

Acquiring Height Data from a Single Image of a Face Using Local Shape Indicators

Mario Castelán^{*†} and Edwin R. Hancock

Department of Computer Science

University of York

York YO1 5DD, UK

Tel: +44 1904 432722

Fax: +44 1904 432767

mario@cs.york.ac.uk

*Supported by National Council of Science and Technology (CONACYT), Mexico, under grant No. 141485.

†Corresponding author

Abstract

This paper describes work aimed at developing a practical scheme for face analysis using shape-from-shading. Existing methods have a tendency to recover surfaces in which convex features such as the nose are imploded. This is a result of the fact that subtle changes in the elements of the field of surface normals can cause significant changes in the corresponding integrated surface. To overcome this problem, in this paper we describe a local-shape based method for imposing convexity constraints. We show how to modify the orientations in the surface gradient field using critical points on the surface and local shape indicators. The method is applied to both surface height recovery and face re-illumination. Experiments show that altering the field of surface normals so as to impose convexity results in greatly improved height reconstructions and more realistic re-illuminations.

1 Introduction

The problem of acquiring surface models of faces is an important one with potentially significant applications in biometrics, computer games and production graphics. There are many ways in which surface models can be acquired, and these include the use of range-scanners [17, 4, 5], stereoscopic cameras [28] and structured light sensors [2, 26]. However, one of the most appealing methods is to use shape-from-shading (SFS), since this is a non-invasive process which mimics the capabilities of the human vision system. Here a field of surface normals and hence the height-map of the viewed surface are reconstructed from a single image. The SFS process was identified by Marr[22] as a key process in the computation of the 2.5D sketch, and was studied in depth by Horn[16]. The topic has also been the focus of recent research in the psychophysics literature [20, 8, 9, 21]. A recent review of techniques for shape-from-shading can be found in the comparative study of Zhang et al [34].

In brief, SFS aims to solve the image irradiance equation, $E(x, y) = R(p, q, s)$, where E is the brightness value of the pixel with position (x, y) , R is a function referred to as *the reflectance map* [13]. The reflectance map uses the surface gradients $p = \frac{\partial Z(x,y)}{\partial x}$ and $q = \frac{\partial Z(x,y)}{\partial y}$ together with the light source direction vector \vec{s} to compute a brightness estimate which can be compared with the observed one using a measure of error. If the surface normal at the image location (x, y) is $\vec{N} = (p, q, -1)^T$, then under Lambertian reflectance model, the image irradiance equation becomes $E(x, y) = \vec{N} \cdot \vec{s}$. In other words, the SFS problem is the one of recovering the surface that, after interaction with the environment (illumination conditions, reflectance properties of the object, inter-reflections) produces the radiances perceived by human eyes as intensities.

In general, though, SFS is an under-constrained problem since the two degrees of freedom for surface orientation (slant and tilt), must be recovered from a single measured brightness value. Hence, it is frequently posed as that of minimizing cost functionals that capture constraints on the gradient field [14]. This is usually carried out through iterative schemes in the discrete domain. Despite sustained research activity in the field for some three decades, no SFS scheme has been demonstrated to work as accurately as the *specialty* constrained SFS cases of photometric stereo [10, 12] and statistical SFS [1].

The first of these methods requires at least three images of the same object illuminated from different view-points, while the second uses an aligned database of accurate surface information belonging to objects of the same class, i.e. faces. When it comes to the original single-image SFS problem, the recovered surfaces are seriously affected by factors such as inaccurately calculated illumination direction and departures from Lambert's law. Both of these are difficult to correct when only a single image is to hand. Moreover, the restrictions imposed by most SFS schemes on the gradient field (smoothness, irradiance, integrability, unit length) are insufficient to overcome the biases introduced by mis-estimated light source direction or non-Lambertian reflectance. Moreover, SFS can not deal with surface discontinuities. All of these factors cause departures from the simple Lambertian reflectance model, and this in turn leads to errors and biases in the calculated field of surface normals ¹. When propagated iteratively these errors can result in error in the topography of the recovered surface. For instance, because of the concave-convex ambiguity, there may be regions where the sign of the curvature becomes reversed. Moreover, since only a single image is to hand, there is no additional evidence that can be used to correct the resulting shape-errors.

For face analysis, the use of SFS has proved to be an elusive task, since the concave-convex ambiguity can result in the inversion of important features such as the nose. To overcome this problem, domain specific constraints have been used. For example, Shimshoni et al.[27] have presented a shape reconstruction method for bilaterally symmetric surfaces from a single image. They combined geometric and photometric information in order to obtain a dense correspondence between pairs of symmetric points and then perform dense shape recovery. Unknown lighting and viewing parameters are also recovered by the process. Lambertian surfaces, unknown constant albedo and weak perspective projection were assumed. Their method gave results of quality depending on the class of input images. Better results were obtained for non-frontal-view examples.

Similarly, Zhao and Chellappa [36] have also exploited the bilateral symmetry of faces. Unlike Shimshoni et al. [27] they dealt with Lambertian surfaces with unknown and varying albedo. The

¹In contrast to the human visual system [20], it seems that computer vision systems find it more difficult to estimate the tilt of a surface from a single image than its slant.

problem of source-from-shading was also addressed in the paper. Orthographic projection and frontal views were assumed. The image irradiance equation was coupled with a newly defined self-ratio image irradiance equation. This was proved to have a unique global solution. The method worked well with synthetic surfaces but did not perform well when tested on images of faces with natural albedo variation. The method has also been used for face recognition [35].

Dogvard and Basri [7] have combined the statistical constraint of Atick[1] et al. and the symmetry constraint of Zhao and Chellappa[36] into a single shape-from-shading framework. Here the aim was to express the surface gradient in terms of a set of deformation coefficients. This allows shape-from-shading to be transformed into a linear system of equations. This system can be simply solved for the shape coefficients and used to reconstruct the height function for the face. Although it uses a statistical model, the method is efficient. However, facial asymmetry can give significant error in the recovered surfaces.

Prados and Faugeras [24] have recently proposed an SFS scheme that relies on the existence of a unique critical point (i.e. the brightest value appears at only one location in the image). Their generic SFS scheme proved to work well with symmetric and non-symmetric surfaces, regardless of facial pose. For face shape recovery, when the light source is at the optical center of the camera, the method gives qualitatively good facial reconstructions.

There are a number of interesting differences between these methods. Only Prados and Faugeras [24] have proposed a generic approach and make SFS a well-posed problem. Zhao and Chellappa [36] and Shimshoni et al [27], on the other hand, have constrained the problem to the domain of faces. The success of the different methods will depend on the correct setting of illumination, pose and albedo. The quality of the facial reconstructions may lack the accuracy demanded by model-based recognition systems. Nonetheless, recognition can also be realized using alternative features such as local shape [32] and the quality of this information can be considerably increased by improved SFS methods.

It is clear that a precise height map is difficult to obtain by integration of the field of surface normals delivered by traditional SFS (i.e., excluding photometric stereo [12] and statistical SFS [1]) due to local errors in direction and potential reversal in the sign of the curvature. However, this does not

imply that the entire field of surface normals is in error. In fact, some regions on the image do provide surface normal information that is sufficiently faithful for qualitatively good surface reconstruction. The problem originates from those locations where the combined effects of image noise, and overreliance on the consistency constraints, results in effects such as inversion of the surface convexity or concavity. As noted previously, when applied to the problem of face reconstruction, the effect can be to cause high curvature features such as the nose to become imploded with respect to the remainder of the surface. As we will demonstrate in this paper, if such regions can be identified and the surface normal directions corrected, then the result is improved overall surface topography.

The outline of this paper is as follows. In Section 2 we commence by providing a brief overview of the SFS method utilized to obtain a first estimate of the surface normal directions. In Section 3, we review Frankot and Chellappa’s global surface integration method which we use to recover facial height maps. Section 4 introduces the local shape indicators which we use to characterize the surface topography. In Section 5, we describe our method for reassigning the surface gradient orientations. We provide experiments to evaluate the method on human faces in Section 6. Finally, in Section 7 we present some conclusions and identify directions for future work.

2 Geometric approach for SFS

Worthington and Hancock [31] have developed an SFS method in which the image irradiance equation is treated as a hard constraint by demanding that the recovered surface normals to lie on the reflectance cone. Suppose that \hat{N}_k is a smoothed² surface normals at step k of the algorithm, then the update equation for the surface normal directions is

$$\hat{N}_{k+1} = \Theta \hat{N}_k \quad (1)$$

where Θ is a rotation matrix computed from the apex angle α and the angle between the current smoothed estimate of the surface normal direction \hat{N}_k and the light source direction. To restore the

²For further details about the suggested method for smoothing the normal field, see [30]

surface normal to the irradiance cone, it must be rotated by an angle

$$\theta = \cos^{-1}(E) - \cos^{-1} \left(\frac{\hat{N}_k \cdot s}{\|\hat{N}_k\| \cdot \|s\|} \right) \quad (2)$$

about the axis $(u, v, w)^T = \tilde{N}_k \times s$. Hence, the rotation matrix is

$$\Theta = \begin{pmatrix} c + u^2c' & -ws + uv c' & vs + uw c' \\ ws + uv c' & c + v^2c' & -us + vw c' \\ -vs + uw c' & us + vw c' & c + w^2c' \end{pmatrix} \quad (3)$$

where $c = \cos(\theta)$, $c' = 1 - c$ and $s = \sin(\theta)$.

The needle maps delivered by this geometric framework have proved to be useful for computing topographic feature characteristics than can be used for 3D object recognition from 2D brightness images [32]. In this paper, we use the method for face-analysis since it has been demonstrated to recover a field of surface normals (or needle-map) that preserves fine topographic detail. The output of the algorithm is used as an initial estimate of the field of surface normals. We show how to use shape-indicators computed from this field of surface normals to correct for convex-concave surface inversions, and hence to improve the quality of the recovered height-map.

3 Integrating a gradient field

Suppose that $Z(x, y)$ denotes the reconstructed height at the image location with co-ordinates (x, y) . The integrability condition for the surface demands that the height function does not depend on the integration path. This in turn means that the surface must satisfy the condition $Z_{xy} = Z_{yx}$ on the second-order mixed partial derivatives of the height function, where the Z_{xy} denotes the partial derivative $\frac{\partial^2 Z(x, y)}{\partial x \partial y}$ and Z_{yx} the partial derivative $\frac{\partial^2 Z(x, y)}{\partial y \partial x}$. This condition can also be regarded as a smoothness constraint, since the partial derivatives of the surface need to be continuous in order that they can be integrable.

There are a number of ways in which a surface may be recovered from a field of surface normals. One approach is to use trigonometry to increment the height function, locally, along a path or a front [33, 6, 25]. On the other hand, by minimizing error functionals the integration process can be carried out globally [15, 11]. An interesting analysis concerning gradient fields integration methods is given in [18].

In [11] Frankot and Chellapa proposed a method to project a gradient field to the nearest integrable solution. They suggested to use a set of integrable basis functions to represent the surface slopes so as to minimize the distance between an ideally integrable gradient field and a non integrable one.

Following [11], if the surface Z is given by

$$\tilde{Z}(x, y) = \sum_{\omega \in \Omega} \tilde{C}(\omega) \phi(x, y, \omega) \quad (4)$$

where ω is a two dimensional index belonging to a domain Ω , and $\phi(x, y, \omega)$ is a set of basis functions which are not necessarily mutually orthogonal. The partial derivatives of \tilde{Z} can also be expressed in terms of this set of basis functions using the formulae

$$\tilde{Z}_x(x, y) = \sum_{\omega \in \Omega} \tilde{C}(\omega) \phi_x(x, y, \omega) \quad \text{and} \quad \tilde{Z}_y(x, y) = \sum_{\omega \in \Omega} \tilde{C}(\omega) \phi_y(x, y, \omega) \quad (5)$$

Given that $\phi_x(x, y, \omega)$ and $\phi_y(x, y, \omega)$ are integrable, then so are the mixed partial derivatives of $\tilde{Z}(x, y)$.

In the same way, the possibly non integrable gradient field (which, indeed, is the only information we have) can be represented as

$$\hat{Z}_x(x, y) = \sum_{\omega \in \Omega} \hat{C}_1(\omega) \phi_x(x, y, \omega) \quad \text{and} \quad \hat{Z}_y(x, y) = \sum_{\omega \in \Omega} \hat{C}_2(\omega) \phi_y(x, y, \omega) \quad (6)$$

Note that, as $\hat{C}_1 \neq \hat{C}_2$, then $\hat{Z}_{xy} \neq \hat{Z}_{yx}$.

The goal then is to find the set of coefficients that minimize the quantity

$$d\{(\hat{Z}_x, \hat{Z}_y), (\tilde{Z}_x, \tilde{Z}_y)\} = \int \int \left\| \tilde{Z}_x - \hat{Z}_x \right\|^2 + \left\| \tilde{Z}_y - \hat{Z}_y \right\|^2 dx dy \quad (7)$$

As Frankot and Chellapa proved, the set of coefficients $\tilde{C}(\omega)$ minimizing the distance given by the above equation is

$$\tilde{C}(\omega) = \frac{P_x(\omega)\hat{C}_1(\omega) + P_y(\omega)\hat{C}_2(\omega)}{P_x(\omega) + P_y(\omega)}, \quad (8)$$

where $P_x(\omega)$ and $P_y(\omega)$ are $\int \int \|\phi_x(x, y, \omega)\|^2 dx dy$ and $\int \int \|\phi_y(x, y, \omega)\|^2 dx dy$ respectively.

If $\phi(x, y, \omega)$ is assumed to be the set of Fourier basis functions $\exp(j\omega_x x + j\omega_y y)$, with $\Omega = (2\pi n, 2\pi m)$, where $n \in \{0, 1, \dots, N-1\}$ and $m \in \{0, 1, \dots, M-1\}$ for an $N \times M$ image, then $P_x = \omega_x^2$, $P_y = \omega_y^2$, $\hat{C}_1(\omega) = \hat{C}_x(\omega)/j\omega_x$, and $\hat{C}_2(\omega) = \hat{C}_y(\omega)/j\omega_y$. Therefore, (8) is represented in the Fourier domain by

$$\tilde{C}(\omega) = \frac{-j\omega_x \hat{C}_x(\omega) - j\omega_y \hat{C}_y(\omega)}{\omega_x^2 + \omega_y^2}. \quad (9)$$

In this manner, by projecting the set of coefficients $\tilde{C}(\omega)$ back to the spatial domain, a height map corresponding to the nearest integrable surface $\tilde{Z}(x, y)$ can be obtained from the input gradient field.

In this paper we use the global integration approach explained above to generate the height maps from the gradient fields of faces.

4 Local shape indicators

Curvature-based information has been widely used in shape analysis, especially for surface segmentation and 3D object recognition. A local shape indicator is a scalar that conveys information concerning the local topography of a surface using its principal curvatures. The principal curvatures may be estimated using the surface normal directions to compute the Hessian matrix, which can be deduced from local changes in the surface normal directions. However, a more robust method is to fit a local quadric surface patch

$$P(u, v) = \frac{1}{2}(\kappa_1 u^2, \kappa_2 v^2)$$

where κ_1 and κ_2 are the principal curvatures (the calculated eigenvalues of the Hessian matrix) with directions u and v respectively.

Local shape indicators are usually coupled. For instance, the HK classification [3] uses the Gaussian and mean curvatures

$$H = \frac{(\kappa_1 + \kappa_2)}{2} \text{ and } K = \kappa_1 * \kappa_2$$

respectively. By distinguishing between the cases in which H and K are individually negative, zero or positive, it is possible on the basis of the joint behavior to assign topographic labels to points on a surface. A different and slightly more convenient set of attributes is the curvedness/shape-index representation developed by Koenderink and Von Doorn [19]. Here the principal curvatures are used to compute the shape index

$$S = -\frac{2}{\pi} * \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right)$$

(for $\kappa_1 \geq \kappa_2$), and the curvedness,

$$C = \sqrt{\kappa_1^2 + \kappa_2^2}.$$

The shape index is an angular variable that relates to the local surface topography. It varies continuously from -1 to $+1$ as the surface changes through cup, rut, saddle-rut, saddle, saddle-ridge, ridge and dome, and cup again. The curvedness relates to the degree of curvature of the surface.

The curvedness is a convenient indicator of potential surface discontinuity. The reason for this is the higher the curvedness, the more likely the presence of a rapid variation in height. A surface with low curvedness corresponds to a highly continuous one, while a complicated surface will give rise to high curvedness at many locations. For instance, in the case of faces, the curvedness is large for features such as the boundaries of the nose, mouth and eyes.

For our experiments, we utilize the local descriptors of shape-index and curvedness to characterize the regions on a gradient field where a change of orientation should be performed to enforce convexity.

5 Using local shape indicators to redirect SFS gradient fields

Inevitably, any surface gradient field delivered by SFS will be inaccurate due to noise or albedo changes, and these in turn cause variations in the intensities of the input image. SFS works well for objects that are uniformly concave or convex. However, if the object under study is more complex, with both concave and convex regions, then SFS can fail. In these situations although the recovered surface normal direction is consistent with the measured image brightness, the recovered surface does not reflect the structure of the object under study. In particular, there may be inversions of the sign of the surface curvature with convex regions appearing concave and vice-versa. However, in the case of faces (and many other objects) the surface under study is largely convex.

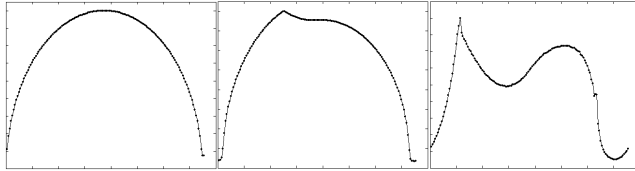


Figure 1: The figure shows the results of applying the method to the derivatives of a sphere with radius 75 units. From left to right, we show transverse plots of the recovered surface height in the direction of the x axis. The global height maximum coordinates (a, b) are both set to 75, 50 and 20 respectively, and τ_x and τ_y are set to zero..

Based on this above observation, in this paper we present a method for enforcing the convexity of the integrated surface. We use the location of the global height maximum on the surface to enforce the condition.

Formally stated, suppose that S is a smooth surface immersed in \mathbb{R}^3 . Let p be a critical point of S and U_p a neighborhood of p . Suppose that S is locally concave over U_p . Then, the new surface \tilde{S} constructed from S by reversing the sign of all its partial derivatives, S_x and S_y , is locally convex within U_p . Moreover, a local maximum on \tilde{S} will be located at that point where the function ceases increasing

and starts decreasing³. Suppose that all the partial derivatives of \tilde{S} with respect to x , \tilde{S}_x , have a negative sign before reaching the position of the critical point p along the x axis and have a positive sign after reaching it. Suppose also that the same occurs for \tilde{S}_y . Under these conditions, then the critical point p on U_p will be the position of the global maximum⁴ of \tilde{S} .

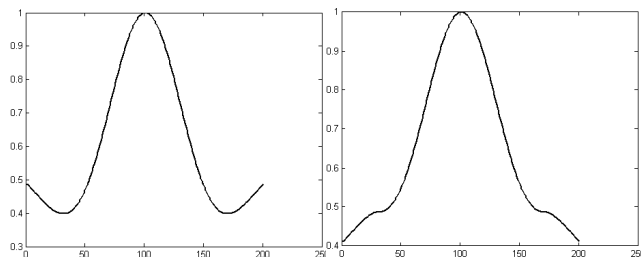


Figure 2: Applying the method to the Mexican hat function. From left to right we show the original surface and recovered surface after applying the method taking as the global height maximum the center of the surface with $\tau_x = \tau_y = 0$. Note how the concave parts of the Mexican hat become convex.

The basic idea underlying this paper is to enforce the condition that the integrated surface has a global height maximum. For the face analysis problem, we select this point to be at the tip of the nose. By choosing such a point we can divide the surface into positively and negatively signed areas of the needle map. To enforce this condition we apply the simple rule:

$$\check{Z}_x(x, y) = \begin{cases} abs(\hat{Z}_x(x, y)) & \text{if } x \leq a \text{ and } C(x, y) \geq \tau_x \\ -abs(\hat{Z}_x(x, y)) & \text{if } x > a \text{ and } C(x, y) \geq \tau_x \\ \hat{Z}_x(x, y) & \text{otherwise} \end{cases}$$

³Of course, \tilde{S} will present many local maxima for a face-like surface.

⁴It might be a maximum or a minimum depending on the integration method.

$$\check{Z}_y(x, y) = \begin{cases} abs(\hat{Z}_y(x, y)) & \text{if } y \leq b \text{ and } C(x, y) \geq \tau_y \\ abs(\hat{Z}_y(x, y)) & \text{if } y > b \text{ and } C(x, y) \geq \tau_y \\ \hat{Z}_y(x, y) & \text{otherwise} \end{cases}$$

where \check{Z}_x and \check{Z}_y are the updated gradients, and \hat{Z}_x and \hat{Z}_y are the original gradients. The desired global height maximum is located at the point with co-ordinates (a, b) (the tip of the nose) and this is assigned manually. To decide whether the element of the gradient field at the location (x, y) will be altered or not, we compare the curvedness indicator, $C(x, y)$ to the thresholds τ_x and τ_y . Similarly, the shape-index (SI) can also be utilized to redirect the surface normals, with the rule:

$$\check{Z}_x(x, y) = \begin{cases} abs(\hat{Z}_x(x, y)) & \text{if } x \leq a \text{ and } SI(x, y) \leq \tau_x \\ -abs(\hat{Z}_x(x, y)) & \text{if } x > a \text{ and } SI(x, y) \leq \tau_x \\ \hat{Z}_x(x, y) & \text{otherwise} \end{cases}$$

$$\check{Z}_y(x, y) = \begin{cases} abs(\hat{Z}_y(x, y)) & \text{if } y \leq b \text{ and } SI(x, y) \leq \tau_y \\ abs(\hat{Z}_y(x, y)) & \text{if } y > b \text{ and } SI(x, y) \leq \tau_y \\ \hat{Z}_y(x, y) & \text{otherwise} \end{cases}$$

In this case, $SI(x, y)$ is the shape-index which is compared to the thresholds τ_x and τ_y for deciding whether the element of the gradient field at the location (x, y) will be altered or not.

The following diagrams show the behavior of the method when applied to the derivatives of a sphere and a Mexican hat. Both experiments were realized using the curvedness indicator.

To illustrate the global height maximum enforcement procedure, Figure 1 shows the results of applying the method to the derivatives of a sphere with radius 75 units. From left to right, we show transverse plots of the recovered surface height in the direction of the x axis. The global height maximum coordinates (a, b) are both set to 75, 50 and 20 respectively, and τ_x and τ_y are set to zero.

The effect on the convexity of the surface is clearer in Figure 2, where the method is applied to the Mexican hat function. Transverse sections of the recovered surface are shown, and from left to right they show the original surface and recovered surface after applying the method taking as the global height maximum the center of the surface with $\tau_x = \tau_y = 0$. Note how the concave parts of the Mexican hat become convex.

It is evident that the peak-enforcement procedure will segment the recovered surface into four quadrants. As a result the curvedness of the recovered surface will be reduced. This is not desirable for surface height recovery from a face. The net effect will be to make the surface structure pyramidal. This problem is overcome by using the thresholding procedure to either force the normals to change direction, or to allow them to remain unchanged. This procedure is rather heuristic, and different thresholds apply to different images of faces. Although we have discovered that a fluctuation between 0.2 and 0.3 is generally successful for most of the cases, it is recommended to commence with the hardest case (all the derivatives change) and gradually modify the thresholds until the best shape is generated, as suggested in Figures 4 and 6.

It is worth commenting that the signs of a needle map are not the only attributes which can alter the resulting integrated surface. Modifying the influence of the z component of the surface normal leads to an alteration which affects the gradient and therefore forces some regions of the height map to be either flatter or more curved after performing the global integration.

Note that we propose a directional correction which can be used in conjunction with already established SFS methods. We focus on face shape recovery for the frontal pose (since the critical point to consider is the tip of the nose). However, other assumptions such as light source direction, albedo, boundary conditions and inter-reflections are incorporated by the SFS method used to compute the derivatives. In this paper we use the geometric framework[31] outlined in Section 2. Although this

framework assumes constant albedo, we test our experiments with variable albedo images of faces.

In the following section some experiments will be presented in order to illustrate these points on an application involving face reconstruction using SFS.

6 Experiments

This section is organized into two parts. We commence by describing experiments focussed on height recovery, and then proceed to describe results obtained by re-illuminating the recovered surfaces.

In first part of the study, which focusses on height recovery, we show examples of the effect of the gradient re-direction process on the recovered surfaces. We also present an analysis of the errors produced by the method on 50 synthetic test images from the MaxPlank database. Results of using the method over real world images are also presented.



Figure 3: Face for analysis.

In the second part of the study, which focusses on surface re-illumination, we show the results of using the recovered surface gradients to synthesize new face images. Here, we investigate the effect of moving the light source direction. We compare the results with ground-truth.

The face database used for our experiments was provided by the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany. As described in [29], this database was constructed using a Cyberware laser scanner. The range images are of the heads of young adults represented in a cylindrical coordinate system. We have converted the cylindrical coordinates into cartesian coordinates and recov-

ered the associated height values. We were also provided with synthetic textures corresponding to each face and these were used to render the range images. We used frontal pose of the recovered surface illuminated by a light source at infinity and parallel to the viewer direction. We also utilized real world face images to complement our experiments. We refer to images from the Max-Planck database unless otherwise stated.

6.1 Height recovery

The first series of tests were carried out on the image shown in Figure 3, and serves as an illustration of the method described in this paper. To compute the surface gradients from the raw image brightness we followed the procedure described in Section 2. As noted previously, this construction ensures that the image irradiance equation is satisfied as a hard constraint. For the surface integration step we used the global method proposed by Frankot and Chellappa [11] and discussed in Section 4. This method recovers surface height using the inverse Fourier transform of the field of surface normal directions.

In Figure 4, we illustrate the effects of using our method for redirecting the field of surface normals. The first column shows the needle map. The second and third columns show the result of re-illuminating the surface normals with light-source directions $(1, 0, 0)^T$ and $(0, -1, 0)^T$. The top row shows the result obtained with the initial field of surface normals. Notice how the re-illuminations suggest that the nose is imploded, since the region surrounding it is shadowed. The subsequent rows show the recovered surface illuminations after applying the method with $\tau_x = \tau_y = 0$, $\tau_x = \tau_y = 0.2$ and finally $\tau_x = 0.3$ and $\tau_y = 0.4$, respectively. For the hardest case (second row), where no data is filtered by the thresholds, the four-quadrant effect is very marked. However, the effect diminishes when the thresholds are increased. The harder the threshold, the stronger the influence of the face features. It is important to note the differences between the needle maps appearing in the top and bottom rows of the figure. The change of signs in the surface normals suggests that the new surface normals relate more strongly to the underlying shape of the face. This is a consequence of inverting the originally concave regions (the ones around the nose, eyes and mouth) to become convex ones, causing the imploded facial areas

to “pop back” so that a better face shape can be recovered.

The correction of the directions of the normals is clearer in Figure 5. In the left-hand column we show the result obtained using the un-modified needle map, while the right-hand column shows the result obtained with the modified field of surface normals with $\tau_x = 0.3$ and $\tau_y = 0.4$. The top row shows the re-illuminations obtained when the light source direction is $(0, 0, 1)^T$. The bottom row shows the recovered height-maps. There are several features that deserve comment. For instance, the intensities around the nose produce the appearance of a convex surface, making it appear more natural.

We take this analysis one step further in Figure 6 where from left to right we show the curvedness map, together with profile (side) and top-down views of the surface wire-frames. In the top row we show the original case. Here the implosion of the nose is clear. The subsequent rows are for $\tau_x = \tau_y = 0$, $\tau_x = \tau_y = 0.2$ and finally $\tau_x = 0.3$ and $\tau_y = 0.4$, respectively. Note in the top row how the curvedness is increased, as well as the poor quality of the recovered surface in terms resemblance to a face. The curvedness is of course minimized when no threshold is taken into account (second row), but the four-quadrant effect again becomes obvious. By incrementing τ_x and τ_y the fine features of the face seem to be preserved and the overall structure of the face is still sound. By choosing an appropriate threshold we are able to enhance the salient features of a particular face while maintaining the overall face composition.

As far as the reconstruction using shape-index is concerned, the height data recovered is very similar to that obtained using curvedness. Figure 7 shows a scatter plot comparison between the original and modified shape-indexes⁵: the x -axis corresponds to the original case while the y -axis corresponds to the $\tau_x = \tau_y = 0.4$ case. This diagram presents only those pixels in the original gradient field with a shape index lower or equal to 0.4 (x -axis), therefore we can analyze the new value of such pixels in the redirected field of surface normals (y -axis). Note how the majority of the points are distributed above the line $x = y$, which shows how the original shape index turned into a greater one, suggesting that the concave regions changed to convex. Such pixels belong to the regions surrounding the nose, mouth and

⁵The shape-index scale was normalized from 0 to 1.

eyes. The small cloud of points below the line, representing the pixels where the shape index remained lower than the threshold is mainly related to the pixels surrounding the face boundary.

We found that for most experiments good qualitative results were obtained with τ_x and τ_y varying between 0.3 and 0.4. A further analysis is shown by Figure 8. Here we compare with ground truth the recovered height maps obtained with (right column) and without (left column) redirection of the normal field. The top row figure shows the recovered height maps superimposed on the image from which they were generated. Note how the imploded features (left image) of the face, such as nose and mouth, become visible after the method is applied (right image). In the bottom row we show longitudinal and transverse sections of the recovered height maps. The left-hand diagram corresponds to the longitudinal section along the x-axis and the right-hand diagram shows the transverse section along the y-axis. In both plots, the thick solid line represents the ground-truth surface, the dotted line illustrates the recovered surface using our method, and the dashed line shows the height data integrated using the original field of surface normals. There is a clear improvement of the recovered height map once the re-direction operation is applied to the surface normals. Although some errors still remain, the rectification of the implosion problem is sufficient to reveal plausible facial shape.

An iso-contour comparison is shown in Figure 9. Here, regions in the images are color coded according to height. Discarding the background, the darkest regions correspond to highest values. The leftmost panel corresponds to the ground-truth surface, the middle panel presents the case when our method was applied, while the rightmost figure represents the recovered surface without gradient redirection. Note the similarities between the ground-truth and the recovered surface after changing the gradient orientations. It is evident that, in the height maps generated by both the ground-truth and re-directed gradients, the peak region is located in the area of the nose, with some discrepancies in the mouth and chin areas though. The un-modified gradient generated height contains equal peak regions over the entire image, i.e., cheeks, chin, front and a tiny one on the tip of the nose.

The images shown in Figure 11 provide an absolute height difference analysis. Here we investigate the effect of attempting height recovery using profile rather than frontal images. The left-most column shows the raw image. The middle and right-most columns show superimposed plots of the recovered

height maps after applying our method on the profile views (i.e. side) of the images from which the original gradient fields were calculated. From the superimposed surfaces, it is clear that the major differences in the recovered height maps and the ground-truth surfaces are near the nose, the cheeks and the mouth area.

To provide a detailed analysis of our method, a more exhaustive set of tests was carried out on fifty images of faces from the database. The average percentage of height difference comparison plot⁶ is shown in Figure 12. We calculated the percentage of height difference error $\|Ground_truth - Recovered_surface\|/Ground_truth$ as an average over all points of the 50 surfaces. The diagram shows a scatter plot comparing the average percentage of height difference from the original gradient integrated surface (x -axis) against the redirected gradient integrated surface (y -axis). Observe how when the original gradient is used, the error is concentrated between 8% and 12%. The error, however, when the redirected gradient is used, is concentrated between 2% and 4%.

Figure 13 illustrates how pre-processing operations might improve the recovery of height maps from images. The left-most panel shows the calculated height map of the image of the face in Figure 16 after removing the eyes and eyebrows by setting their pixel brightness values to be that of the average over the skin. The right-most image shows the calculated height map on the un-edited face image. Note how this pre-processing step generates more accurate height maps in the eye-region when our re-direction method is used.

However, some regions present a higher degree of error for the new surfaces. This is illustrated by the analysis given by Figure 14. The first column shows the longitudinal sections along the face for the recovered surfaces both before (dotted line) and after (dashed line) redirection of surface normals. The recovered height map after integration of the ground-truth normals is shown as a solid line. The second and third columns respectively show the iso-contour representations of the absolute height differences before and after re-directing the surface normals. From these pictures it is clear that the absolute height

⁶For these experiments, the ground-truth surface was generated by integrating the known ground-truth gradient from each image, using the Frankot and Chellappa method. This was done so that all the surfaces were generated on the same basis for comparison purposes. This reduces the biases involved in the integration method.

difference is reduced after changing the direction of the surface normals. However, there are still errors and these occur mainly in the proximity of the chin and eyes areas. These problems can be attributed to changes of albedo (eye area) and instabilities produced by the boundaries of the chin and neck . It is also interesting to note how the separation between the lips tends to disappear after redirection. This can be explained as a consequence of erroneously enforcing convexity.

Finally, Figure 10 presents some wire-frame views of the recovered surface obtained from an image of a smiling face (leftmost). Here we have used a real world (non-synthetic) image which was taken with a digital camera. This is a challenging example since the face is in a non-frontal pose. The overall structure of the face was well recovered, however there is again some error in the area of the chin and eyes.

6.2 Re-illumination

Dealing with variation in illumination direction is a topic of central concern in face recognition. The reason for this is that light-source effects are responsible for more variability in the appearance of face images than changes in identity [23]. In this section we investigate how the surfaces recovered using our method can be used for synthesizing new facial images under different lighting conditions. Here we use a simple Lambertian re-illumination model using albedo maps derived from the input images.

Figure 15 presents re-illumination experiments for two example faces. The first column corresponds to the input image. The remaining columns show the generated re-illuminations after applying our method. For the second and third columns, the light source vector is nearly parallel to the x axis in both negative and positive directions, while nearly parallel to the y axis for the fourth and fifth columns.

A more exhaustive analysis is shown in Figure 16. For both sets of images, the top row represents the re-illumination results obtained using the ground-truth normals, the middle row shows those obtained using the unaltered gradient field and the bottom row those obtained using the re-directed surface normals. From left to right, the light source direction makes an angle of -45 , -25 , 25 and $+45$ degrees to the image normal in the horizontal (x) direction for the upper set of images. In the lower set

of images, the light source is moved in the same manner in the vertical (y) direction. It is interesting to note the similarities between the ground-truth and re-directed gradient re-illuminations. Although the recovered surface does not accurately represent the shape of the image from which it was acquired, the overall shape is sufficiently accurate to create realistic re-illuminations provided that the light source is not moved by more than 45° . The results are best when the light source is moved in the horizontal direction. This is a consequence of the vertical symmetry of human faces. On the other hand, the re-illumination results for the un-modified gradient fields show artifacts of implosion in the area around the nose and mouth. This becomes more severe when the light source moves further away from the viewer direction.

A second analysis is shown in figure 17, where an unprocessed image of a face (single image in between two rows) was used for the experiments. This image is a real world one, taken by a digital camera. The first row presents the results for the modified gradient field while second shows those obtained with the original gradient field. Note how in the top row the quality of the re-illuminations is improved. This contrasts with the imploded features presented in the second row. It is important to note, however, that errors appear in some areas of the face, i.e. those surrounding the mouth. This can be explained as the consequence of the change in reflectance properties on the lips. Of course, the eyes and the mustache area also present different reflectance properties. This suggests that more attention should be paid to the reflectance model used for re-illumination.

7 Conclusions

We have presented a method for correcting a gradient field of a face. The aims in doing this are twofold. First, we wish to generate a height map with a global maximum located at a critical point located at the tip of the nose. Second, we aim to force the recovered surface to be convex in accordance with evidence provided by local shape indicators. We have proved that the simple idea of modifying the surface normal directions so as to restore the convexity of imploded features using the constraints derived from the location of a point of global maximum height seems to work well with the recovery

of face surfaces. After integration, the recovered shape preserves most of the salient facial features, including the nose lips and eye-sockets. As the accuracy of the reconstruction will depend on the gradient data, pre-processing steps for correcting intensities (i.e. removing specularities and areas of albedo variation such as the eye areas) would probably improve the quality of the results, if a more accurate height map is required.

Although the method reported here is effective in correcting feature implosion when the surface gradient is computed using a geometric SFS approach, there clearly remains scope for further improvement in computing accurate surface gradients. The improved height recovery from single images of faces also suggests the possibility of generating coarse extrapolation for rendering novel views with few degrees of rotation. Considering that our only input information is a single frontal image of a face, the raw height maps resulting from our algorithm can help as a starting point for other refinements aimed at generating more precise information. Of course, the gradient maximum constraint is natural for faces, since the tip of the nose is a global height maximum. However, the constraint could be used for more general surfaces in a local manner for surface height recovery and where there are local regions of implosion.

References

- [1] Atick, J., Griffin, P. and Redlich, N. (1996), Statistical Approach to Shape from Shading: Reconstruction of Three-Dimensional Face Surfaces from Single Two-Dimensional Images, *Neural Computation*, Vol. 8, pp. 1321-1340.
- [2] Battle, J., Mouaddib, E. and Salvi, J. (1998) Recent Progress in Coded Structured Light as a Technique to Solve the Correspondence Problem: a Survey, *Pattern Recognition*, Vol. 31, No. 7, pp. 963-982.
- [3] Besl, P.J. and Jain, R.C. (1986) Invariant surface characteristics for 3-d object recognition in range images, *Comput. Vision Graphics Image Proc.*, Vol. 33, pp. 33 - 80.
- [4] Besl, P. (1989) Active Optical Range Imaging Sensors, *Advances in Machine Vision*, Chapter 1, pp. 1-63.

- [5] Blais, F. (2003), A review of 20 years of range sensor development, *Videometrics VII, SPIE Proc.*, Vol. 5013, pp. 62-76
- [6] Bors, A.G., Hancock, E.R. and Wilson, R.C. (2003) Terrain Analysis Using Radar Shape-from-Shading, *IEEE Trans, on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5.
- [7] Dovgand, R. and Basri, R. (2004), Statistical symmetric shape from shading for 3D structure recovery of faces, *European Conf. on Computer Vision (ECCV 04)*, Prague, May 2004.
- [8] Erens, R.G.F., Kappers, A.M.L. and Koenderink, J.J. (1993), Perception of Local Shape from Shading. *Perception and Psychophysics*, Vol. 54, No. 2, pp. 145 - 156.
- [9] Erens, R.G.F., Kappers, A.M.L. and Koenderink, J.J. (1993), Estimating Local Shape from Shading in the Presence of Global Shading. *Perception and Psychophysics*, Vol. 54, No. 3, pp. 334 - 342.
- [10] Forsythe, D. and Ponce, J. (2001), *Computer Vision: a Modern Approach*, Prentice-Hall.
- [11] Frankot, R.T. and Chellapa, R. (1988), A Method for Enforcing Integrability in Shape from Shading Algorithms, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, pp. 438 - 451.
- [12] Georghiades, A., Belhumeur, D. and Kriegman, D. (2001), From Few to Many: Illumination Cone Models fro Face Recognition under Variable Lighting and Pose, *IEEE Trans. Pattern Analysis and Machine Intelligence*, pp. 643 - 660.
- [13] Horn, B.K.P. (1977), Understanding Image Intensities, *Artificial Intelligence*, Vol. 8, pp. 201-231.
- [14] Horn, B.K.P. and Brooks, M.J. (1986), The Variational approach to Shape-from-Shading, *Computer Vision Graphics Image Processing*, Vol. 33. pp 174-208.
- [15] Horn, B.K.P. (1989), Height and Gradient from Shading, *International Journal of Computer Vision*, Vol. 5, No. 1, pp. 37-75.
- [16] Horn, B.K.P. and Brooks, M.J., (1989), *Shape from Shading*, MIT Press, Cambridge, MA.

- [17] Jarvis, R., (1983), A Perspective on Range Finding Techniques for Computer Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 5, No. 2, pp. 122-139.
- [18] Klette, R. and Schluns, K. (1996), Height Data from Gradient Fields, *Proc. Machine Vision Applications, Architectures, and Systems Integration*. Vol. 5, SPIE 2908, pp. 204-215.
- [19] Koenderink, J.J., and Van Doorn, A.J. (1992), Surface Shape and Curvature Scales, *Image en Vision Computing*, Vol. 10, pp. 557-565.
- [20] Koenderink, J.J., and Van Doorn, A.J. (1992), Surface Perception in Pictures, *Perception and Psychophysics*, Vol. 52, No. 5, pp. 487 - 496.
- [21] Koenderink, J.J., Van Doorn, A.J., Christou, C.G. and Lappin, J.S. (1996), Perturbation Study of Shading in Pictures, *Perception*, Vol. 25, No. 9, pp. 1009 - 1026.
- [22] Marr, D. (1982), *Vision*, W.H. Freeman & Co., San Francisco.
- [23] Moses, Y., Adini, Y. and Ullman, S. (1994), Face Recognition: the Problem of Compensating for Changes in Illumination Direction, *European Conference in Computer Vision*, pp. 286-296.
- [24] Prados, E. and Faugeras, O. (2004), A rigorous and realistic Shape From Shading method and some of its applications, *INRIA Research Report, RR-5133*, March 2004.
- [25] Robles-Kelly, A. and Hancock, E. R. (2004) A Graph-Spectral Approach to Shape-from-shading, *IEEE Transactions on Image Processing*, Vol. 13(7), pages 912-926, 2004.
- [26] Scharstein, D. and Szeliski, R. (2003) High-accuracy stereo depth maps using structured light, *IEEE Proceedings on International Conference on Computer Vision and Pattern Recognition*.
- [27] Shimshoni, Mose, Y. and Lindenbaum, M. (2003) Shape Reconstruction of 3D Bilaterally Symmetric Surfaces, *IJCV*2:1-15.
- [28] Starks, M. (1995) Stereoscopic imaging technology: Review of patents and literature, *Int Jour Virtual Reality*. Vol. 1, No. 2, pp. 2-25

- [29] Blanz, V. and Vetter, T. (1999), A Morphable model for the synthesis of 3D Faces, *Proceedings of SIG-GRAPH '99*, pp. 187 - 194.
- [30] Worthington, P. L. and Hancock, E. R. (1998), Needle Map Recovery Using Robust Regularizers, *Image and Vision Computing*, Vol. 17, No. 8, pp. 545-559.
- [31] Worthington, P. L. and Hancock, E. R. (1999), New Constraints on Data-closeness and Needle Map Consistency for Shape-from-shading, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 12, pp. 1250-1267.
- [32] Worthington, P. L. and Hancock, E. R. (2001), Object Recognition Using Shape-from-Shading, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 5, pp. 535-542.
- [33] Wu, Z. and Li, L. (1988), A Line-Integration Based Method for Depth Recovery from Surface Normals, *Computer Vision, Graphics, and Image Processing*, 43, pp. 53-66.
- [34] Zhang, R., Tsai, P.S., Cryer, J.E. and Shah, M. (1999), Shape from Shading: A Survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 8, pp 690-706.
- [35] Zhao, W. and Chellapa, R. (2000), Illumination-insensitive Face Recognition Using Symmetric Shape-from-Shading. *Conference on Computer Vision and Pattern Recognition*, pp. 286-293.
- [36] Zhao, W.Y. and Chellapa, R. (2001), Symmetric Shape-from-Shading Using Self-ratio Image. *IJCV*, 45(1):55-75.

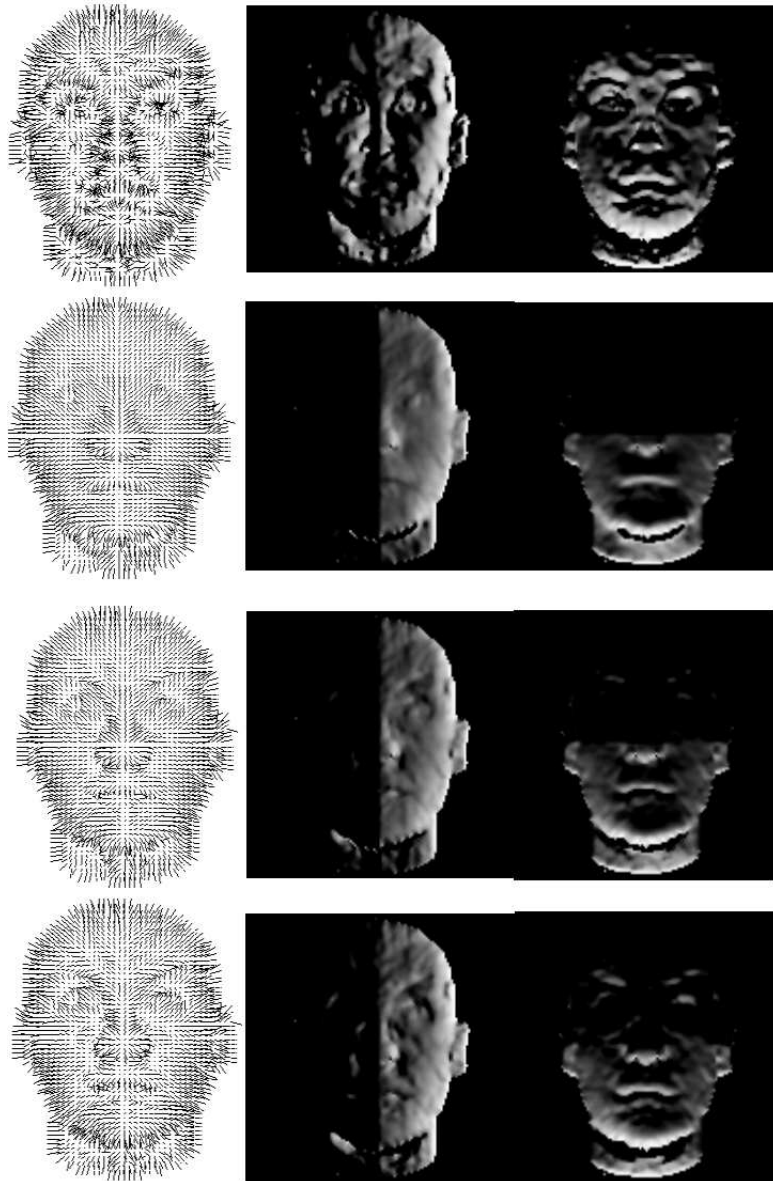


Figure 4: Gradient analysis. The first column shows the field of surface normals. The second and third columns show the result of re-illuminating the surface normals with light-source directions $(1, 0, 0)^T$ and $(0, -1, 0)^T$. The top row shows the result obtained with the initial field of surface normals. The subsequent rows show the recovered surface illuminations after applying the method with $\tau_x = \tau_y = 0$, $\tau_x = \tau_y = 0.2$ and finally $\tau_x = 0.3$ and $\tau_y = 0.4$, respectively.

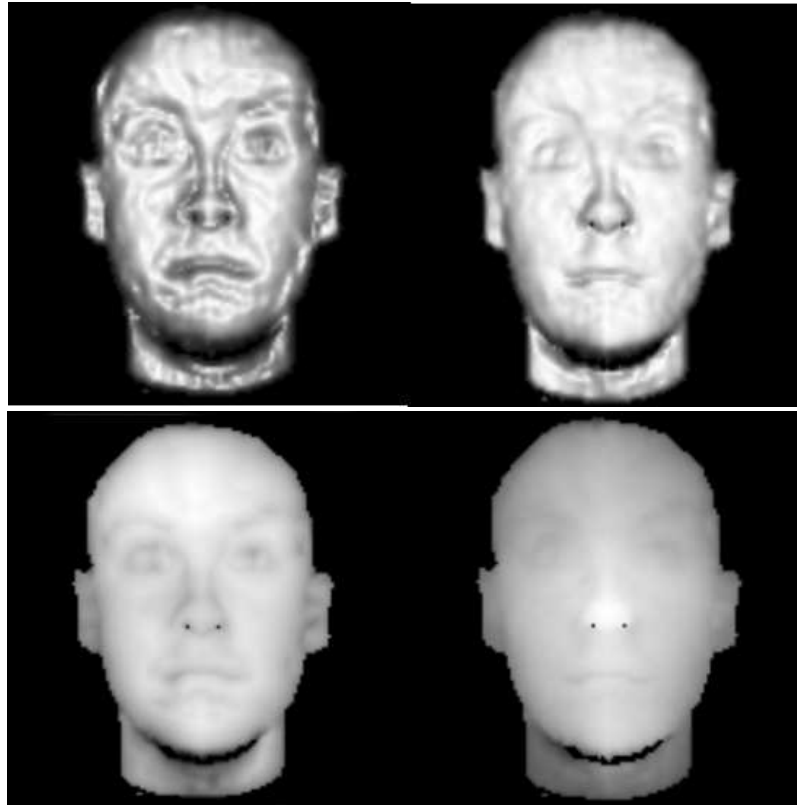


Figure 5: Illumination and range analysis. In the left-hand column of the figure we show the result obtained using the un-modified needle map, while the right-hand column shows the result obtained with the modified field of surface normals with $\tau_x = 0.3$ and $\tau_y = 0.4$. The top row shows the re-illuminations obtained when the light source direction is $(0, 0, 1)^T$. The bottom row shows the recovered height-maps as intensity plots.

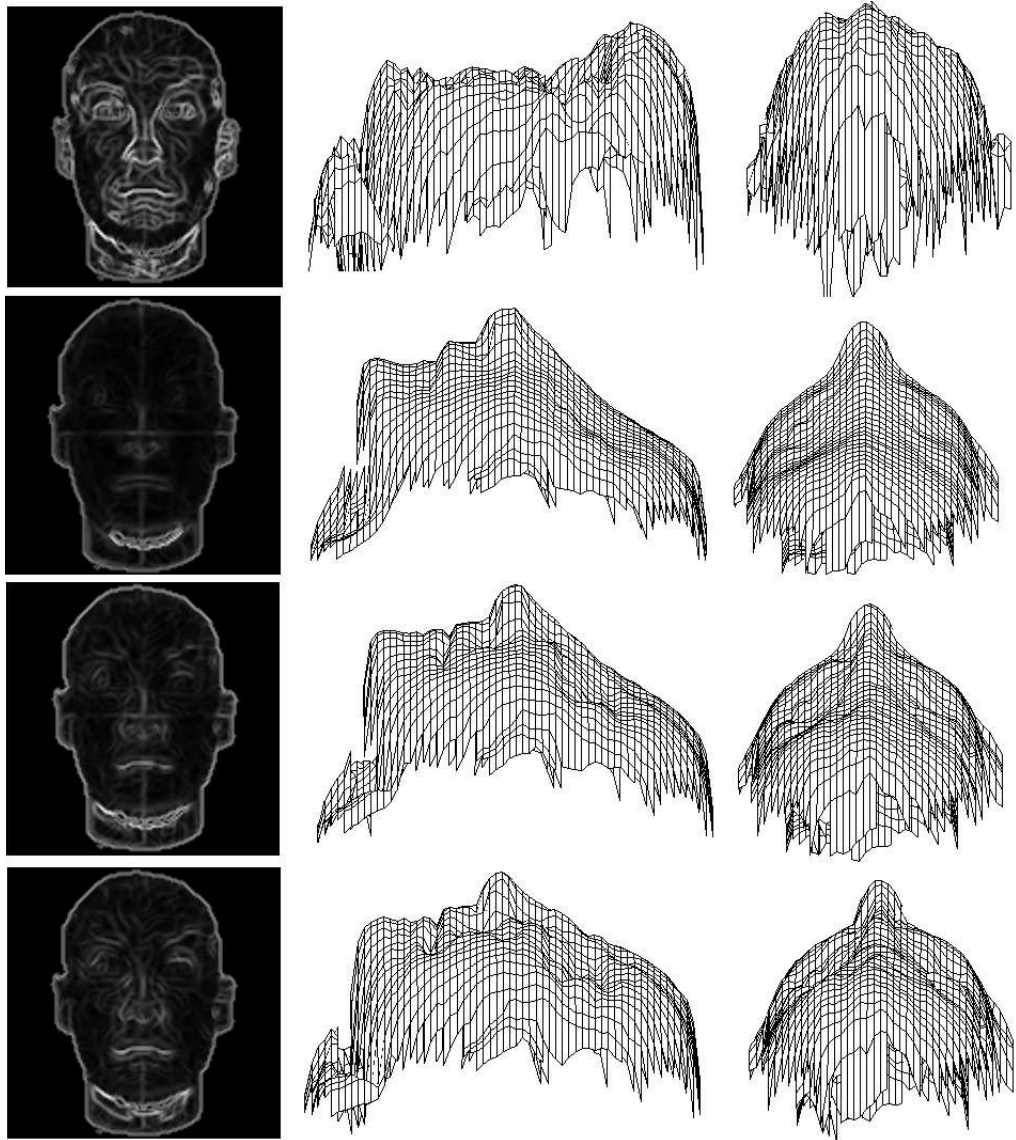


Figure 6: Curvedness and surface analysis. From left to right we show the curvedness map, together with profile (side) and top-down views of the surface wire-frames. In the top row we show the original case. Here the implosion of the nose is clear. The subsequent rows are for $\tau_x = \tau_y = 0$, $\tau_x = \tau_y = 0.2$ and finally $\tau_x = 0.3$ and $\tau_y = 0.4$, respectively.

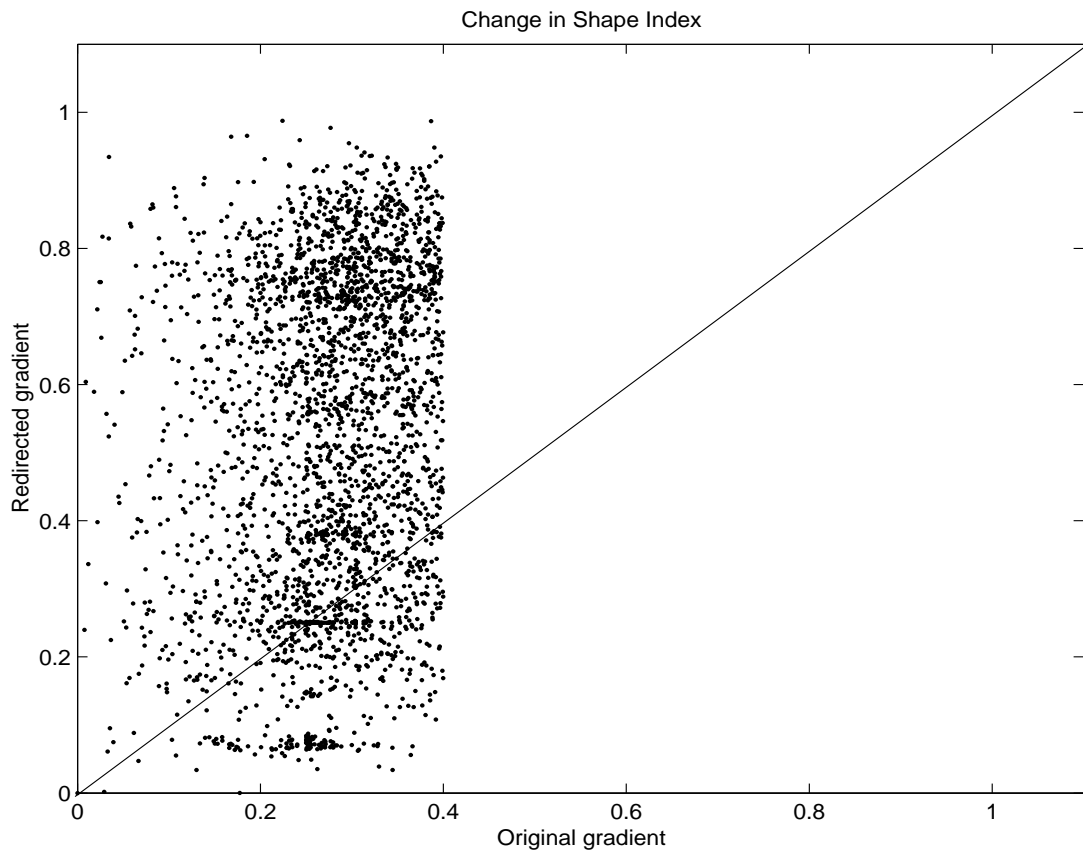


Figure 7: Convexity enforcement using shape index. Scatter plot comparison between the original and modified shape-indexes (the shape-index scale was normalized from 0 to 1). The x -axis corresponds to the shape index of the input field of normals. The y -axis corresponds to the shape index of that input needle map after redirecting it using $\tau_x = \tau_y = 0.4$ case

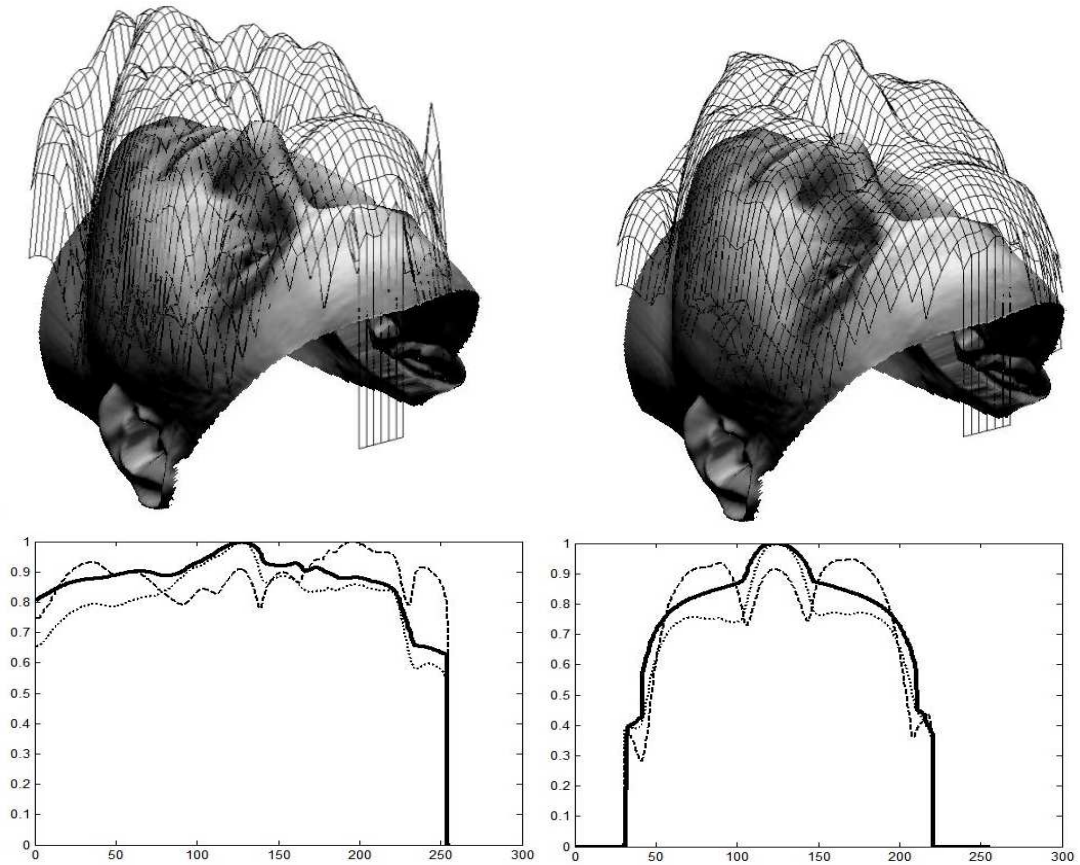


Figure 8: Height maps and cross plots comparison. We compare with ground truth the recovered height maps obtained with (right column) and without (left column) redirection of the normal field. The top row figure shows the recovered height maps superimposed on the image from which they were generated. In the bottom row we show longitudinal and transverse sections of the recovered height maps. The left-hand diagram corresponds to the longitudinal section along the x -axis and the right-hand diagram shows the transverse section along the y -axis. In both plots, the thick solid line represents the ground-truth surface, the dotted line illustrates the recovered surface using our method, and the dashed line shows the height data integrated using the original field of surface normals.

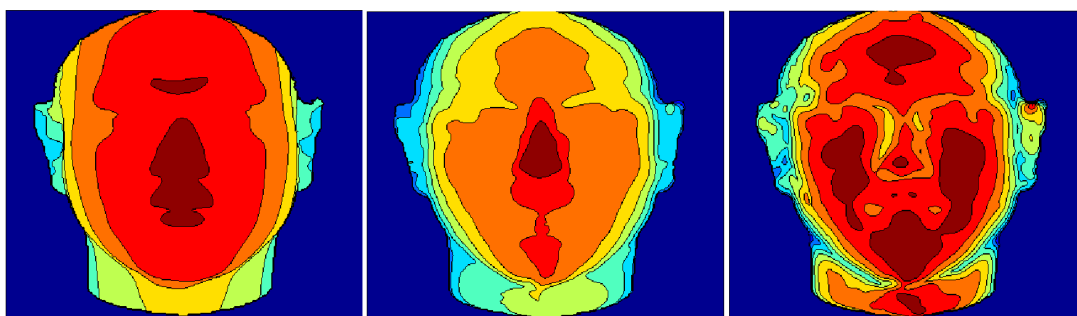


Figure 9: Iso-contour comparison. Regions in the images are color coded according to height. Discarding the background, the darkest regions correspond to highest values. The leftmost panel corresponds to the ground-truth surface, the middle panel presents the case when our method was applied, while the rightmost figure represents the recovered surface without gradient redirection.

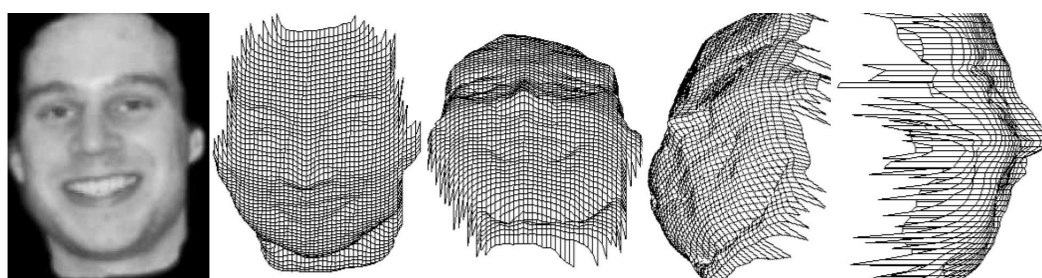


Figure 10: Recovered surface for a smiling face.

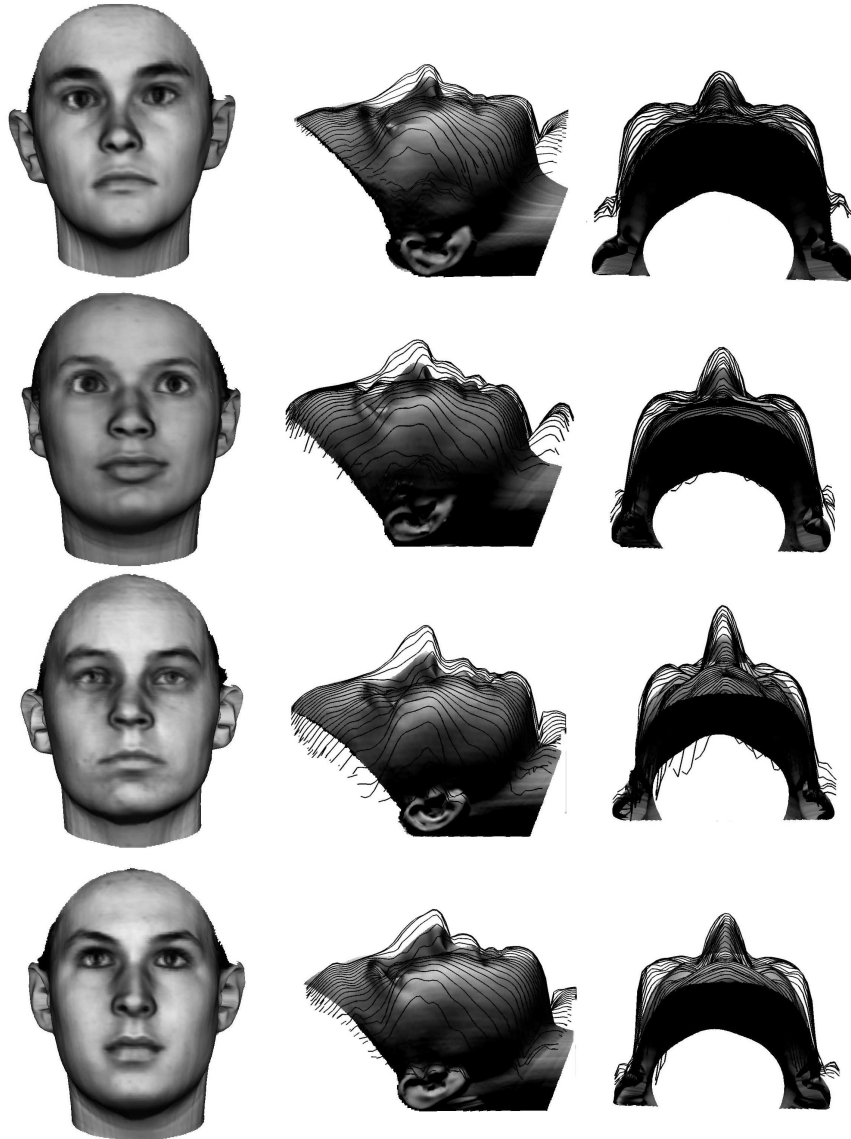


Figure 11: Height map analysis. The left-most column shows the raw image. The middle and right-most columns show superimposed plots of the recovered height maps after applying our method on the profile views (i.e. side) of the images from which the original gradient fields were calculated.

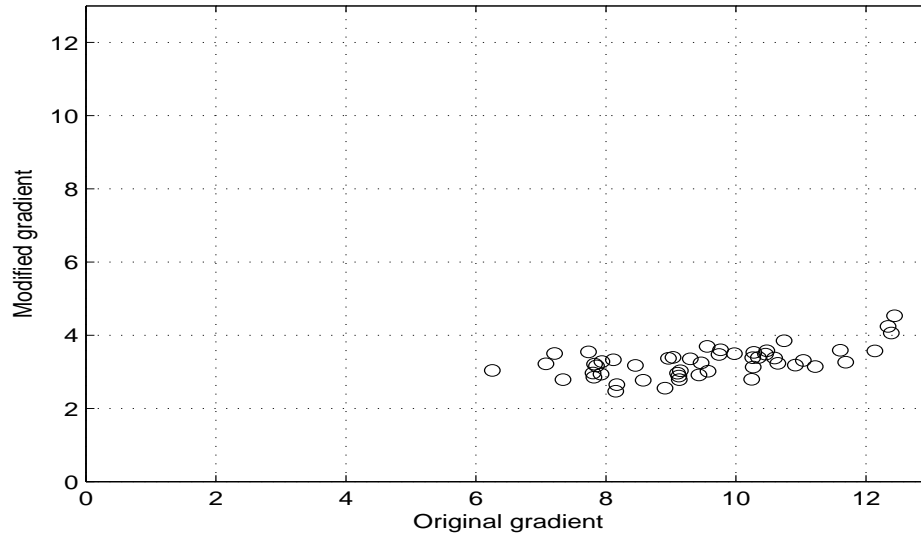


Figure 12: Average percentage of height difference. The diagram shows a scatter plot comparing the average percentage of height difference from the original gradient integrated surface (x -axis) against the redirected gradient integrated surface (y -axis).

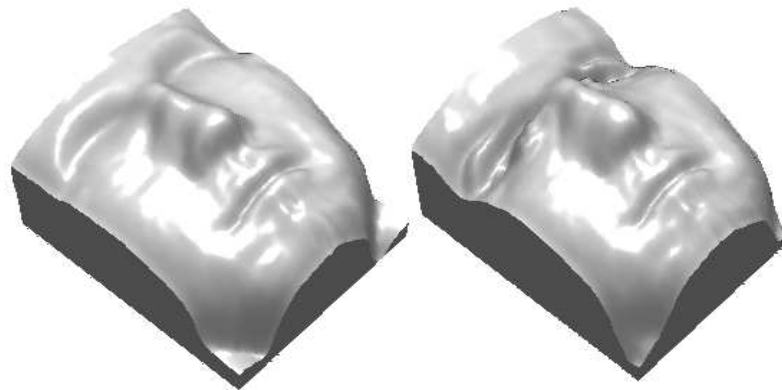


Figure 13: Results on an edited image. The left-most panel shows the calculated height map of the image of the face in Figure 16 after removing the eyes and eyebrows by setting their pixel brightness values to be that of the average over the skin. The right-most image shows the calculated height map on the un-edited face image.

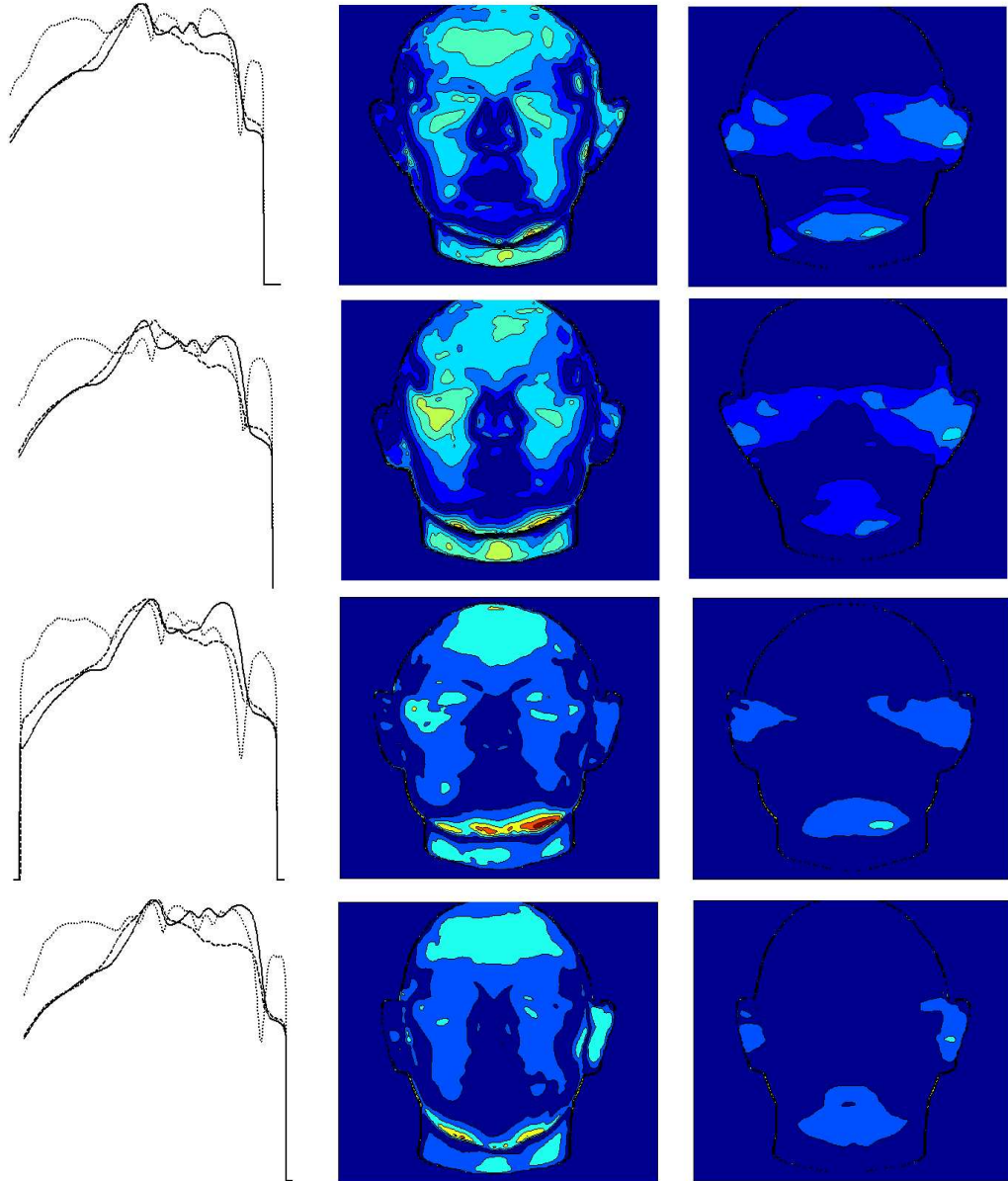


Figure 14: Height difference cross sections and iso-contour plots. The first column shows the longitudinal sections along the face for the recovered surfaces both before (dotted line) and after (dashed line) redirection of surface normals. The recovered height map after integration of the ground-truth normals is shown as a solid line. The second and third columns respectively show the iso-contour representations of the absolute height differences before and after re-directing the surface normals.

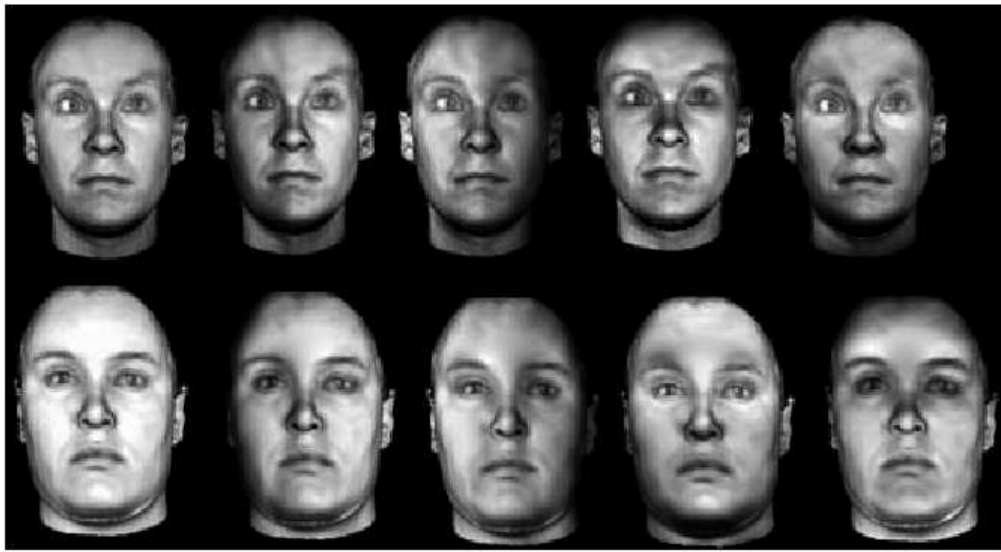


Figure 15: Re-illumination tests for two different images. The first column corresponds to the input image. The remaining columns show the generated re-illuminations after applying our method. For the second and third columns, the light source vector is nearly parallel to the x axis in both negative and positive directions, while nearly parallel to the y axis for the fourth and fifth columns.

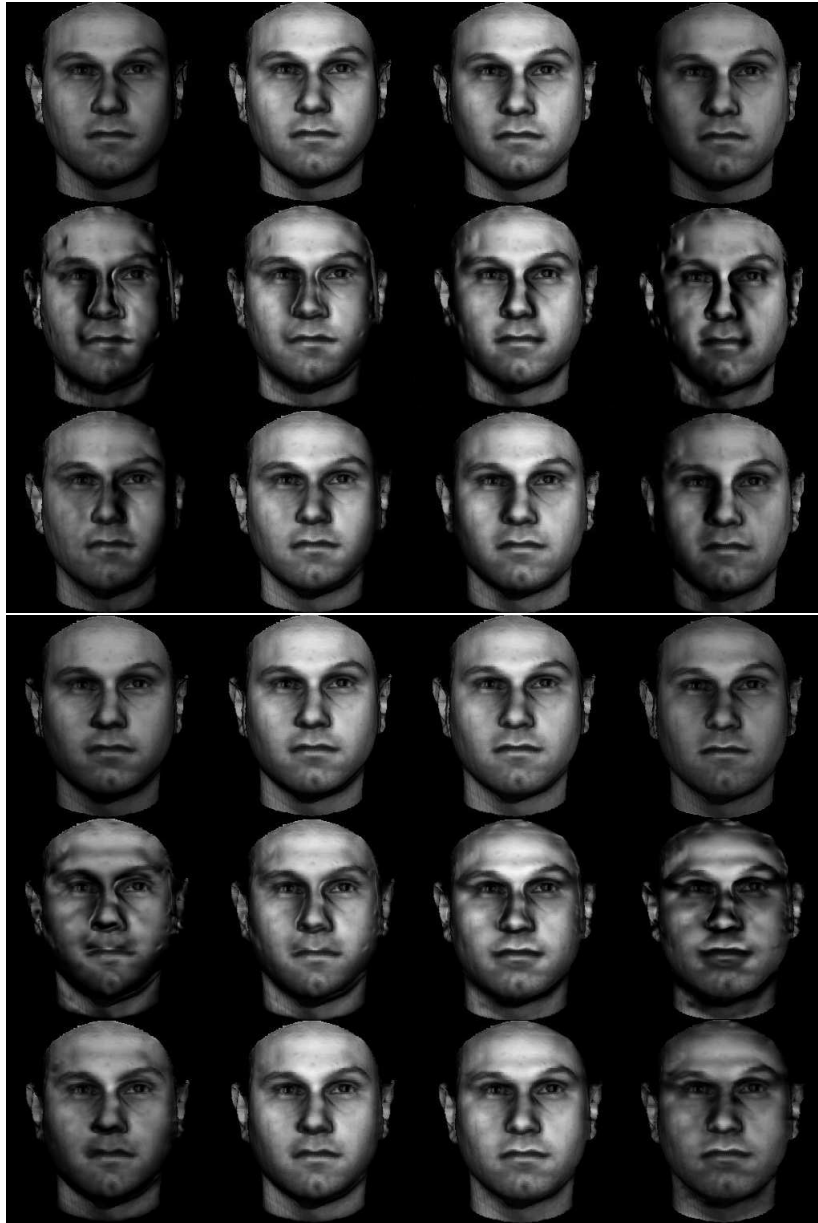


Figure 16: Comparison for re-illumination tests. For both sets of images, the top row represents the re-illumination results from the ground-truth normals, the middle row shows those obtained from the unaltered gradient field and the bottom row those obtained from the re-directed surface normals. From left to right, the light source direction makes an angle of -45 , -25 , 25 and $+45$ degrees to the image normal in the horizontal (x) direction for the upper set of images. In the lower set of images, the light source is moved in the same manner in the vertical (y) direction.



Figure 17: Comparison of re-illuminations using an unprocessed image of a face (single image in between two rows). The first row presents the results for the modified gradient field while second shows those obtained with the original gradient field.