (1987). Adaptive Histogram Equalization and Its Variations. Comput. Vision Graphics, Image Process, 39, pp. 355-368.

Rahman, Z.; Woodell, G. & Jobson, D. (1997). A Comparison of the Multiscale Retinex with Other Image Enhancement Techniques, Proceedings of the IS&T 50th Anniversary Conference.

Ramamoorthi, R. & Hanrahan, P. (2001). On the Relationship between Radiance and Irradiance: Determining the Illumination from Images of a Convex Lambertian Object, Journal of Optical Society of American, 18(10), pp. 2448-2459.

Samsung, S. (2005). Integral Normalized Gradient Image a Novel Illumination Insensitive Representation, CVPRW, pp.166.

Shan, S.; Gao, W.; Cao, B. & Zhao, D. (2003). Illumination Normalization for Robust Face Recognition against Varying Lighting Conditions, Proc. of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 17, pp. 157-164.

Shashua, A. & Riklin-Raviv, T. (2001). The Quotient Image: Class Based Re-rendering and Recognition with Varying Illuminations, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 23(2), pp. 129-139.

Short, J.; Kittler, J. & Messer, K. (2004). A Comparison of Photometric Normalisation Algorithms for Face Verification, Proc. Automatic Face and Gesture Recognition, pp. 254-259.

Sim, T.; Baker, S. & Bsat, M. (2002). The CMU Pose, Illumination, and Expression (PIE) Database, Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition.

Socolinsky, D.A. & Selinger, A. (2004), a. Thermal Face Recognition over Time, ICPR, 4, pp.187-190.

Socolinsky, D.A. & Selinger, A. (2004), b. Thermal Face Recognition in an Operational Scenario. CVPR, 2, pp.1012-1019.

Suzuki, Y. & Shibata, T. (2006). Illumination-invariant Face Identification Using Edge-based Feature Vectors in Pseudo-2D Hidden Markov Models, In Proc. EUSIPCO, Florence, Italy.

Villegas, M. & Paredes, R. (2005). Comparison of Illumination Normalization Methods for Face Recognition, In COST275, pp. 27–30, University of Hertfordshire, UK.

Wagner, A.; Wright, J.; Ganesh, A.; Zhou, Z.H. & Ma, Y. (2009). Towards A Practical Face Recognition System: Robust Registration and Illumination by Sparse Representation, IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 597-604..

Wang, C & Li, Y.P. (2008). An Efficient Illumination Compensation based on Plane-Fit for Face Recognition, The 10th International Conference on Control, Automation, Robotics and Vision, pp. 939-943.

Wang, C. & Li, Y.P. (2009). Combine Image Quality Fusion and Illumination Compensation for Video-based Face Recognition, Neurocomputing (To appear).

Wang, H.; Li, S.Z. & Wang, Y.S. (2004). Generalized Quotient Image, CVPR, 2, pp.498-505.

Wang, Y.; Zhang, L.; Liu, Z.C.; Hua, G.; Wen, Z.; Zhang, Z.Y. & Samaras, D. (2008). Face Re-Lighting from a Single Image under Arbitrary Unknown Lighting Conditions, IEEE Transactions on Pattern Analysis and Machine Intelligence

Weyrauch, B.; Heisele, B.; Huang, J. & Blanz, V. (2004). Component-based Face Recognition with 3D Morphable Models. CVPRW.

Wright, J.; Ganesh, A.; Yang, A. & Ma, Y. (2009). Robust Face Recognition via Sparse Representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(2), pp. 210-227.

Xie, X. & Lam, K.M. (2005). Face Recognition under Varying Illumination based on a 2D Face Shape Model, Pattern Recognition, 38(2), pp. 221-230.

Xie, X.H.; Zheng, W.S.; Lai, J.H. & Yuen, P.C. (2008). Face Illumination Normalization on Large and Small Scale Features. CVPR, pp.1-8.

Xu, Y.L. & Roy-Chowdhury, A.K. (2005). Integrating the Effects of Motion, Illumination and Structure in Video Sequences, ICCV, 2, pp.1675-1682.

Xu, Y.L. & Roy-Chowdhury, A.K. (2007). Integrating Motion, Illumination, and Structure in Video Sequences with Applications in Illumination-Invariant Tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(5), pp. 793-806.

Zhang, L. & Samaras, D. (2004). Pose Invariant Face Recognition under Arbitrary Unknown Lighting using Spherical Harmonics, In Proceedings, Biometric Authentication Workshop (in conjunction with ECCV 2004).

Zhang, L. & Samaras, D. (2006). Face Recognition from a Single Training Image under Arbitrary Unknown Lighting Using Spherical Harmonics, IEEE Trans. PAMI, 29.

Zhang, T.P.; Tang, Y.Y.; Fang, B.; Shang, Z.W. & Liu, X.Y. (2009). Face Recognition under Varying Illumination Using Gradientfaces, IEEE Transactions on Image Processing.

Zhang, Y.Y.; Tian, J.; He, X.G. & Yang, X. (2007). MQI based Face Recognition under Uneven Illumination, ICB, pp. 290-298.

Zhou, S.H.; Krueger, V. & Chellappa, R. (2003). Probabilistic Recognition of Human Faces from Video, Computer Vision and Image Understanding, 91(1-2), pp. 214-245.

Zou, X.; Kittler, J. & Messer, K. (2005). Face Recognition Using Active Near-IR Illumination, Proc. British Machine Vision Conference.

Zou, X.; Kittler, J. & Messer, K. (2007), a. Illumination Invariant Face Recognition: A Survey, First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1-8.

Zou, X.; Kittler, J. & Messer, K. (2007), b. Motion Compensation for Face Recognition Based on Active Differential Imaging, ICB, pp. 39-48.