We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

From Canonical Face to Synthesis – An Illumination Invariant Face Recognition Approach

Tele Tan
Curtin University of Technology
Australia

1. Introduction

The need to further develop robust face recognition techniques to meet real world requirements is still an open research challenge. It is widely understood that the two main contributions of poor recognition performances are that caused by variations in face pose and lighting. We will deal with the problem of illumination in this chapter. Approaches addressing the illumination-related problems can be broadly classified into two categories; feature-based approach and exemplar- or appearance- based approach. Feature-based approaches aim to define a feature space that exhibits some broad invariance over the lighting variations. Examples of these are (Adini & Ullman, 1997), (Belhumeur et al., 1997) and (Yang et al., 2004) which uses different image representations like 2D Gabor-like filters, first and second derivatives of the image, and the logarithmic transformation. Although these features may exhibit intensity immunity, none of these are found to be reliable to cope with significantly large variations in illumination changes (Manjunath et al.1992) (Yang et al., 2004).

Exemplar- or appearance- based approaches use a set of sample images taken of a class object (in this case a face) as a basis to compute an intermediary image. The intermediate image can then be used either directly as the probe image or be used to synthesize novel views of the face under different lighting conditions (Mariani, 2002). For example, (Riklin-Raviv & Shashua, 2001) reported a method to compute the Quotient Image from a small sample of bootstrap images representing a minimum of two class objects. The illumination invariant signature of the Quotient Image enables an analytic generation of the novel image space with varying illumination. However, this technique is highly dependent on the types of bootstrap images used which has the undesirable effect of generating diversely looking Quotient Images even from the same person. (Sim & Kanade, 2001) used a statistical shape-from-shading model to estimate the 3D face shape from a single image. The 3D recovery model is based on the symmetric shape-from-shading algorithm proposed by (Zhao & Chellappa, 1999). They used the 3D face model to synthesize novel faces under new illumination conditions using computer graphics algorithms. The approach produce high recognition rate on the illumination subset of the CMU PIE database (Sim et al., 2003). However, it was not evident how their synthesis technique can cope with extreme illumination conditions (Sim & Kanade, 2001). (Debevec et al., 2000) presented a method to

acquire the reflectance field of a human face and use these measurements to render the face under arbitrary changes in lighting and viewpoint. However, the need to generate a large sample of images using the light stage is unfeasible for face recognition purposes. A parameter-free method of estimating the bi-directional reflectance distribution of a subject's skin was proposed by (Smith et al., 2004). They estimated the radiance function by exploiting differential geometry and making use of the Gauss map from the surface onto a unit sphere. They demonstrated the approach by applying it to the re-rendering of faces with different skin reflectance models.

As in (Riklin-Raviv & Shashua, 2001) and (Mariani, 2002), we address the problem of class-based image synthesis and recognition with varying illumination conditions. We define an ideal class as a collection of 3D objects that have approximately the same shape but different albedo functions. For recognition purposes, we can broadly assume all human faces to belong to a certain class structure. This assumption was similarly adopted by (Riklin-Raviv & Shashua, 2001) and (Mariani, 2002). Our approach is based on the dual recovery of the canonical face model and lighting models given a set of images taken with varying lighting conditions and from a minimum of two distinct subjects within the class. The canonical image is equivalent to the reflectance field of the face that is invariant to illumination. The lighting model is the image representation of the ambient lighting independent of the face input. We will first formulate the problem with an over-determined set of equations and propose a method in solving them over every pixel location in the image. We will demonstrate the quality of the recovered canonical face for generating novel appearances using both subjective and objective measures.

2. The Illumination Model

The intensity of reflected light at a point on a surface is the integral over the hemisphere above the surface of a light function L times a reflectance function R. The pixel equation at point (x,y,z) can be expressed as

$$I(x,y,z) = \iiiint\limits_{t\ \lambda\ \theta\ \phi} L(t,x,y,z,\theta,\phi,\lambda)R(t,\theta,\phi,\lambda)\,\mathrm{d}\theta\,\mathrm{d}\phi\,\mathrm{d}\lambda\,\mathrm{d}t \tag{1}$$

where

x,y,z = the co-ordinate of the point on the surface

ϕ and θ = azimuth and yaw angle from the z axis respectively

t and λ = time and wavelength of the light source

This equation is computationally too complex to solve in many real-time applications. We need to make further simplification of the equation without significantly affecting the goal of our work. Firstly, z, t and λ can be eliminated because we are dealing with the projected intensity value of a 3D point onto a single frame grey scale digital image. Additionally, if one considers fixing the relative location of the camera and the light source, θ and ϕ both become constants and the reflectance function collapses to point (x, y) in the image plane. This is a valid condition since we assume the camera to be positioned directly in front of the human subject at all times. Therefore, the first-order approximation of equation (1) for a digital image $I(x,y)$ can be further written as:

$$I(x,y) \approx R(x,y)\,L(x,y) \tag{2}$$

where $R(x,y)$ is the reflectance and $L(x,y)$ is the illumination at each image sample point (x,y). Our approach is to use exemplar images taken over different fixed lighting directions to recover both the reflectance model and illumination source at the same time. It is not the scope of this work to accurately model the skin reflectance property according to specificity like the melanin content of the skin, skin haemoglobin concentration and level of perspirations. These are important for visually accurate skin rendering application but less so for face recognition purposes.

3. Computing the Canonical Face and the Illumination Models

In our case, only the measured intensity images are available. Therefore, there are twice as many unknown data (RHS) as there are known data (LHS) making equation (2) ill-posed. The reflectance surface essentially comprises the combination of the reflectance property associated with the pigmentation of the skin, mouth, eyes and artifacts like facial hair. We define the reflectance model as the canonical face and represent it as a grey level intensity image. We will formulate the dual recovery technique for the canonical faces and illumination models given a set of intensity images $I_{ij}(x,y) \approx R_j(x,y) \, L_i(x,y)$, where i and j are indices to the collection of bootstrap[1] faces taken from M distinct persons ($j = 1, ..., M$) and N different lighting directions ($i = 1, ..., N$).

3.1 Defining and Solving the Systems of Equations

As explained in the previous section, equation (2) has more unknown terms than known. In order to make the equation solvable in a least square sense, we need to introduce additional measurements thus making the system of equations ov*er determined*. We further note that the bootstrap image, $I_{ij}(x,y)$ has two variable components. They are the reflectance component which is unique to the individual person and the illumination model which is dependent on the lighting source and direction. Suppose we have M distinct persons which we use in the bootstrap collection (i.e. R_j, $j = 1, …, M$) and N spatially distributed illumination sources whose direction with respect to the person is fixed at all instances (i.e. L_i, $i = 1, …, N$), we will have therefore a total of MxN known terms and M+N unknown terms. These *over-determined* systems of equations can be solved by selecting any values of M and N that are greater than 1. For example, if we use M persons from the bootstrap collection, and collect N images for each person by varying the illumination, we get the following system of equations;

$$I_{i1}(x,y) \approx R_1(x,y) \, L_i(x,y)$$
$$\vdots$$
$$I_{iM}(x,y) \approx R_M(x,y) \, L_i(x,y)$$

(3)

where $i = 1,…,N$. The terms on the left hand side of these equations are the bootstrap images from the M number of persons. If the illuminations used to generate these bootstrap images are the same, the illumination models, L_i will be common for every person as is reflected in equation (3).

---

[1] The bootstrap collection comprises of face sample images taken of various person over multiple illumination directions, the relative location of which are fixed.

Numerous non-linear minimization algorithms exist and are usually problem dependent (Yeredor, 2000). We chose to use the Levenberg-Marquardt non-linear least square fitting algorithm (More, 1978) as it is fast and suited to problems of high dimensionality. The solver takes as input the set of equations shown in (3) to minimize, a Jacobian matrix of derivatives, a set of known data (i.e. I_{ij}) and seed values for the unknowns. We chose to set the seed value to 128 since there are 256 possible grey values for both the reflectance and illumination models. The internal functions of the solver are iterated until the change in computed values falls below a threshold. At this point the algorithm is said to have converged, and the current computed values for the unknown data are taken as the solution. The algorithm is extremely fast and can recover the unknown values (for most practical values of M and N) in near real-time.

Figure 1 shows the schematic block diagram of the canonical face and illumination model recovery process. The input of the system are the M (M > 1) distinct individuals with each of them taken under N (N > 1) illumination conditions. The outputs are the canonical faces of the M individuals and N illumination models.
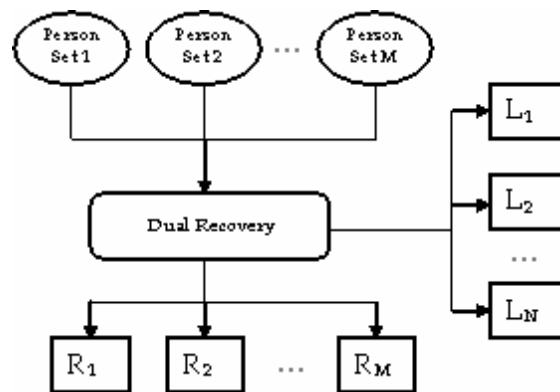

Figure 1. The schematic block diagram showing the canonical face and illumination model recovery process. M is the number of distinct persons and N is the total number of lightings

3.2 Appearance Synthesis
The recovery of the canonical and illumination models is the important first step to enable the following appearance synthesis functions to be performed:
1. New illumination models can be generated by combining the subset of the recovered illumination models. This is possible since mixing lights is an additive function and therefore the new illumination model is simply an addition of the component lights. We can therefore potentially generate significantly greater than N possible variations of the illumination conditions to reflect more accurately the actual real-world lighting conditions.
2. Novel appearance views for each person can be generated by combining the expanded set of illumination models mentioned in point (1) to closely match the actual illumination conditions. (Mariani, 2002) synthesizes the appearance of the face from a single source image by varying both the pose and illumination conditions and reported good recognition rate by matching a probe face with these appearances.

It is not economical and computationally feasible to store specific illumination models for specific faces. To make this approach viable, we need to define a set of generic illumination models that is applicable to people of different skin types and bone structures. We compute each generic illumination model as such;

1. Select a well represented genre of people and recover the canonical face representation for these people using the approach explained in the previous section.
2. Recover the corresponding illumination model for each canonical face. The illumination model should be different for different individual.
3. Estimate the covariance matrix of the intensity measurement between the sample individuals.
4. Diagonalise the covariance matrix using the Singular Value Decomposition (SVD) algorithm to recover the Eigen values and vectors.
5. Use the coefficient of the normalized Eigen vector associated with the highest Eigen value as weights to sum the illumination contribution for each sample individuals. We call this final model the generic illumination model, Lg.

We will subsequently use Lg_i to represent the illumination models in this work. The process of creating the synthesize face images is shown in Figure 2.
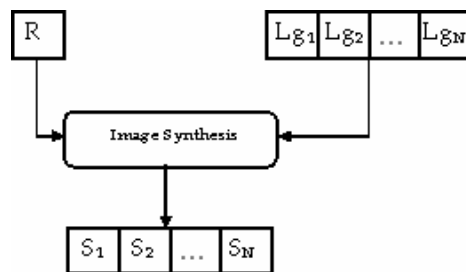

Figure 2. The face synthesis process where S_i ($i = 1, ..., N$) are the synthesized images

4. Experiments

4.1 The Database

In our experiments, we make use of the illumination subset of the CMU PIE database (Sim et al., 2003). The original database comprises 41,368 images taken from 68 people taken under 13 different face poses and 43 different lighting conditions. Each subject were taken with 4 different facial expressions. The images are taken under two ambient conditions; one with the ambient lightings turned on, and the other with the ambient lightings turned off. All the color images are first transformed to grey-level, pre-processed and the faces cropped. The final size for all images is 110 x 90 pixels. Data sets that were used in this experiment are divided into 11 sets of different number of lighting conditions. Lights are selected so that they are as evenly distributed as possible. Table 1 and Table 2 show the individual ID and lighting groupings used in the experiments.

04000	04007	04017	04026	04042	04048
04001	04008	04018	04034	04043	04050
04002	04012	04019	04035	04045	04053
04004	04014	04020	04041	04047	

Table 1. Individual ID (as in PIE Database) use in the experiment

4.2 Canonical Face Recovery

We use equation (3) to recover the canonical faces with different values of M and N and a subset of them are shown in Figure 3. In order to measure the quality of the recovered canonical face, we first define a set of measures that describes the properties of an acceptable canonical face. These measures are; (1) Elimination of lighting effects like specular reflections and casted shadows. (2) Preservation of the visual distinctiveness of the underlying face. (3) Well-balanced intensity distribution. Based on these measures, we can see that in general the recovery of the canonical faces for different values of M and N are very good. In fact the quality is largely independent on the number of bootstrap images (i.e. N) used in the estimation. This is a significant improvement over the Quotient Image reported in (Riklin-Raviv and Shashua, 2001). To illustrate the ability of the technique to extract the canonical image, Figure 4 shows the 5 bootstrap images used to generate the canonical face of subject 04002 as highlighted in red in Fig 3. It is interesting to note that the shadow and highlight regions as seen in these bootstrap images have been significantly reduced, if not eliminated in the canonical face image. The slight reflection on the nose region of subject 04002 may be attributed to oily skin deposits.

Set	# Lights	Flash Positions
1	2	f04, f15
2	3	f01, f11, f16
3	5	f05, f06, f11, f12, f14
4	7	f04, f05, f06, f11, f12, f14, f15
5	9	f04, f05, f06, f11, f12, f14, f15, f19, f21
6	11	f04, f05, f06, f08, f11, f12, f14, f15, f19, f20, f21
7	13	f04, f05, f06, f07, f08, f09, f11, f12, f14, f15, f19, f20, f21
8	15	All except f11, f18, f19, f20, f21, f22
9	17	All except f18, f19, f21, f22
10	19	All except f18 and f22
11	21	All

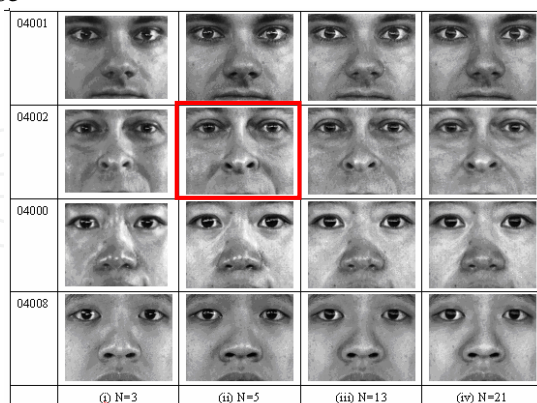Table 2. Groupings of 11 lighting sets and their associated flash positions as determined in the CMU PIE database

Figure 3. Canonical faces generated for candidate samples 04001, 04002, 04000 and 04008 using (i) N=3 (Set 2), (ii) N=5 (Set 3), (iii) N=13 (Set 7) and (iv) N=21 (Set 11)

Figure 4. The bootstrap images used to generate the canonical face for candidate 04002 for N=5

To further support the significance of the recovered canonical face, we will next describe a face recognition experiment that will quantitatively show the ability of our approach to deal with illumination variation problem.

4.3 Face Appearance Synthesis

For each recovered canonical face, the corresponding set of 21 illumination models can then be computed. We further estimated the generic illumination models as defined in Section 3.2 by using 10 candidate samples from the CMU PIE database. We then use these generic illumination models and the canonical faces from the remaining samples to generate novel appearance faces. Figure 5a shows the synthesized views of a subject generated using 7 different illumination models. The corresponding images captured by the actual illuminations are shown in Figure 5b. As can be seen, the appearances of the synthetic images broadly resemble the actual images in relation to the location of the shadow and highlight structures. However, this alone cannot be used as justification for the synthesis approach. One way to measure the significance of the appearance synthesis is by using quantitative face recognition performance measures. This will be elaborated in the next section.
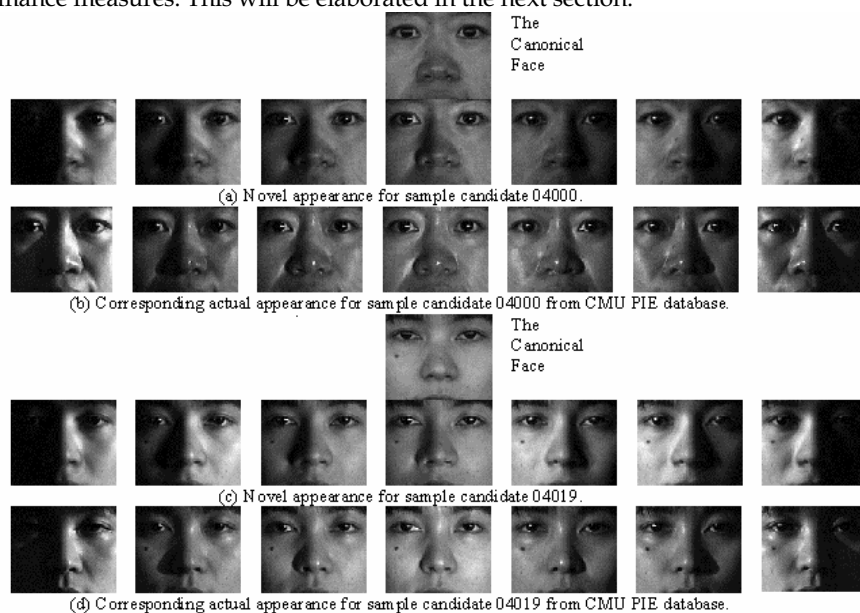

Figure 5. Novel appearance synthesis results using a subset of the generic illumination models and its comparison with the actual appearance. The canonical faces used to generate these images are shown at the top of (a) and (c). The corresponding images of (b) and (d) show the actual illuminations taken from the CMU PIE database.

4.4 Face Recognition

To demonstrate the feasibility of the face appearance synthesis for recognition, we implement a simple classifier based on template matching. This is equivalent to the nearest neighbor classifier reported by (Sim & Kanade, 2001). We use only frontal pose faces throughout the experiment. The generic illumination models used here is the same as in Section 4.3. To maintain unbiased recognition outcome, the test samples used for recognition does not come from any of the samples used to produce the generic illumination models. There are 21 persons in the test samples. From each person we compute the canonical representation and use it to synthesize 21 appearances of the person under different lighting conditions. These images collectively form the registry representation of the person in the database. We use actual illumination samples of the PIE database as the test images. There are a total of 441 (i.e. 21x21) test sample images. We construct different registry databases for different combination of M (number of person) and N (number of lighting) values. We then perform the face recognition experiments on the test samples over the different registries. Figure 6 shows the summary of recognition rate for different values of M and N. We observe several important behaviors. They are:

1. For a fixed value of M, the recognition rate increases monotonically when N increases.
2. However when M increases, N has to consequentially increase for the canonical face to be recovered with reasonable quality. The minimum (M,N) pair needed to establish good recognition rates are (2,3), (4,5), (6,7), (8,9) and (10,11).
3. The recognition rate for N=2 is very poor for all values of M.
4. The range of recognition rates for different values of M and N (ex N=2) are between 85.5% and 88.8%.

As can be seen, the results obtained here is significantly better than (Sim & Kanade, 2001) which reported an accuracy of 39% with the nearest neighbor classifier on a similar dataset. The general trend of the recognition rates which flatten off as N increases for all values of M suggest a wide perimeter for the choices of these values. However, from the computation, data acquisition and hardware standpoint, it would be effective to keep the M and N values small, without negatively impacting the recognition rate.
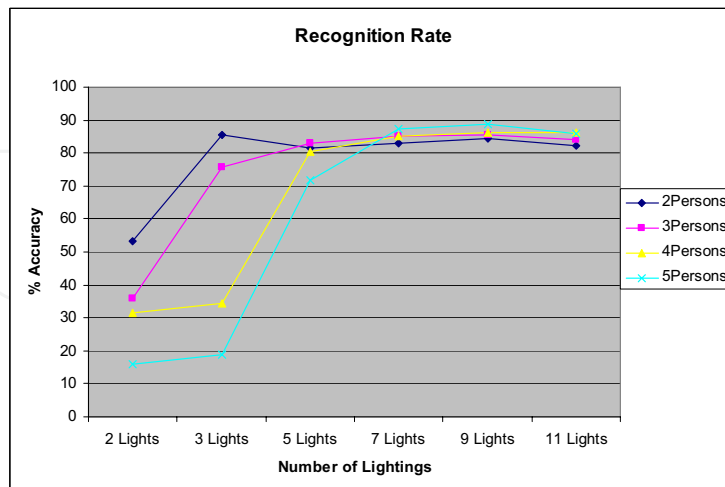

Figure 6. Recognition rates (in percentage) when varying the values of M and N

5. Discussion

The results obtained using the canonical face recovery algorithm is very encouraging. We have shown that the minimum number of illumination specified bootstrap images (i.e. N) needed to generate a stable canonical face ranges between 3 and 5. This makes good hardware design sense as an elaborate light stage setup (Sim et al., 2003) (Debevec et al., 2000) becomes unnecessary. Currently we are fabricating a low-cost portable light array module used to implement the canonical face recovery. Depending on the role, this lighting module can be embedded into the different stages of the face recognition system. For example, the module can be introduced during the registration stage where the objective is to capture a good quality neutral image of the subject (i.e. its canonical representation) to be registered into the system irregardless of the ambient lighting condition. Another possible use is to incorporate it into the image capture system at the front end of the recognition system. This will ensure that the picture taken and used for matching will not be affected again by external light sources. Besides using the images captured by the lighting module as described here, we can explore using shape-from-shading techniques to recover the 3D shape of the face (Zhao and Chellappa, 1999). The range information will be an added boost to improve on the illumination rendering quality as well as for recognition.

Although the illumination models recovered using the CMU PIE database generates 21 different variations they are inadequate as some important lighting directions (i.e. especially those coming from the top) are glaringly missing. We will next consider using computer graphics tools to develop a virtual light stage that has the ability to render any arbitrary lighting conditions on a 3D face. These new variations can then be used to extract finer quality illumination models which in turn can be use to synthesis more realistic novel appearance views. Finally, we will explore how the systems of canonical face recovery and appearance synthesis can play a further role in enhancing the performances of illumination challenged real world analysis systems. One possible use of this would be in the area of improving data acquisition for dermatology-based analysis where maintaining colour integrity of the image taken is extremely important.

6. Conclusion

We have developed an exemplar-based approach aim at recovering the canonical face of a person as well as the lighting models responsble for the input image. The recovery is not dependent on the number of person (i.e. M) and number of lighting positions (i.e. N). In fact, we have demonstrated that a low value of M=2 and N=2 are in fact adequate for most cases to achieve a good recovery outcome. The canonical face can either be use as a probe face for recognition or use as a base image to generate novel appearance models under new illumination conditions. We have shown subjectively that the canonical faces recovered with this approach are very stable and not heavily dependent on the types and numbers of the bootstrap images. The strength of the view synthesis algorithm based on the canonical face was further demonstrated by a series of face recognition tests using the CMU PIE images which yielded a 2.3 times recognition improvement over the existing technique.

7. Acknowledgement

8. References

Adini Y. and Ullman S. (1997). Face Recognition: the Problem of Compensating for Changes in Illumination Direction, in *Proc. of IEEE Trans. on PAMI*. Vol. 19, No. 7, pp. 721-732.

Belhumeur P., Hespanha J. and Kriegman D. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, in *Proc. of IEEE Trans. on Pattern Analysis and Machine Intelligence*. Vol. 19, No. 7, pp. 711-720.

Debevec P., Hawkins T., Tchou C., Sarokin W., and Sagar M. (2000). Acquiring the Reflectance Field of a Human Face, in Proc. of SIGGRAPH 2000, pp. 145-156.

Manjunath B., Chellappa R. and Malsburg C. (1992). A feature based approach to face recognition, in *Proc. of IEEE Computer Society. Confr. On Computer Vision and Pattern Recognition*, pp. 373-378.

Mariani R., (2002). A Face Location and Recognition System Based on Tangent Distance, *Multimodal Interface for Human-Machine Communication*, Vol. 48. Ed. Yuen P., Tang Y. and Wang P., World Scientific, pp 3-31.

More J. (1978). The Levenberg-Marquardt Algorithm: Implementation and Theory, in G. Watson, Ed, *Lecture Notes in Mathematics*, Springer Verlag, pp. 105-116.

Riklin-Raviv T. and Shashua A. (2001). The Quotient Image: Class based Re-rendering and Recognition with Varying Illuminations, in *Proc. of IEEE Trans. on PAMI*. Vol. 23, No. 2, pp. 129-139.

Sim T. and Kanade T. (2001). Combining Models and Exemplars for Face Recognition: an Illumination Example, in *Proc. of Workshop on Models versus Exemplars in Computer Vision*, Dec.

Sim T., Baker S. and Bsat M. (2003). The CMU Pose, Illumination and Expression Database, in *Proc. of IEEE Trans. on PAMI*, Vol. 25, No. 12, pp. 1615-1618.

Smith W., Robles-Kelly A. and Hancock E. (2004). Skin Reflectance Modelling for Face Recognition, in *Proc. of the Int'l Confr. on Pattern Recognition*, pp. 210-213.

Yang P., Shan S., Gao W., Li S. and Zhang D. (2004). *Face recognition using Ada-boosted Gabor features*, FG 2004, pp. 356-360.

Yeredor A. (2000). The Extended Least Squares Criterion: Minimisation Algorithms and Applications, *IEEE Trans. on Signal Processing*, Vol. 49, No. 1, pp. 74-86.

Zhao W. and Chellappa R. (1999). Robust Face Recognition using Symmetric Shape-from-Shading, *Technical Report* CARTR-919, Centre for Automation Research, University of Maryland, College Park, MD.

Face Recognition

Edited by Kresimir Delac and Mislav Grgic

ISBN 978-3-902613-03-5

Hard cover, 558 pages

Publisher I-Tech Education and Publishing

Published online 01, July, 2007

Published in print edition July, 2007

This book will serve as a handbook for students, researchers and practitioners in the area of automatic (computer) face recognition and inspire some future research ideas by identifying potential research directions. The book consists of 28 chapters, each focusing on a certain aspect of the problem. Within every chapter the reader will be given an overview of background information on the subject at hand and in many cases a description of the authors' original proposed solution. The chapters in this book are sorted alphabetically, according to the first author's surname. They should give the reader a general idea where the current research efforts are heading, both within the face recognition area itself and in interdisciplinary approaches.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Tele Tan (2007). From Canonical Face to Synthesis - an Illumination Invariant Face Recognition Approach, Face Recognition, Kresimir Delac and Mislav Grgic (Ed.), ISBN: 978-3-902613-03-5, InTech, Available from: http://www.intechopen.com/books/face_recognition/from_canonical_face_to_synthesis_-_an_illumination_invariant_face_recognition_approach

INTECH
open science | open minds