Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 2 (2010) 11–19

Procedia Computer Science

www.elsevier.com/locate/procedia

ICEBT 2010

A efficient and practical 3D face scanner using near infrared and visible photometric stereo

Gary A. Atkinson, Mark F. Hansen, Melvyn L. Smith, Lyndon N. Smith

The University of the West of England, Bristol BS16 1QY

Abstract

This paper is concerned with the acquisition of model data for automatic 3D face recognition applications. As 3D methods become progressively more popular in face recognition research, the need for fast and accurate data capture has become crucial. This paper is motivated by this need and offers three primary contributions. Firstly, the paper demonstrates that four-source photometric stereo offers a potential means for data capture that is computationally and financially viable and easily deployable in commercial settings. We have shown that both visible light and less intrusive near infrared light is suitable for facial illumination. The second contribution is a detailed set of experimental results that compare the accuracy of the device to ground truth, which was captured using a commercial projected pattern range finder. Importantly, we show that not only is near infrared light a valid alternative to the more commonly exploited visible light, but that it actually gives more accurate reconstructions. Finally, we assess the validity of the Lambertian assumption on skin reflectance data and show that better results may be obtained by incorporating more advanced reflectance functions, such as the Oren-Nayar model.

Keywords: 3D reconstruction, face recognition, photometric stereo

1. Introduction

The variability of face images with lighting and pose result in one of the most significant pitfalls of current 2D face recognition methods. This is the prime motivation for the move of recent years to incorporate 3D information [1]. The 3D shape of a face is completely invariant to the illumination distribution and pose changes can more readily be corrected for if the 3D structure of the face is known.

One problem with 3D face recognition however, is the difficulty of acquiring the necessary 3D face models in order to perform the actual recognition. This problem is the motivation for our work. The paper aims to use a relatively inexpensive and easily deployable system to capture 3D face geometries and assess the accuracy of reconstructions for face recognition. The device uses four-source photometric stereo (PS) synchronised at high-speed to a camera and triggered by an ultrasound proximity sensor [2]. We apply our experiments using both visible-light flash-gun sources and the more covert near-infrared (NIR) LED cluster sources. The results in this paper make the system attractive for use in many commercial and industrial settings such as at entrances to high security areas, airport check-in and border control.

Email address: gary.atkinson@uwe.ac.uk (Gary A. Atkinson)

1877-0509
doi:10.1016/j.procs.2010.11.003

1.1. Related Work

The use of 3D information for face recognition has been attracting increasing attention in recent years [1, 3, 4]. Popular methods include the 3D morphable model of Blanz and Vetter [5] and the geodesic representations of Bronstein et al. [6] and Mpiperis et al. [7]. Research that directly compares 2D and 3D recognition frequently reports improved success rates for 3D recognition and that the best results occur when 2D and 3D information is fused [1]. As demand for practical face recognition systems is likely to increase, it is important that the most accurate methods are used and that the acquisition devices are both practical and affordable. There are a number of existing ways to capture and reconstruct 3D face information and the benefits and limitations of the most common approaches will now be discussed with the aim of putting our device into context.

Structured light scanning is perhaps the best known approach to generating 3D models of faces. This was used for generating the morphable head model in [5] and also for all the 3D faces used in the renowned FRGC2.0 dataset [8]. The Minolta Vivid 910 device [9] used for FRGC2.0 has a quoted accuracy of ±0.10mm. However, these devices take about 2.5s to capture the data, during which time the subject could move, thus distorting the reconstruction and are sensitive to high levels of ambient illumination. A much faster alternative is to project patterns across the entire face, as in the 3dMD system [10], which is used in this paper to acquire ground truth models. This device uses a number of cameras to take images of an object from different positions and uses the projected patterns to solve the correspondence problem between the images. The benefits of this system are its high accuracy (reported as <0.2mm) and the speed of image acquisition (1.5ms). However, the processing time is much longer, the system is expensive and it requires a time consuming calibration procedure.

A much cheaper method to acquire the 3D shape of a face is to use shape-from-shading (SFS) where the naturally occuring intensity patterns of a face are used to extract the 3D geometry from a single image [11]. However the problem is ill-posed, meaning that there is no guarantee of a unique solution for a given image [12]. Smith and Hancock [13] use a statistical model in the SFS paradigm to recover geometry. An alternative is to photograph the object multiple times under different illumination. This technique is known as Photometric Stereo (PS) and was first devised by Woodham [14] who showed that for any Lambertian surface, three differently illuminated images are sufficient to remove the ambiguity associated with a single image. Several researchers have attempted to address issues of shadows, specularities and non-Lambertian reflectance with PS [15, 16, 17, 18, 19, 20].

Of the vast amount of research into automatic face recognition during the last two decades [21], relatively little work has involved PS. Kee et al. investigate the use of 3-source PS under dark room conditions [22]. They were able to determine the optimal light source arrangement and demonstrate a working recognition system. Zhou, Chellappa and Jacobs apply rank, integrability and symmetry constraints to adapt PS to face-specific applications [23]. Zhou et al. extended a PS approach to unknown light sources [24]. Georghiades et al. show how reconstructions from PS can be used to form a generative model to synthesise images under novel pose and illumination [25].

Comparing point clouds, the shape index, depth maps, profiles and surface normals in terms of face recognition performance, Gökberk et al. [4] concluded that surface normals provide the best features for face recognition. It is surprising therefore, that so few applications to date utilise PS, which inherently generates surface normals.

1.2. Contributions

The first contribution of this paper is a qualitative demonstration of the capabilities of our PS-based face scanner. The device, shown in Fig. 1, is suitable for practical recognition environments and consists of four illumination sources placed evenly around a high-speed camera. Subjects walk through the archway towards the camera located on the back panel and exit through the side. Compared to existing technologies, our device is cheap to build and involves exceptionally short image capture and processing times. The device is also able to operate at high resolution, is robust to ambient illumination and requires only minimal calibration. All images are captured in approximately 20ms, resulting in only very small misalignment between frames. This allows subjects to be imaged as they casually walk through the archway. We have tested our device using both visible and NIR illumination sources.

The second contribution is a thorough analysis of the accuracy of the device using both visible and NIR illumination. We found that the latter yields more accurate reconstructions when compared with ground truth. To the best of our knowledge, no published research has looked at using NIR light sources in PS for the purpose of face recognition. All ground truth experiments are based on models generated by a commercial 3dMD scanner [10] and are presented using RMS height errors and ℓ_2-norm errors.

Figure 1: The geometry capture device. Enlarged areas from top to bottom: a NIR light source, a visible light source and the ultrasound proximity trigger. The camera can be seen on the back panel.

The final contribution is a quantitative analysis of the validity of the Lambertian assumption on skin reflectance. The extent of the discrepancies between the measured skin reflectance and Lambert's Law are demonstrated graphically and shown to be relatively minor for non-grazing angles. We also show that skin is more Lambertian under NIR illumination. Lastly, the reflectance analysis demonstrates the possibilities of improving the reconstructions by incorporating the Oren-Nayar reflectance model into the method.

2. Data capture

Figure 1 shows a photograph of the device that we have constructed for data acquisition. The person walks towards camera on the back panel from the left. An ultrasound proximity switch detects the presence of the individual and triggers the acquisition procedure. Four images are captured using the high speed camera on the back panel with the face illuminated by four light sources in sequence. Our rig allows either Jessops 100M visible-light flashguns to be used (colour temperature 5600K) or stripped down NIR X-vision VIS080IR lensed 7-LED clusters ($\approx$ 850nm).

It is generally expected that the face will be moving at the time of acquisition. For this reason, it is necessary to use a high speed camera to rapidly acquire the images as the light sources change before significant motion is possible. We therefore use a Basler A504kc 1280×1024 pixel camera operating at 200 fps. It was found experimentally that this was the frame rate necessary to avoid significant face movement between images. The light sources are synchronised to the camera frames using FPGA interfacing [2]. All interfacing and synchronisation is programmed in LabVIEW (although the image processing and shape estimation are performed in MATLAB). The sensor used to initiate the entire process is a highly directional Baumer ultrasound proximity switch.

Four greyscale images are captured by the camera with each corresponding to one of the four visible or NIR light sources. The regions containing the actual face are extracted from the background using the method described by Lienhart and Maydt [26]. We then estimate of the field of surface normals using a standard photometric stereo technique and assuming known light source directions [27, §5.4]. Finally, we integrate these surface normals using the well-established Frankot-Chellappa method [28] to recover the height map estimate. Figure 2 shows an example of four raw images and the resultant height estimate. For each test presented in this paper, we either use the four visible light sources or the four NIR sources.

One disadvantage of the visible light set-up is that the firing of flashguns is obvious to the subject and possibly intrusive to any surrounding people. NIR light by contrast, is more covert for a face recognition environment and subjects are less inclined to "pose" for the camera, meaning that more neutral expressions are likely. It is also worth noting the advantage that many camera sensors are inherently more sensitive to NIR light. One disadvantage of NIR illumination is the relative difficulty in obtaining the necessary brightness for the required short exposure times. While the flashguns were easily bright enough with an exposure time of 1ms, an exposure of 5ms was needed for the NIR

Figure 2: Example of four raw images and the resultant surface reconstruction.

Figure 3: Example raw images and reconstructions using visible (top) and NIR light sources for four subjects. For these experiments only, the subjects were asked to rest their chin on a support in order to ensure that all subjects are compared to each other in fair conditions.

LEDs (i.e. the maximum possible exposure for the given frame rate). Although this was adequate for our experiments, we had to use LED lenses that provided a narrow divergence angle, meaning that the face had to be more precisely positioned to obtain full illumination. For the visible light sources, the images were bright enough even for large diversion angles, removing the need for accurate positioning of apparatus and allowing subjects to pass through the archway without having to consider their exact location with respect to the camera.

To account for ambient illumination, a control image is taken after the final light source and subtracted from the other four images. The resultant images are then normalised in terms of intensity before reconstruction takes place. This is done by linearly scaling the greylevels of each image so that the mean intensity was equal for each image.

3. Reconstruction Analysis

Figure 3 shows a set of reconstructions using visible and NIR light. The general 3D structures of the faces have clearly been well estimated qualitatively. In order to quantitatively assess the accuracy of the reconstructions, we scanned eight different faces using a commercial 3dMD projected pattern range finder [10]. The 3dMD models were rescaled so that the distance between tear ducts was the same as in the corresponding PS reconstruction. All reconstructions were then cropped to 160×200px regions centred on the nose tip that encompass the eyebrows and mouth. Part of the forehead is omitted by this choice of cropping region as it is frequently occluded by hair and is therefore deemed unreliable for face recognition. An example of the face regions used for comparison can be seen in Fig. 4, which also shows a ground truth reconstruction acquired using a 3dMD scanner. The face regions from visible and NIR light sources are then aligned to ground truth using the Iterative Closest Point (ICP) algorithm [29].

Individual RMS and ℓ_2-norm error results between the reconstructions and ground truth are displayed in Fig. 5. The eight subjects consist of 6 males and 2 females and a mixture of Caucasian and Asian ethnicities. The variations in residual errors and ℓ_2-norm distances between visible and NIR reconstructions are significant according to paired t-tests ($p = 0.05$). This demonstrates that PS using NIR as a light source is a perfectly valid approach and leads to more accurate reconstructions.

In order to obtain an indication of the regions where the greatest differences occur between ground truth and PS reconstructions, the residuals and ℓ_2-norm errors at each pixel are plotted in Fig. 6. Typically, the largest variations occur in regions with the highest curvatures, such as eye sockets, nose tips and the sides of the nose.

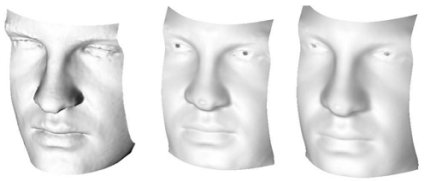

Figure 4: 3D Reconstructions for one subject from a 3dMD scanner (left) which is used as ground truth, PS using visible light sources (middle), and PS using NIR sources (right).

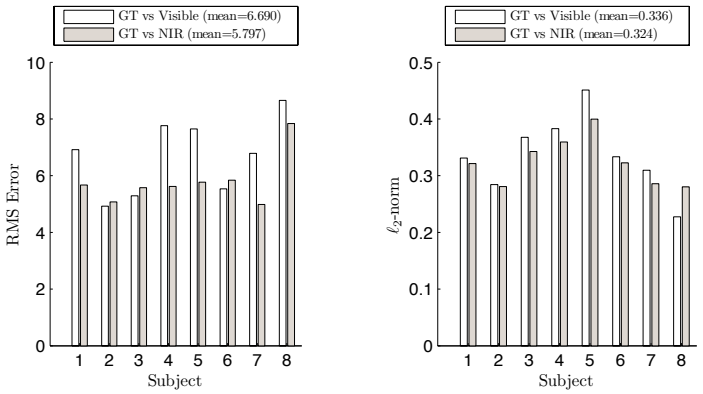

Figure 5: RMS (left) and ℓ_2-norm errors between Ground Truth (GT) and visible PS and NIR PS for each subject. The order of subjects is arbitrary.

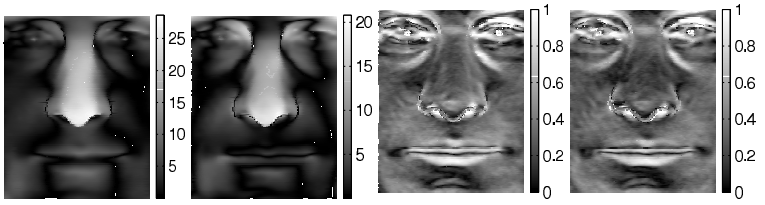

Figure 6: Representative examples of the residuals and the ℓ_2-norm errors at each pixel. Left to right: residuals for visible and NIR respectively, ℓ_2-norm errors for visible and NIR respectively. Lighter areas represent larger errors.

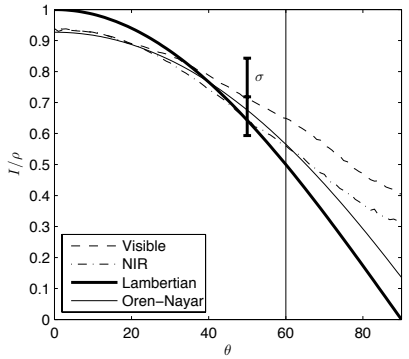

Figure 7: Mean I/ρ values averaged over 8 subjects against θ. To the right of the vertical line at $\theta = 60°$, data were too sparse to be of significance. For reference, *one* standard deviation is shown to give an indication of spread.

In attempting to produce the most accurate reconstructions possible via PS, it was found that the estimated surface normals could be enhanced by using normals acquired by re-differentiating the reconstructed height map estimate. It is unclear as to why this should be the case but preliminary analysis indicates that the reason may be due to the imposition of integrability constraints and the fitting of limited basis functions in the Fourier domain [30], as required by our adopted integration method. These factors may cause errant normals to be "smoothed out" leading to a more accurate reconstruction. However, if this method of improving reconstructions is used, a second integration step would be needed thus removing one of the benefits of PS for face recognition: that the surface normals (and hence distinctive differential surface features) are recovered directly. More research is required into this area in order to confirm that the improvements result from the imposed integrability constraints.

4. Reflectance Analysis

To determine whether Lambert's law is obeyed more strictly under NIR light than visible light, we have plotted graphs of I/ρ against the angle between the light source and the normal vector, θ. For a purely Lambertian surface, the relationship between the two should follow a cosine law. The results can be seen in Fig. 7. To generate the graph, values of I, ρ and θ were estimated for each pixel of each image for each of eight faces. The angle θ is calculated for each point of the face from the 3dMD scan data and the known light source vectors. The average values of I/ρ are used for each 1° increment in θ. The line at $\theta = 60°$ indicates a reasonable cut-off point after which data points become too sparse to be significant. The RMS difference between the measured curves and the cosine curve in the range of $0 \leq \theta \leq 60$ is 0.04 (s.d. 0.11) for NIR light and 0.06 (s.d. 0.12) for visible. For completeness, the RMS difference across the whole curve is 0.11 (s.d. =0.13) for NIR light and 0.17 (s.d. =0.12) for visible. The figure demonstrates that skin under NIR light is marginally more Lambertian than under visible light.

Although the data suffers from significant noise levels, the NIR condition has a lower RMS error and is therefore closer to the Lambertian curve than for visible light. This difference is significant given the large numbers of pixels and subjects used in the trials. This represents an average pixel intensity error of 10 greylevels for NIR and 15 for visible light across the image, assuming a maximum of 256 grey level intensities. We believe that this result is related to the fact that NIR light penetrates more deeply into the skin than visible light [31], which facilitates a more uniform scattering than surface reflection. Note however, that neither the Lambertian model nor the Oren-Nayar model (see below) take account of internal scattering or Fresnel effects. The results above demonstrate that the more Lambertian behaviour associated with NIR light also leads to more accurate reconstructions. A more detailed analysis for two individual subjects is shown in Fig. 8 and Table 1. There are small differences in the I/ρ curves caused by different light sources but this appears to have little negative impact on the reconstructions and is likely to be due to environmental effects.

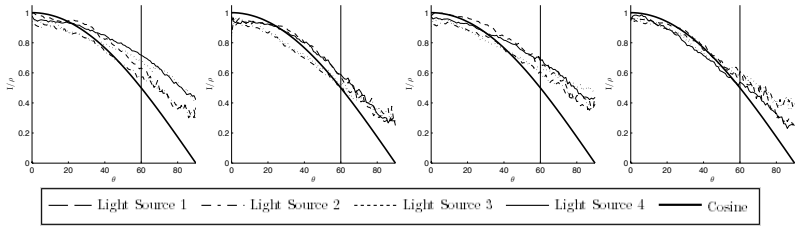

Figure 8: I/ρ values from individual light sources plotted against θ for the first two reconstructions shown in Fig. 3. Left to right: Subject 1 under visible, Subject 1 under NIR, Subject 2 under visible, Subject 2 under NIR. The light sources are labelled clockwise from the bottom-left in Fig. 1.

	Visible		NIR	
	RMS, $\theta \leq 60°$	RMS, overall	RMS, $\theta \leq 60°$	RMS, overall
All Faces	0.06, (σ = 0.11)	0.16 (σ = 0.12)	0.04 (σ = 0.12)	0.11 (σ = 0.13)
Subject 1	0.07, (σ = 0.09)	0.16 (σ = 0.18)	0.05 (σ = 0.12)	0.10 (σ = 0.22)
Subject 2	0.07, (σ = 0.10)	0.17 (σ = 0.18)	0.04 (σ = 0.13)	0.12 (σ = 0.21)

Table 1: The RMS collective error across all eight reconstructions and for the first two reconstructions shown in Fig. 3 separately. The standard deviations are shown in brackets.

Comparison to the Oren-Nayar model

We have also compared our reflection measurements to the Oren-Nayar reflectance model [32], as shown in Fig. 7. The Oren-Nayar model represents the reflecting surface as an array of V-shaped groves of random orientation, commonly called "microfacets". The distribution of microfacet orientations is characterised by a roughness parameter and each facet is assumed to act as perfect Lambertian reflector. This model is able to account for the common feature of limb-brightening and is itself based on the earlier Torrance-Sparrow model [33] where each microfacet is assumed to be mirror-like.

We have chosen to use the Oren-Nayar model as skin is not a smooth surface and it has been shown previously to be successful on a range of materials of varying degrees of roughness [32]. We do not believe that the microscopic structure of skin closely matches the Oren-Nayar model, but are merely demonstrating how alternate methods for reflection may improve our framework in future work. Investigating the various degrees of freedom of the BRDFs is also reserved for future work. Furthermore, there are additional models for skin reflectance which take account of a huge range of physical phenomena [34, 35], but these are out of the scope of this paper.

The Oren-Nayar curve in Fig. 7 represents an example intensity profile for reference with a roughness parameter of 0.2. Clearly, this model fits the measured reflectance data significantly more accurately than the Lambertian curve, suggesting that the model could be incorporated into the method in the future. This will however, add significant complexity and computation time to the algorithm. This is because a minimisation method must be implemented in order to recover all the model parameters and to accommodate the increased number of angular degrees of freedom in the model.

5. Discussion

The results presented in this paper demonstrate that PS is an effective method for producing 3D facial reconstructions for automatic recognition. Using the device with a standard PS algorithm, LabVIEW interfacing, MatLab processing and a typical modern PC, the time between device trigger and the reconstructed height map was approximately four seconds. The construction of the hardware also lends itself well to relatively unobtrusive data capture with a minimum amount of effort from the subject.

Our system offers several benefits over commonly used existing laser triangulation and projected pattern 3D shape capture devices:

1. It is significantly cheaper to construct.

2. Acquisition time is shorter than laser triangulation systems.
3. Data processing time is shorter than projected pattern systems.
4. The method is robust to typical ambient illumination conditions.
5. It is very robust against accidental collisions (because it is tolerant to errors in the light source vectors).
6. Very fine details of the face can be reconstructed.
7. Calibration is very quick and simple and only needs to be performed after the initial light source positioning.
8. Although our system cannot reconstruct hair with high levels of accuracy, it can at least provide some details of its overall shape (see Fig. 3, for example). In contrast, laser triangulation and projected pattern systems usually fail completely with hair.

At present, the 3D reconstructions are not yet as accurate as those from projected pattern range finders. They do however provide extremely fine detail of a face such as wrinkles and pores. The reconstructions under NIR were shown to be more accurate than those under visible light and diminish the need for flashing lights, making the system less intrusive compared to visible light. Zivanov et al. [36] offer an alternative argument to ours, stating that shorter wavelength light gives better results. Their justification is that shorter wavelengths undergo less internal scattering and thus provide a crisper, more defined reconstruction. It would appear therefore that a compromise must be reached in deciding between fine detail (using Zivanov's short wavelength suggestion) and overall geometry and covertness (using our NIR method).

One current limitation of the hardware described in this paper is that it does not cope with large deviations of peoples' height. Extremely tall or short people, or wheelchair bound persons would probably trigger the device correctly, but the location of the face could be outside of the field of view of the camera. Two possible solutions for this are (1) to use two cameras and trigger sensors at different heights or (2) to increase the field of view of the camera. Another improvement which could be made involves detecting the coordinates of the face and adjusting the light source vectors accordingly to improve the accuracy of the PS reconstruction. In the current system, the light source unit vectors are calculated from a point at the centre of the camera's field of view and this is used for all reconstructions regardless of where the face is actually located. For this reason, the light source unit vectors are less accurate if the person walking through the device does not locate their face near the centre of the camera's field of view. The exact error caused by this inaccuracy is unknown, but amending the light source angles on a per person basis will improve the surface normal estimates.

6. References

[1] K. W. Bowyer, K. Chang, P. Flynn, A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition, Comput. Vis. Image Understand. 101 (1) (2006) 1–15.
[2] G. A. Atkinson, A. R. Farooq, M. L. Smith, L. N. Smith, Facial reconstruction and alignment using photometric stereo and surface fitting, in: Proc. IbPRIA, 2009, pp. 88–95.
[3] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld, Face recognition: A literature survey, ACM Comput. Surv. (2003) 399–458.
[4] B. Gökberk, M. O. İrfanoğlu, L. Akarun, 3D shape-based face representation and feature extraction for face recognition, Image and Vision Comput. 24 (8) (2006) 857–869.
[5] V. Blanz, T. Vetter, Face recognition based on fitting a 3D morphable model, IEEE Trans. Patt. Anal. Mach. Intell. (2003) 1063–1074.
[6] A. M. Bronstein, M. M. Bronstein, R. Kimmel, Three-Dimensional face recognition, Intl. J. Comp. Vis. (2005) 5–30.
[7] I. Mpiperis, S. Malassiotis, M. G. Strintzis, 3-D face recognition with the geodesic polar representation, IEEE Trans. Information Forensics and Security (2007) 537–547.
[8] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, Overview of the face recognition grand challenge, in: Proc. CVPR, Vol. 1, 2005.
[9] www.konicaminolta.com/sensingusa/products/3d/non-contact/vivid910 [9 June 2010].
[10] www.3dmd.com/3dmdface.html [9 June 2010].
[11] B. K. P. Horn, Shape from shading: A method for obtaining the shape of a smooth opaque object from one view, Ph.D. thesis, MIT (1970).
[12] P. N. Belhumeur, D. J. Kriegman, A. L. Yuille, The bas-relief ambiguity, Intl. J. Comp. Vis. 35 (1999) 33–44.
[13] W. Smith, E. Hancock, Facial shape-from-shading and recognition using principal geodesic analysis and robust statistics, Intl. J. Comp. Vis. 76 (2008) 71–91.
[14] R. J. Woodham, Photometric method for determining surface orientation from multiple images, Opt. Eng. 19 (1) (1980) 139–144.
[15] E. N. Coleman, R. Jain, Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry, Comput. Vision and Image Proc. (1982) 309–328.
[16] F. Solomon, K. Ikeuchi, Extracting the shape and roughness of specular lobe objects using four light photometric stereo, IEEE Trans. Patt. Anal. Mach. Intell. 18 (1996) 449–454.

[17] S. Barsky, M. Petrou, The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows, IEEE Trans. Patt. Anal. Mach. Intell. 25 (2003) 1239–1252.
[18] A. S. Georghiades, Recovering 3-D shape and reflectance from a small number of photographs, Leuven, Belgium, 2003, pp. 230–240.
[19] J. Sun, M. Smith, L. Smith, S. Midha, J. Bamber, Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities, Image Vision Comput. 25 (2007) 1050–1057.
[20] C. Hernández, G. Vogiatzis, R. Cipolla, Shadows in three-source photometric stereo, in: Proc. ECCV, 2008, pp. 290–303.
[21] W. Zhao, R. Chellappa, Face processing: advanced modeling and methods, Academic Press, 2006.
[22] S. C. Kee, K. M. Lee, S. U. Lee, Illumination invariant face recognition using photometric stereo, IEICE T. Inf. Syst. E Ser. D 83 (7) (2000) 1466–1474.
[23] S. K. Zhou, R. Chellappa, D. W. Jacobs, Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints, Proc. ECCV (2004) 588–601.
[24] S. K. Zhou, G. Aggarwal, R. Chellappa, D. W. Jacobs, Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition, IEEE Trans. Patt. Anal. Mach. Intell. 29 (2) (2007) 230–245.
[25] A. S. Georghiades, P. N. Belhumeur, D. J. Kriegman, From few to many: Illumination cone models for face recognition under variable lighting and pose, IEEE Trans. Patt. Anal. Mach. Intell. (2001) 643–660.
[26] R. Lienhart, J. Maydt, An extended set of Haar-like features for rapid object detection, in: IEEE ICIP, 2002, pp. 900–903.
[27] D. A. Forsyth, J. Ponce, Computer Vision, A Modern Approach, Prentice Hall, Upper Saddle River, NJ, 2003.
[28] R. T. Frankot, R. Chellappa, A method for enforcing integrability in shape from shading algorithms, IEEE Trans. Patt. Anal. Mach. Intell. 10 (1988) 439–451.
[29] P. J. Besl, H. D. McKay, A method for registration of 3-D shapes, IEEE Trans. Patt. Anal. Mach. Intell. 14 (2) (1992) 239–256.
[30] R. T. Frankot, R. Chellappa, A method for enforcing integrability in shape from shading algorithms, IEEE Trans. Patt. Anal. Mach. Intell. 10 (4) (1988) 439–451.
[31] C. Fredembach, N. Barbuscia, S. Süsstrunk, Combining visible and near-infrared images for realistic skin smoothing, in: In Proc. IS&T/SID Color Imaging Conference, 2009.
[32] M. Oren, S. K. Nayar, Generalization of the Lambertian model and implications for machine vision, Intl. J. Comp. Vis. 14 (1995) 227–251.
[33] K. Torrance, M. Sparrow, Theory for off-specular reflection from roughened surfaces, J. Opt. Soc. Am. 57 (1967) 1105–1114.
[34] C. Donner, H. W. Jensen, Light diffusion in multi-layered translucent materials, ACM Trans. Graphics 24 (2005) 1032–1039.
[35] L. Li, C. S. Ng, Rendering human skin using a multi-layer reflection model, Int'l J. Mathematics and Computers in Simulation 3 (2009) 44–53.
[36] J. Zivanov, P. Paysan, T. Vetter, Facial normal map capture using four lights - an effective and inexpensive method of capturing the fine scale detail of human faces using four point lights, in: GRAPP, 2009, pp. 13–20.