Model-Based Illumination Correction for Face Images in Uncontrolled Scenarios

Bas Boom, Luuk Spreeuwers, and Raymond Veldhuis

University of Twente, EEMSC, Signals & Systems,
P.O. box 217, 7500 AE, Enschede, The Netherlands
{b.j.boom,l.j.spreeuwers,r.n.j.veldhuis}@utwente.nl

Abstract. Face Recognition under uncontrolled illumination conditions is partly an unsolved problem. Several illumination correction methods have been proposed, but these are usually tested on illumination conditions created in a laboratory. Our focus is more on uncontrolled conditions. We use the Phong model which allows us to model ambient light in shadow areas. By estimating the face surface and illumination conditions, we are able to reconstruct a face image containing frontal illumination. The reconstructed face images give a large improvement in performance of face recognition in uncontrolled conditions.

1 Introduction

One of the major problems with face recognition in uncontrolled scenarios is the illumination variation, which is often larger than the variations between individuals. We want to correct for these illumination variations in a single face image. In literature, several methods have been proposed to make face images invariant to illumination. These methods can be divided into two categories: The first category contains methods that perform preprocessing based on the local regions, like Histogram Equalization [1] or (Simplified) Local Binary Patterns [2,3]. These methods are direct and simple, but fail to model the global illumination conditions. The methods in second category estimate a global physical model of the illumination mechanism and its interaction with the facial surface. One of the earlier methods in this category is the Quotient Image [4], which estimates illumination in a single image allowing the computation of a quotient image. More recent correction methods [5,6] are also able to deal with shadows and reflections using an addition error term. In our experience, these methods work on images with illumination conditions created in a laboratory, but fail in uncontrolled scenarios. In [7], 3D morphable models are used to simulate the illumination conditions in a single images, calculating shadows and reflections properly. The disadvantage of this method is the computational cost for a single image. In [8], a illumination normalization is proposed for uncontrolled conditions which requires a color image together a 3D range image.

We developed a new method for illumination correction in [9] which used only a single grey level image, but this method improved the recognition for

X. Jiang and N. Petkov (Eds.): CAIP 2009, LNCS 5702, pp. 33–40, 2009.

face images taken under uncontrolled conditions. During these experiments, we discovered that both our method and [5] have problems modelling shadow areas which still contain some reflection. This often occurs in face images taken under uncontrolled conditions. Furthermore, we observed that the found surface normals were not restricted by the geometrical constrains. In this paper, we tried to solve these issues by improving our previous method.

2 Illumination Correction Method

2.1 Phong Model

To model the shadow areas that contain some reflections, we use the Phong model, which explains these areas using the ambient reflection term. In our previous work and in [5], the Lambertian model with a summed error term was used to model the shadows. This however fails when both the intensities of the light source on the face and the reflections in the shadow areas vary. The Phong model in combination with a shadow expectation is able to model these effects. If we assume a single diffuse light source l, the Phong model is given by the following Equation:

$$b(\mathbf{p}) = c_a(\mathbf{p})i_a + c_d(\mathbf{p})\mathbf{n}(\mathbf{p})^T\mathbf{s}_l i_d + \text{specular reflections} \tag{1}$$

The image $b(\mathbf{p})$ at location $\mathbf{p}$ can be modelled using three parts namely: the ambient, diffuse and specular reflections. The ambient reflections exist of the albedo $c_a(\mathbf{p})$ and the intensity of the ambient light i_a. The ambient reflections are still visible if there is no diffuse light, for instance in shadow areas which are not entirely dark. The diffuse reflections are similar to the Lambertian model, where the surface normals $\mathbf{n} \in \mathcal{R}^3$ define the direction of the reflection and together with the albedo $c_d(\mathbf{p})$ give the shape $\mathbf{h}(\mathbf{p}) = c_d(\mathbf{p})\mathbf{n}(\mathbf{p})^T$. The diffuse light can be modelled by a normalized vector $\mathbf{s} \in \mathcal{R}^3$, which gives the light direction and the intensity of the diffuse light i_d. The final term contains the specular reflections, which explain the highlights in the image, but because this phenomenon is usually only present in a very small part of the image we will ignore this term.

The shadow can be modelled as a hard binary decision. If a light source can not reach a certain region, it makes a shadow. This holds expect for areas which contain the transition between light and shadow. Using a 3D range map of a face, we can compute the shadow area given a certain light direction using a ray tracer. Computing these shadow areas for multiple images, allows us to calculate an expectation of shadow $e_l(\mathbf{p})$ on the position $\mathbf{p}$ for the light directions $\mathbf{s}_l$. This gives us a user independent shadow model given the light direction.

$$b(\mathbf{p}) = c(\mathbf{p})i_a + c(\mathbf{p})\mathbf{n}(\mathbf{p})^T\mathbf{s}_l i_d e_l(\mathbf{p}) \tag{2}$$

$$\mathbf{b} = \mathbf{c}i_a + H\mathbf{s}_l i_d \star \mathbf{e}_l \tag{3}$$

In Equation 2, we simplified the Phong model and we added the expectation term $e_l(\mathbf{p})$ to model shadows. We also use the same albedo term for ambient and diffuse illumination, which is common practice [7]. In Equation 3, we vectorized all the terms, where the $\star$ denotes the Cartesian product. Our goal is to find the face shape and the light conditions given only a single image.

2.2 Search Strategy for Light Conditions and Face Shape

An method to estimate both the face shape and the light conditions is to vary one of the variables and calculate the others. In our case, we chose to vary the light direction allowing us to calculate the other variables. After obtaining the other variable, e.g. light intensity, surface and albedo, we use an evaluation criteria to see which light direction gives the best estimates. The pseudo-code of our correction method is given below:

- For a grid of light directions $\mathbf{s}_l$
  - Estimate the light intensities i_a and i_d
  - Estimate the initial face shape
  - Estimate the surface using geometrical constrains and a 3D surface model
  - Computing the albedo and its variations
  - Evaluation of the found parameters
- Refine the search to find the best light direction.
- Reconstruct a face images under frontal illumination.

We start with a grid where we vary the azimuth and elevation of the light direction with 20 degrees. The grid allows us to locate the global minimum, from there we can refine the search using the downhill simplex search method [10] to find the light direction with an accuracy of ± 2 degrees. Using the found parameters like light conditions and face shape, we can reconstruct a face image under frontal illumination, which can be used in face recognition. In the next sections, we will discuss the different components mentioned in the pseudo-code.

2.3 Estimate the Light Intensities

Given the light direction $\mathbf{s}_l$ and the shadow expectation $\mathbf{e}_l(\mathbf{p})$, we can estimate the light intensities using the mean face shape $\overline{\mathbf{h}}(\mathbf{p})$ and mean albedo $\overline{c}(\mathbf{p})$. The mean face shape and albedo are determined using a set of face images together with there 3D range maps. This gives us the following linearly solvable equation, allow us to obtain the light intensities $\{i_a, i_d\}$:

$$\{i_a, i_d\} = \arg \min_{\{i_a, i_d\}} \sum_{\mathbf{p}} \|b(\mathbf{p}) - \overline{c}(\mathbf{p}) i_a - \overline{\mathbf{h}}^T \mathbf{s}_l i_d e_l(\mathbf{p})\|^2 \tag{4}$$

Because this is an over-determined system, we can use the mean face shape and mean albedo to estimate the light intensities, which still gives a very accurate estimation. However, this might normalize the difference in intensity of the skin color. If the light intensities are negative, we skip the rest of the computations.

2.4 Estimate the Initial Face Shape

To estimate the initial face shape given the light conditions $\{\mathbf{s}_l, \mathbf{e}_l(\mathbf{p}), i_a, i_d\}$, we use the following two assumptions: Firstly, the Phong model must hold, which gives us the following equations:

$$b(\mathbf{p}) = c(\mathbf{p})i_a(\mathbf{p}) + h_x(\mathbf{p})s_{x,l}i_d e_l(\mathbf{p}) + h_y(\mathbf{p})s_{y,l}i_d e_l(\mathbf{p}) + h_z(\mathbf{p})s_{z,l}i_d e_l(\mathbf{p}) \quad (5)$$

Secondly, the face shape should be similar to the mean face shape. This can be measure by taking the Mahalanobis distance between the face shape $\mathbf{h}(\mathbf{p})$ and the mean face shape $\overline{\mathbf{h}}(\mathbf{p})$. Using Lagrange multipliers, we can minimize the distance with Equation 5 as a constrain. This allows us to find an initial face shape $\hat{\mathbf{h}}(\mathbf{p})$, which we will improve in the next steps using a surface model together with geometrical constrains.

2.5 Estimate Surface Using Geometrical Constrains and a 3D Surface Model

Given an estimate of the face shape $\hat{\mathbf{h}}$, we want to determine the surface $\mathbf{z}$, which is a depth map of the face image. Given a set of 3D range images of faces, we can calculate depth maps $\{\mathbf{z}^t\}_{t=1}^T$ and we can obtain the mean surface $\overline{\mathbf{z}}$ and a covariance matrix $\Sigma_\mathbf{z}$. Using Principal Component Analysis (PCA), we computer the subspace by solving the eigenvalue problem:

$$\Lambda_\mathbf{z} = \Phi^T \Sigma_\mathbf{z} \Phi \qquad\qquad \hat{\mathbf{z}} = \overline{\mathbf{z}} + \sum_{k=0}^{K} \Phi_k u_\mathbf{z}(k) \quad (6)$$

where $\Lambda_\mathbf{z}$ are the eigenvalues and Φ are the eigenvectors of the covariance matrix $\Sigma_\mathbf{z}$, which allows to express the surface in variations $\mathbf{u_z}$ for the mean surface $\overline{\mathbf{z}}$. We also know that $h_{zx}(\mathbf{p}) = \frac{h_z(\mathbf{p})}{h_x(\mathbf{p})} = \nabla_x z(p)$ and $h_{zy}(\mathbf{p}) = \frac{h_z(\mathbf{p})}{h_y(\mathbf{p})} = \nabla_y z(\mathbf{p})$ holds, where ∇_x and ∇_y denote the gradient in x and y direction. This allows us to calculate the variations of the surface $\mathbf{u_z}$ using the following equation:

$$\mathbf{u_z} = \arg\min_{\mathbf{u_z}} ||\nabla_x \overline{\mathbf{z}} + \nabla_x \Phi \mathbf{u_z} - \hat{\mathbf{h}}_{zx}||^2 + ||\nabla_y \overline{\mathbf{z}} + \nabla_y \Phi \mathbf{u_z} - \hat{\mathbf{h}}_{zy}||^2 \quad (7)$$

The surface $\hat{\mathbf{z}}$ can be found using Equation 6 and from this surface we can also find the surface normals $\mathbf{n}(\mathbf{p})$. In this case, the surface normals are restricted by geometrical constrains. Using only the geometrical constrains does not have to be sufficient to determine the face surface, therefore, we use the surface model to ensure the convergence.

2.6 Computing the Albedo and Its Variations

In the previous sections, we obtained the surface normals $\mathbf{n}(\mathbf{p})$ and the illumination conditions $\{\mathbf{s}_l, \mathbf{e}_l(\mathbf{p}), i_a, i_d\}$. This allows us to calculate the albedo $\mathbf{c}$ from Equation 2. In order to find out whether the albedo is correct, we also create a PCA model of the albedo. Given a set of face images together with their 3D range maps, we estimated the albedo, see [9]. Vectorizing the albedo $\{\mathbf{c}^t\}_{t=1}^T$ allows us to calculate a PCA model and find the variations $\mathbf{u_c}$, which is also used for the surface model. Using the variations $\mathbf{u_c}$, we calculated also a projection of albedo $\hat{\mathbf{c}}$ to PCA model. The projection $\hat{\mathbf{c}}$ does not contain all details necessary for the face recognition. For this reason, we use the albedo $\hat{\mathbf{c}}$ from the PCA model in the evaluation criteria, while we use the albedo $\mathbf{c}$ obtained from Equation 2 in the reconstructed image.

2.7 Evaluation of the Found Parameters

Because we calculate the face shape for multiple light directions, we have to determine which light direction results in the best face shape. Furthermore, the downhill simplex algorithms also needs an evaluation criteria to be able to find the light direction more accurately. Using the found light conditions and face shape, we can reconstruct an image $\mathbf{b}_r$ which should be similar to the original image. This can be measured using the sum of the square differences between the pixels values. Minimizing this may cause overfitting of our models at certain light directions. For this reason, we use the maximum a posterior probability estimator given by $P(\mathbf{u_c}, \mathbf{u_z}|\mathbf{b})$, which can be minimized by the following equations, see [7]:

$$E = \frac{1}{\sigma_b} \sum_p \|b(\mathbf{p}) + b_r(\mathbf{p})\|^2 + \sum_{k=1}^{K} \frac{u_{\mathbf{z}}^2(k)}{\lambda_{\mathbf{z}}(k)} + \sum_{j=1}^{J} \frac{u_{\mathbf{c}}^2(j)}{\lambda_{\mathbf{c}}(j)} \tag{8}$$

In this case, σ_b controls the relative weight of the prior probability, which is the most important factor to minimize. $\lambda_{\mathbf{z}}$ and $\lambda_{\mathbf{c}}$ are the eigenvalues of the surface and albedo. The light directions that minimizes Equation 8, give us the parameters from which we can reconstruct a face image with frontal illumination.

3 Experiments and Results

We correct for illumination by estimating both the illumination conditions and the face surface. In this section, we will show some of the estimate surfaces together with their corrected images. The main purpose of the illumination correction is to improve the performance of the face recognition method. Our goal is therefore to demonstrate that our face recognition method indeed benefits from the improvement in the illumination correction. For this purpose, we use the FRGCv1 database where we have controlled face images in the enrollment and uncontrolled face images as probe images.

3.1 3D Database to Train the Illumination Correction Models

For our method, a database is needed that contains both face images and 3D range maps to compute the surface, shape, shadow and albedo models. In this case, we used the Spring 2003 subset of Face Recognition Grand Challenge (FRGC) database, which contains face images together with their 3D range maps. These face images contain almost frontal illumination and no shadows, making this subset of the database ideal to compute the surface and albedo. The exact method to retrieve the albedo is describe in [9].

3.2 Recognition Experiment on FRGCv1 Database

The FRGCv1 database contains frontal face images taken under both controlled and uncontrolled illumination conditions as is shown in Figure 1. The first image in Figure 1 is taken under controlled conditions, while the other images are taken under uncontrolled conditions. For the first person, we show that our method is able to correct for different unknown illumination conditions. In case of the last image, we observe more highlighted areas directly under the eyes, this is caused by the reflection of the glasses which are not modelled by our method.

In order to test if this illumination correction method improves the performance in face recognition, we performed the following experiment to see if illumination conditions are removed in the images taken under uncontrolled conditions. In this case, we use the images with uncontrolled illumination as probe image and make one user template for every person with the images taken under controlled conditions. To train our face recognition method, we randomly divided both the controlled and uncontrolled set of the FRGCv1 database into two parts, each containing approximately half of the face images. The first halves are used to train the face recognition method. The second half of the controlled set is used to compute the user templates, while the second half of the uncontrolled set is used as probe images. We repeat this experiment 20 times using different images in both halves to become invariant against statistical fluctuations. The Receiver operating characteristic (ROC) in Figure 2 is obtained using the PCA-LDA likelihood ratio [11] for face recognition. The first three lines in Figure 2 are also stated in our previous work [9], where the last line depicts the improvements obtained using the method described in this paper. The ROC curves in Figure 2 shows only the improvements due to illumination correction. In the

Fig. 1. First row contains the original images from the FRGCv1 database, second and third row show the resulting surface, the fourth row depicts the reconstructed frontal illumination

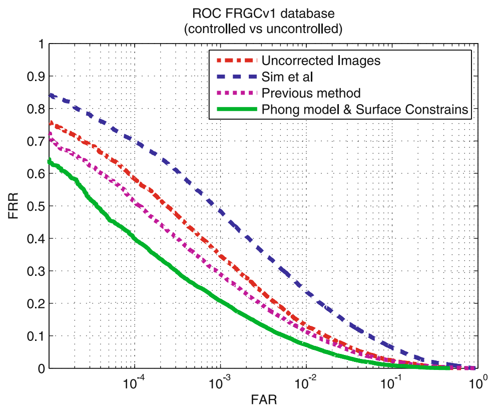

Fig. 2. ROC on the FRGCv1 database with a comparison to our previous work [9] and to work of Sim et al [5]

case of our previous method, the False Reject Rate (FRR) becomes significantly better at a False Accept Rate (FAR) smaller than 1%, while the last line is overall better. Most other illumination correction methods like [6, 8] evaluated their method only on a database create in a laboratory or do not perform a recognition experiment, which makes the comparison with other methods difficult.

4 Discussion

In figure 1, we observe that the phenomenon, where the shadow areas are still not completely dark, often occurs in uncontrolled illumination conditions. Improving our model on this point gave also improvements in the recognition results, which was the main purpose of our illumination correction. We choose to ignore other illumination effects like specular reflections, because we expected a small performance gain in face recognition and a large increase in computation time.

The second improvement is a restriction to the face shape by computing the surface instead of the surface normals, which also slightly improved the face recognition results. Another benefit is that we obtained an estimation of the surface of the face, which might be handy in other applications. In our research, the focus has not been on the quality of the estimated surfaces. Although we expect that this can be an interesting result to improve for instance 3D face acquisition and recognition.

5 Conclusion

We present two major improvements for our illumination correction method for face images, where the purpose of our method is to improve the recognition results of images taken under uncontrolled illumination conditions. The first improvement uses a better illumination model, which allows us to model the ambient light in the shadow areas. The second improvement computes an estimate of the surface given a single face image. This surface gives us a more accurate face shape and might also be useful in other applications. Because of both improvements, the performance in face recognition becomes significantly better for face images with uncontrolled illumination conditions.

References

1. Shan, S., Gao, W., Cao, B., Zhao, D.: Illumination normalization for robust face recognition against varying lighting conditions. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, AMFG 2003, October 17, pp. 157–164 (2003)
2. Heusch, G., Rodriguez, Y., Marcel, S.: Local binary patterns as an image preprocessing for face authentication. In: 7th International Conference on Automatic Face and Gesture Recognition, FGR 2006, April 10-12, pp. 6–14 (2006)
3. Tao, Q., Veldhuis, R.N.J.: Illumination normalization based on simplified local binary patterns for a face verification system. In: Biometrics Symposium 2007 at The Biometrics Consortium Conference, Baltimore, Maryland, USA, September 2007, pp. 1–6. IEEE Computational Intelligence Society, Los Alamitos (2007)
4. Shashua, A., Riklin-Raviv, T.: The quotient image: class-based re-rendering and recognition with varying illuminations. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(2), 129–139 (2001)
5. Sim, T., Kanade, T.: Combining models and exemplars for face recognition: An illuminating example. In: Proc. CVPR Workshop on Models versus Exemplars in Computer Vision (December 2001)
6. Basri, R., Jacobs, D.: Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(2), 218–233 (2003)
7. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3d faces. In: IGGRAPH 1999: Proceedings of the 26th annual Conference on Computer Graphics and Interactive Techniques, pp. 187–194 (1999)
8. Al-Osaimi, F.R., Bennamoun, M., Mian, A.: Illumination normalization for color face images. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Nefian, A., Meenakshisundaram, G., Pascucci, V., Zara, J., Molineros, J., Theisel, H., Malzbender, T. (eds.) ISVC 2006. LNCS, vol. 4291, pp. 90–101. Springer, Heidelberg (2006)
9. Boom, B., Spreeuwers, L., Veldhuis, R.: Model-based reconstruction for illumination variation in face images. In: 8th International Conference on Automatic Face and Gesture Recognition, FGR 2008 (2008)
10. Nelder, J., Mead, R.: A simplex method for function minimization. The Computer Journal 7, 308–313 (1965)
11. Veldhuis, R., Bazen, A., Booij, W., Hendrikse, A.: Hand-geometry recognition based on contour parameters. In: Proceedings of SPIE Biometric Technology for Human Identification II, Orlando, FL, USA, March 2005, pp. 344–353 (2005)