Variable Albedo Surface Reconstruction from Stereo and Shape from Shading

Dimitrios Samaras, Dimitris Metaxas, P ascalFua and Yvan G. Leclerc
VAST Laboratory, CIS Department, University of P ennsylvania, Philadelphia PA
LIG, EPFL, Lausanne, Switzerland
AIC, SRI International, Menlo Park, CA
samaras,dnm@graphics.cis.upenn.edu, fua@lig.dl.epfl.ch, leclerc@ai.sri.com

Abstract

We presen t a multiview method for the computation of object shape and reflectance characteristics based on the integration of shape from shading (SFS) and stereo, for nonconstan talbedo and non-uniformly Lambertian surfaces. First we perform stereo fitting on the input stereo pairs or image sequences. When the images are uncalibrated, w e recover the camera parameters using bundle adjustment. Based on the stereo result, we can automatically segment the albedo map (which is taken to be piece-wise constant) using a Minimum Description Length (MDL) based metric, to identify areas suitable for SFS (typically smooth textureless areas) and to derive illumination information. The shape and the illumination parameter estimates are refined using a deformable model SFS algorithm, which iterates bet w een computing shape and illumination parameters. Our method takes into accoun tthe viewing angle dependent foreshortening and specularity effects, and compensates as much as possible by utilizing information from more than one images. We demonstrate that we can extend the applicability of SFS algorithms to real world situations when some of its traditional assumptions are violated. We demonstrate our method by applying it to face shape reconstruction. Experimental results indicate a significant improvement over SFS-only or stereo-only based reconstruction. Model accuracy and detail are improved, especially in areas of low texture detail. Albedo information is retrieved and can be used to accurately re-render the model under different illumination conditions.

1 Introduction

In this paper we propose a multiview approach for the recovery of object shape based on the integration of shape from shading (SFS) and stereo, in the presence of non-constant albedo and surfaces that are non-uniformly Lambertian. We use a deformable model approach, in order to impose illumination constraints derived from multiple images on data that have been acquired using a stereo method. We apply our method to estimate object shape and reflectance characteristics. The input images may exhibit a number of characteristics, each of which could be potentially crippling for a traditional SFS algorithm, such as specularities, changes in albedo, either on larger areas or as noise, non-point source illumination, non-Lambertian reflectance characteristics. The application of SFS improves the stereo results, both by complementing them in areas where they are not very precise and also by recovering a dense set of normals which vary smoothly across the surface. The latter is useful in order to distinguish albedo changes from illumination artifacts in the texture map. That way the model can be rendered under different illumination conditions and texture-mapped appropriately.

We use the problem of face reconstruction to demonstrate the applicability of our method. In recent years there has been a lot of w orkon modeling faces from image and range data. A number of approaches have been proposed for the recovery of facial geometry based on stereo [3], structured light [18], or laser scans. In [23] face shape is recovered based on images and an extensive database of laser scans. F ace reconstruction has been a favorite example to demonstrate shape-from-shading algorithms but most of these algorithms use a single image and assume a lambertian model and constant albedo [8, 12, 24]. In practice, however, these approaches are plagued by the fact that even faces do not have constant albedoes—some parts are darker than others and facial hair creates problems—and are not really Lambertian as demonstrated by the presence of specularities. It is appealing to integrate stereo and SFS, since they complement each other's strengths. Stereo w orksbetter on images with a lot of texture, SFS on images of smooth, textureless objects [2]. In earlier w ork [4] an approach was proposed that combined stereo and shading, and could handle smoothly varying albedoes. In this paper, instead, albedo is considered piecewise constant, and an attempt is made to segment the various areas. In the current approach w euse the stereo results to derive initial information about the illumination, the surface reflectance and the initial shape reliability. We then fit a much more densely sampled mesh by using SFS,

and reestimate the illumination parameters, during the fitting process.

In previous work [6], a facial animation model was fitted to stereo data. This requires a number of images (typically a short video sequence to derive models that are both precise and realistic). When only a few images are available, there might not be enough data for a perfect reconstruction. Here we show how SFS can be used to greatly improve the quality of the reconstruction. We also discuss the use of photometric information from multiple views.

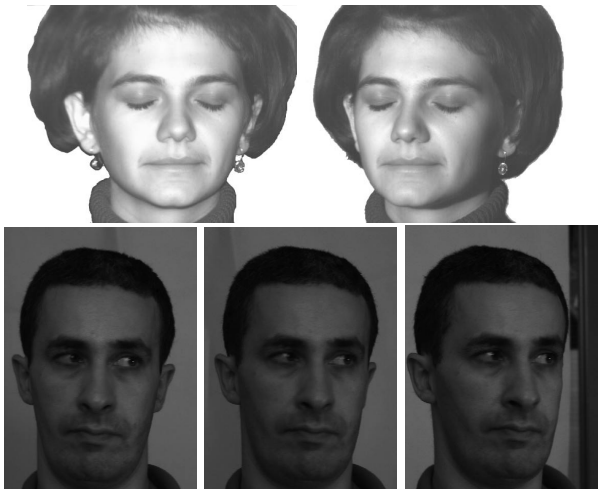

Figure 1: Typical input images for our method

2 Our Framework

1) We use stereo pairs or triplets, such as those in Figure 1 or short image sequences. These examples were acquired using uncalibrated cameras; the camera parameters were recovered using bundle adjustment [6].

2) We compute disparity maps for each image pair and fit a face model [9] to the corresponding 3D points [6], using least squares adjustment. The shape thus recovered is quite accurate, but it lacks detail in the textureless parts of the face. In some cases artifacts are produced by the mask, where they did not exist in the images. The output of the stereo fitting of the mask can be seen in Figure 2.

3) To improve the stereo results, we need to identify the areas where the stereo matches are most reliable. Since we usually do not have access to ground truth to estimate the reliability of the stereo information, we use a minimum description length (MDL) metric of stereo confidence, monotonically related to the self consistency of the matches [11]. This is described in Section 3.1

4) We use the MDL score to identify the areas where the stereo results are less reliable, i.e. the textureless parts of the skin, mostly on the cheeks and the sides of the nose, as explained in Section 3.2. We combine these results with image gradient information, in order to identify the areas where albedo is piecewise constant or smoothly varying, thus most suitable for SFS. Instead of using the SFS algorithm blindly, we only apply it to the parts of the face where it can serve the most. Existence of a model which is not far from the truth allows for the smooth integration of the piecewise shape estimates.

5) Using these areas only, we compute the illuminant direction, ambient light strength and albedo estimate. Furthermore we identify areas of specularity. Based on the stereo results, we can compute least squares estimates of the direction of the illuminant, the amount of ambient light and a uniform albedo of the image. If we compute the albedo locally at this stage, visible artifacts are introduced by the illumination effects and the shape inaccuracies. This is described in Section 4.1.

6) We consider the recovered shape to be a deformable surface which can be deformed in order to minimize edge, stereo and shading constraints. We perform illumination computations periodically to take advantage of the improved shape, and continue fitting using the new light estimates [22]. Our model-based approach allows us to overcome the requirement of SFS algorithms for a single orthographically projected image as an input, by giving us a framework on which to combine shading constraints from multiple images, taking occlusions and specularities into account, as described in Section 5. We modify the formulation of [21] for meshes with a high number of local degrees of freedom, in Section 4.2, in order to use much more efficient sparse matrix routines instead of SVD, in Section 4.3

Our results show improved shape estimates, especially in areas of low texture detail. Furthermore we recover a dense set of normals which vary smoothly across the surface. The latter is useful in order to separate albedo changes from illumination artifacts in the texture map. That way the model can be rendered under different illumination conditions and texture-mapped appropriately.

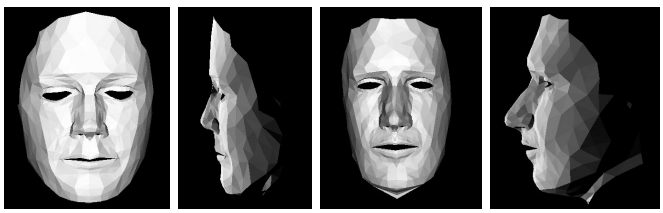

Figure 2: Face models fitted to stereo data for the subjects in Fig. 1

3 Evaluating the stereo results

We now introduce the MDL-based stereo reliability metric that we utilize, to evaluate the reliability of the initial stereo-based reconstruction for the areas whose shape we trust most. We also explain how we use this metric to identify areas which are potential candidates for SFS. We expand upon the MDL-based stereo reliability metric that was introduced in [11] and show its applicability to the problem of detecting the boundaries of albedo changes on a surface of piecewise constant albedo.

3.1 MDL-Based Coding Loss

The image-matching measure that we use for matching windows is the traditional sum of squared differences (SSD) [5]. For multi-image mesh optimization [4], it is replaced by the sum, over all visible sample points within a facet, of the squared variance of the gray level of the projections of these sample points. When using only two images, these two measures are equivalent. For normalization purposes, we typically compute it using differences of Gaussians rather than the images themselves.

The problem with this SSD measure is that it is ambiguous. That is, a low SSD measure can occur not only when the facet is correctly located (as expected), but also when the facet is incorrectly located and the object's surface is spatially uniform.

Intuitively, then, we want an image-matching measure that is low only when the match between the predicted and observed pixel values is close and the pixel values form a sufficiently complex pattern that it is unlikely to be matched elsewhere. We have developed a measure that satisfies this intuitive requirement, which we call the coding loss. It is based on Minimum Description Length (MDL) theory [20]. In MDL theory, quantized observations of a random process are encoded using a model of that process. This model is typically divided into two components: a parameterized predictor function and the residuals (differences) between the observations and the values predicted by that function. The residuals are typically encoded using an i.i.d. noise model [10]. MDL is basically a methodology for computing the parameters that yield the optimal lossless code length for this model and for a given encoding scheme.

Here, the process we are observing is the object's surface. The observations are the pixel values in all of the images covered by a given correlation window or facet. In the first case, the parameters are the disparity values and surface-reflectance estimates of the individual pixels. In the second case, the disparity values are replaced by the facets' vertex coordinates.

In both cases, let us assume that we have N images with $N = 2$ in the case of the window-based score, $N >= 2$ in the case of the mesh score. Let M be either the number of pixels in the correlation window or the number of samples in a facet and let g_i^j be the image gray level of the i^{th} pixel observed in image j.

These gray levels can be described independently in each image. For image j, the length of such a description can be approximated by:

$$C_j = M(\log \sigma_j + c), \tag{1}$$

where σ_j is the measured variance of the $g_{i\,1 \le i \le N}^j$ and $c = (1/2)\log(2\pi e)$.

Alternatively, these gray levels can be expressed in terms of the mean gray level $\overline{g_i}$ across images and the deviations $g_i^j - \overline{g_i}$ from this average in each individual image. The cost of describing the means, can be approximated by

$$\overline{C} = M(\log \overline{\sigma} + c), \tag{2}$$

where $\overline{\sigma}$ is the measured variance of the mean gray levels. Similarly, the coding length of describing deviations from the mean is given by

$$C_j^d = M(\log \sigma_j^d + c), \tag{3}$$

where σ_j^d is the measured variance of those deviations in image j. In the case $N = 2$, the variance simply becomes the squared difference. Note that, because we describe the mean across the images, we need only describe $N - 1$ of the C_j^d. The description of the Nth one is implicit.

We take the coding loss per pixel to be the difference between these two codings lengths, normalized by the number of samples, that is

$$Loss = \overline{C} + \sum_{1 \le j \le N-1} C_j^d - \sum_{1 \le j \le N} C_j \ . \tag{4}$$

When there exists a good match between images, the $g_{i\,1 \le j \le N}^j$ have a small variance. Consequently the C_j^d should be small, $\overline{C}$ should be approximately equal to any of the C_j and $Loss$ should be negative. However, C_j can only be strongly negative if these costs are large enough, that is, if there is enough texture for a reliable match.

3.2 Segmenting the face

We use the results of the MDL metric based above to segment the albedo maps of the face as follows: We get consistently better (i.e. lower) scores in the areas where we have both a good match and high texture, such as the eyebrows, the mouth, the eyes, the ridge

of the nose, etc. We choose not to fit those areas with SFS, since they are the areas where we expect stereo to do best. Conversely, the MDL score is generally higher on the cheeks, the forehead, the sides of the nose, and these areas become potential candidates for SFS fitting.

The fact that the MDL score is lower where there is a lot of texture, can be used to segment changes of texture on otherwise smooth geometry. In the case of faces, this has an interesting use for male subjects with beginnings of facial hair growth (i.e. "five o'clock shadows"). This shadow is not visible enough to create strong edges, the way eyebrows do, and cannot be easily detected by simple image gradient based methods. However, it causes the reflectance properties of the skin to change and hence needs to be identified. We notice that the MDL scores for male subjects exhibit a very distinct band of lower scores where this shadow line starts. This is not exhibited in female subjects, as can be seen in Figure 3. This can be explained because the skin is not as smooth anymore, and it is a little darker. That creates enough texture to make a clear difference in the MDL scores. Since no drastic change in curvature at that part of the geometry is expected that would cause such texture in the images, the texture can only be attributed to the change of the reflectance properties.

4 Using Shading Information

In this section we present the use of illumination information in our images. We compute the parameters of the illumination information, based on the already computed shape. We re-evaluate these parameters as more information on the surface reflectance properties becomes available, or as we update the shape.

Based on the updated illumination parameters, we derive the illumination constraints that we will impose to refine the shape. This is done using a generalization of the methods in [21, 22] to incorporate constraints from multiple images under perspective projection. We also introduce a new fast method for computing the illumination constraint forces.

4.1 Computing the illumination parameters

We use the reconstructed stereo data to compute the light source direction using the Levenberg-Marquart [14], least squares minimization method, under the assumption of a Lambertian reflectance model with the skin having uniform albedo. Although simplistic, the Lambertian model is a reasonable assumption, when the incidence and emittance angles of the light are less than seventydegrees as experimentally reported in [13]. Since we have more than one view, for somewhat frontally facing faces we can usually find at least one image for each patch for which this holds.

Given the number of data points and the general accuracy of the model, local violations of the assumptions do not influence the computation of the global lighting parameters.

The lighting model we use is the Lambertian model with an ambient light term.

$$I_L = B\rho\ (\hat{\mathbf{s}} \cdot \hat{\mathbf{n}} + A), \tag{5}$$

where θ is the angle between the surface unit normal vector, $\hat{\mathbf{n}} = \mathbf{n}/\|\mathbf{n}\|$, and the unit light source direction vector $\hat{\mathbf{s}}$. B is the strength of the light source and ρ is the albedo of the surface, which represents the fraction of the incident light to the surface that is reflected. A is the amount of ambient light, also a constant. All the constants are normalized for unit light source strength.

Using the derived illumination values, and the computed shape, we can now relax the uniform albedo assumption and compute an albedo map for the face, by computing an albedo value per pixel (or per shape facet). The piecewise uniformity and smoothness of this map is a good indication of the quality of the recovered shape and light. Given the inaccuracies of the shape locally, the map exhibits artifacts that do not correspond to realistic albedo variations, as can be seen in Figure 7(a),(b). Thus, it is not safe to use this map to segment the regions of the face where we expect the albedo of our model to be smooth and constant. Instead, we use the pixel gradient information on the input images to identify parts of the model where intensity changes do not correspond to the shape changes. The MDL-based method described in this work provides an extra criterion.

4.2 Illumination Constraints

In this section we describe how to incorporate illumination constraints from a single image, into a deformable model. We assume that an object's illumination is described by a lighting model of the general form $L(\mathbf{l}_p, \mathbf{q})$, where $\mathbf{l}_p$ are the lighting model parameters (both light source parameters and shape reflectance properties, such as the surface albedo) and $\mathbf{q}$ the shape parameters. This model must be differentiable with respect to the shape and light parameters, as is the case with the lighting models commonly used in computer vision, e.g. the Lambertian model, or [16]. Then, in order to recover the shape parameters, we need to minimize the lighting constraint $C = L(\mathbf{l}_p, \mathbf{q}) - I_L$, where I_L is the measured light intensity at every pixel of the image. If we assemble all the constraints in vector $\mathbf{C}$ then, using a Baumgarte [1] stabilization approach, as

(a) (b) (c) (d) (e)

Figure 3: MDL scores for a male and female subjects: Eyebrows, eyes and mouth are clearly segmented, as well as the ridge of the nose. (a) Input view of female subject, (b) MDL score for female subject, (c) Input view of male subject, (d) MDL score for male subject. Notice the strong line in the cheek of (d) which defines two different textures on the face. No such lines exist in (b) when there is no change in geometry. (e) Profile photo where the change in texture is very visible. This photo was not used in the experiments and is provided only for reference.

shown in [21], we need to estimate the object shape subject to

$$\mathbf{C_q}\dot{\mathbf{q}} + \alpha\mathbf{C} = 0, \tag{6}$$

where $\mathbf{C_q}$ is the constraint Jacobian matrix.

Following the formulation of [15], at every point of the model there are external forces $\mathbf{f}$ exerted by the data, in this case by the shape estimates already computed, and internal forces $\mathbf{K}_d\mathbf{q}$ due to the smoothness constraints derived from a finite element formulation of the model. In this case we assume C_0 continuity.

As described in [21], the change in model shape parameters is given by:

$$\dot{\mathbf{q}} = \mathbf{f} - \mathbf{K}_d\mathbf{q} - \mathbf{C}_\mathbf{q}^\top\boldsymbol{\lambda}, \tag{7}$$

where $\mathbf{f}$ are generalized edge based forces, $\mathbf{C_q}$ is the Jacobian matrix of the constraints $\mathbf{C}$ w.r.t. the shape parameters and $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_n]^\top$ are the Lagrange multipliers.

The matrix $\mathbf{C}_\mathbf{q}^+$ is the pseudo-inverse of the matrix $\mathbf{C_q}$, and if we define the vector of all non-constraint (generalized) forces as $\mathbf{b} = \mathbf{f} - \mathbf{K}_d\mathbf{q}$, (7) becomes

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{b} - \mathbf{C}_\mathbf{q}^+(\alpha\mathbf{C} + \mathbf{C_q}\mathbf{b}) && (8)\\ &= -\mathbf{C}_\mathbf{q}^+\alpha\mathbf{C} + (\mathbf{I} - \mathbf{C}_\mathbf{q}^+\mathbf{C_q})\mathbf{b}. && (9) \end{aligned}$$

4.3 Fast Computation of Constraint Forces

In the above method we calculate the pseudoinverse $\mathbf{C}_\mathbf{q}^+ = \mathbf{C}_\mathbf{q}^\top(\mathbf{C_q}\mathbf{C}_\mathbf{q}^\top)^{-1}$ using Singular Value Decomposition (SVD). In a high resolution mesh with local deformations only, which has a few thousand nodal parameters, the computational cost of SVD becomes prohibitive, as SVD cannot take advantage of the sparsity of the constraint Jacobian matrix. Approximate SVD techniques which calculate only the most significant singular values, are not useful when there are no global parameters to dominate the shape description. Here we propose a new technique for the case of large deformable meshes with only local degrees of freedom. Motivated by the desire to use sparse matrix inversion methods, we notice an alternative formulation of (9) as

$$\dot{\mathbf{q}} = \mathbf{b} - (\mathbf{C}_\mathbf{q}^\top\mathbf{C_q})^{-1}\mathbf{C}_\mathbf{q}^\top(\alpha\mathbf{C} + \mathbf{C_q}\mathbf{b}). \tag{10}$$

$(\mathbf{C}_\mathbf{q}^\top\mathbf{C_q})$ is an $n \times n$ matrix where n is the number of nodes in the mesh. Since the system is over-constrained [21], the rank of this matrix is the rank of $\mathbf{C_q}$, which is the number of degrees of freedom in the orientation of the facets of the object. Consequently, singularities can occur only if we have more nodal parameters than degrees of freedom in the orientation of the facets of the shape. In general, because of redundancies in the shape parameterization, we can have more parameters than degrees of freedom in the shape of the object. In the case where all the parameters are local nodal variables, possible degeneracies in this matrix can be caused only by nodes in the boundaries of the mesh which belong to only one triangle. The reason is that the orientation of that triangle can already be determined by its other two nodes which belong to other triangles also. By not including these nodes in $\mathbf{C_q}$ we guarantee that the dimensionality of $(\mathbf{C}_\mathbf{q}^\top\mathbf{C_q})$ will not exceed the dimensionality of the mesh and thus $(\mathbf{C}_\mathbf{q}^\top\mathbf{C_q})^{-1}$ will exist. This inverse matrix can be calculated with standard numerical methods. such as LU decomposition, which can be much faster. Furthermore, the matrix is highly sparse and thus sparse matrix routines can speed up computation significantly. If we have a regularly sampled mesh the maximum distance of non-zero entries will be on the order of $\sqrt{n}$, making banded matrix routines attractive.

5 Selective Application of SFS

Once we know the projection transformation of each image, incorporating illumination constraints from multiple images to our model, is a straightforward extension of the single view case. In order to calculate the value of the illumination constraints on our model, we

project the model using the camera transformations under which each of the images was captured. When doing this, we perform a visibility check for each model point $\mathbf{p}$ with respect to every image. The image value at the projection point $\mathbf{p}_i$ of $\mathbf{p}$ on image i gives rise to a new constraint in $\mathbf{C}$ in (6). This visibility check ensures that the lighting constraints are applied to the proper model points, and it takes care of occlusions. Each model point has constraints derived only from the images in which it is visible. Furthermore, we do not sample facets of the model, for which the incidence or emittance angles of the light are more than seventy degrees, since the Lambertian model is much more unreliable, and foreshortening effects are significant. In the ideal Lambertian case, constraints on the same model point from different images would have the same value (once the image values have been normalized for camera gain), since the radiance of a point is independent of viewing direction. In real images though, this is not always the case. The least-squares formulation of (9) allows us to estimate the parameters that "best" satisfy all constraints. The images have been acquired under perspective projection, therefore, as the model iteratively deforms to fit the data, the visibility of its points will change. Consequently, the visibility check has to be performed at every iteration.

The SFS method is not applied to the whole face model automatically, but only to regions where the reliance on stereo is low, as measured by the MDL metric, and the gradient changes in the image indicate a smooth surface. Once we identify these regions, we recalculate the light parameters based on these regions. We detect specularities in the areas where the image brightness is greater than the possible range of diffuse values of the lighting model. The accuracy of the light model parameters actually affect whether an image point would be characterized as specular. One indication that the recomputed lighting parameters are more accurate than before is the fact that in our experiments the points that fall outside the model reduce by about thirty percent. Weak smoothness forces in the model guarantee the blending of areas fitted under different modalities.

6 Results

6.1 Shape Reconstructions

We show results for the two faces in Figure 1. One is a very smooth and textureless face of a female, the other one is a significantly textured face of a male with a five o'clock shadow.

In the first case (Fig.6), SFS is able to recover almost the whole surface. The MDL metric as shown in Figure 3(b) segments only the eyes, mouth and nose. The

(a) (b) (c)

Figure 4: Top row: Recovered Shape for male subject (a) stereo result (b) SFS result without light re-estimation (c) SFS result after light re-estimation and more fitting. Middle row: Textures of the above. Bottom row: Textured details of the above

results are very smooth. Problem areas remain near discontinuities, e.g. the chin and around the eyes.

In the second case (Fig.4), significant parts - especially in the lower part of the face - can be fitted only with stereo. There are significant albedo changes because of the stubble, mostly on the mouth and chin area. This is good for stereo, as can be seen in Figure 3(d). These darker texture areas significantly influence the illumination computations. Segmenting the albedo map before illumination calculations is necessary, yielding improved estimates of the light and reflectance. Despite the reduced area where SFS is applicable, we can still use it to obtain smooth reconstructions of the cheeks, forehead, and nose.

In order to evaluate the reconstruction results, we provide profile views of the subjects in Fig. 5. These

views are provided for comparison only and were not used for reconstruction.

In both reconstructions, recomputing the illumination while fitting was crucial. Due to the least squares approach, the reflectance values at a number of points cannot be explained by the lighting model as initially computed; these points are considered outliers. As the shape error decreases, we can recompute more accurate values for the illumination parameters, thus covering some of the outlying points too. In both examples, recomputing the light parameters increased the total area fitted by SFS by approximately twenty percent, yielding much smoother surfaces.

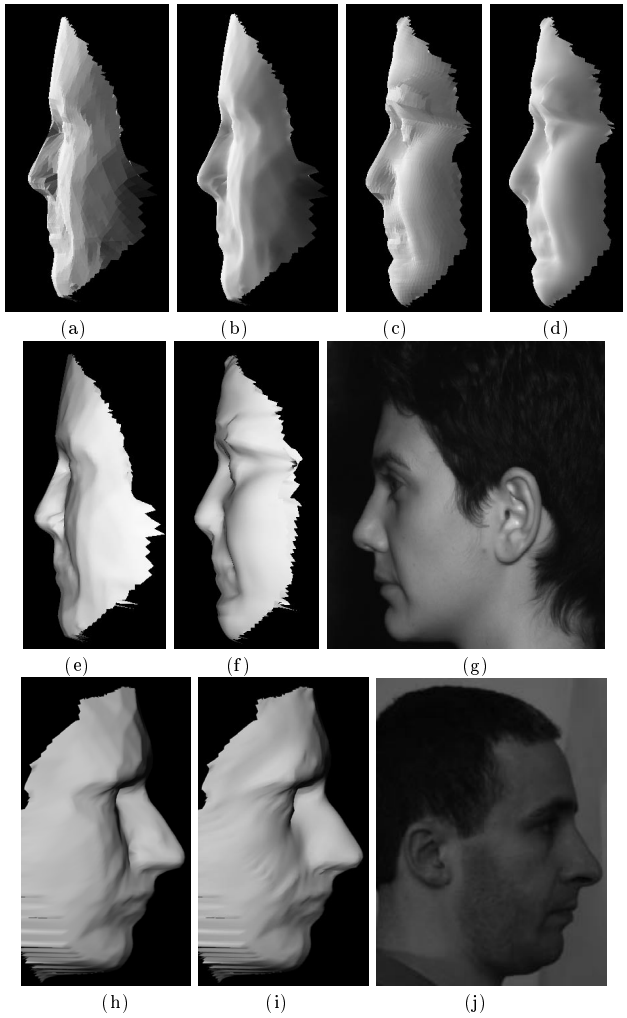

Figure 5: Profile views:(a) Flat-shaded stereo results (c) Flat-shaded SFS results(b),(e),(h) Smooth-shaded stereo results (d),(f),(i) Smooth-shaded SFS results. Profile photos (g),(j) were not used in reconstruction

6.2 De-lighting texture maps

Face modeling applications rely on texture mapping facial appearance onto a face model. The texture maps are often generated from a set of registered images of the model [7], [17]. A common problem is the fact that these maps contain lighting artifacts as part of the facial texture. Our method can be used to "de-light" texture maps. Our method returns accurate light parameter estimates and surface geometry for a dense set of points. With these we can calculate a shaded image S of the geometry, using the input image I camera parameters. Each pixel of the corresponding "albedo" image A is $A = I/S$. Due to better shape estimates, these images are superior to the ones produced using the stereo only results, as shown in Figure 7.

As can be seen in Figures 5(c),(d) and 6 the difference between the flat-shaded and smooth-shaded versions is very small. This demonstrates the high density of the surface normals that we obtain, allowing for detailed texture recovery. The resulting texture maps are much smoother, and more accurate. As an example, dimples on the cheek are correctly recovered as a geometry effect, dependent on the viewing angle and not as texture which appears the same from all viewing locations.

7 Conclusions

We have presented a method for the recovery of surface shape, based initially on stereo and then SFS. We segment the surface into areas of piecewise constant albedo where we can apply SFS. Our results significantly improve the recovered shape and allow the generation of more accurate texture maps. Further extensions of this workinclude the ability to recover surface shape from specularities and the use of better reflectance models.

References

[1] J. Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1:1-16, 1972.

[2] Cryer, J.E., Tsai, P.S., Shah, M., Integration of Shape from Shading and Stereo, *PR(28)*, No. 7, 1995

[3] Devernay, F., Faugeras, O.D., Computing Differential Properties of 3D Shapes from Stereoscopic Images without 3D Models, *CVPR94*.

[4] P. Fua and Y.G. Leclerc. Object-Centered Surface Reconstruction: Combining Multi-Image Stereo and Shading, *IJCV(16)*, September 1995.

[5] P. Fua, Combining Stereo and Monocular Information to Compute Dense Depth Maps that Preserve Depth Discontinuities, *IJCAI 1991*, pg. 1292-1298

[6] P. Fua, C. Miccio, From Regular Images to Animated Heads: A Least Squares Approach, *ECCV 1998*.

[7] P. Fua. Using Model-Driven Bundle-Adjustment to Model Heads from Raw Video Sequences, *ICCV 1999*.

[8] B.K.P. Horn and M.J. Brooks. eds. *Shape from Shading*. Cambridge, MA, MIT Press, 1989

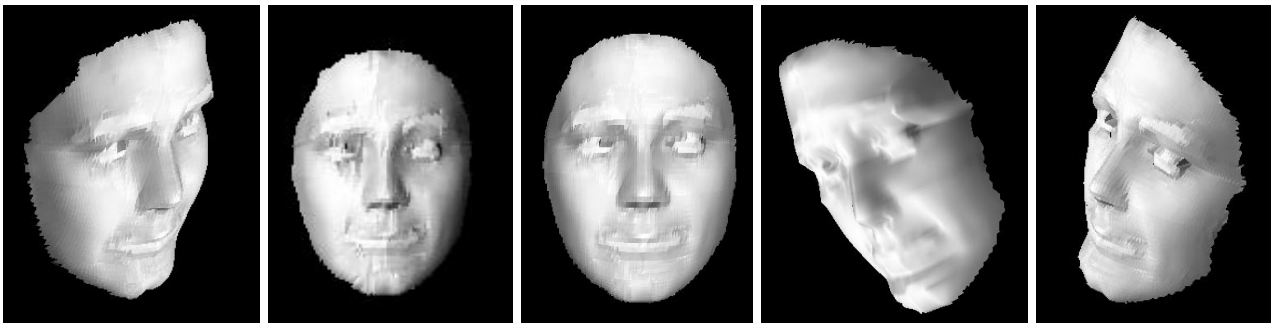

Figure 6: Views of recovered shape under various illumination directions

(a) (b) (c) (d)

Figure 7: Top row: (a), (b), Albedo for initial mesh (c), (d) Albedo for final shape. Bottom row, Texturing (a), (b), the initial mesh and (c), (d), the final shape using the above albedoes

[9] P. Kalra, A. Mangili, N. Magnenat Thalmann, D. Thalmann, Simulation of Facial Muscle Actions Based on Rational Free Form Deformations, *Eurographics 1992*.

[10] Y. G. Leclerc, Constructing Simple Stable Descriptions for Image Partitioning, *IJCV(3)* 1, 1989.

[11] Y. G. Leclerc, Q.T. Luong, P. Fua, Characterizing the Performance of Multiple-image Point-correspondence Algorithms using Self-Consistency, in *Vision Algorithms: Theory and Practice*, Workshop 1999

[12] Y.G. Leclerc and A.F. Bobick. The Direct Computation of Height from Shading. In *CVPR 1991*, 1991.

[13] S. Marschner, S. Westin, E. Lafortune, K. Torrance, and D. Greenberg. Reflectance measurements of human skin. Technical report PCG-99-2, Program of Computer Graphics, Cornell University, Jan 1999.

[14] D.W. Marquart *SIAM* 11:431-441, 1963

[15] D. Metaxas, Physics-Based Deformable Models: Applications to Computer Vision, Graphics and Medical Imaging, Kluwer-Academic Publishers, 1996.

[16] M. Oren and S.K. Nayar.Diffuse reflectance from rough surfaces. In *ICCV 1993* 1993.

[17] F.Pighin, J. Hecker, D. Lischinski, R. Szeliski, D. H. Salesin, Synthesizing realistic facial expressions from photographs, *SIGGRAPH 1998*.

[18] Proesmans, M., VanGool, L., Oosterlinck, A., Active Acquisition of 3D Shape for Moving Objects, *ICIP96(19P1)*.

[19] W.H. Press, S.A. Teukolsky , W.T. Vettering and B.P. Flannery. *Numerical Recipes in C*, Cambridge University Press, 1992.

[20] J. Rissanen, *Encyclopedia of Statistical Sciences*, chap. Minimum-Description-Length Principle, pg.523-527, John Wiley and Sons, New York, New York, 1987.

[21] D. Samaras and D. Metaxas Incorporating Illumination Constraints in Deformable Models, *CVPR 1998*

[22] D. Samaras and D. Metaxas Coupled Lighting Direction and Shape Estimation from Single Images, *ICCV 1999*.

[23] Blanz, V. and Vetter, T. A Morphable Model for the Synthesis of 3D Faces. *SIGGRAPH 1999*.

[24] R. Zhang, P. Tsai, J.E. Cryer,M. Shah. Analysis of shape from shading techniques. In *CVPR 1994*.