3D Shape Reconstruction of Mooney Faces

Ira Kemelmacher-Shlizerman[1] Ronen Basri[1,2] Boaz Nadler[1]

[1]Dept. of Computer Science and Applied Math.
The Weizmann Institute of Science
Rehovot 76100, Israel

[2]Toyota Technological Institute at Chicago
1427 East 60th Street
Chicago, Illinois 60637

Abstract

Two-tone ("Mooney") images seem to arouse vivid 3D percept of faces, both familiar and unfamiliar, despite their seemingly poor content. Recent psychological and fMRI studies suggest that this percept is guided primarily by top-down procedures in which recognition precedes reconstruction. In this paper we investigate this hypothesis from a mathematical standpoint. We show that indeed, under standard shape from shading assumptions, a Mooney image can give rise to multiple different 3D reconstructions even if reconstruction is restricted to the Mooney transition curve (the boundary curve between black and white) alone. We then use top-down reconstruction methods to recover the shape of novel faces from single Mooney images exploiting prior knowledge of the structure of at least one face of a different individual. We apply these methods to thresholded images of real faces and compare the reconstruction quality relative to reconstruction from gray level images.

1. Introduction

Two-tone (thresholded, black and white) images of faces were first introduced in the 1950s by Craig Mooney [17] to test the ability of children to form a coherent percept of shape (termed "perceptual closure") on the basis of very little visual detail. Such images (see examples in Figure 1) may seem initially difficult to interpret, but eventually lead to a rich and stable percept of the objects in the image.

Mooney images have fascinated psychologists and neurobiologists throughout the past half a century. Their ambiguous nature (is it an object or a random collection of blobs?), face specificity (faces seem more readily identifiable than other objects), and sudden interpretability led to a flurry of studies that used Mooney images to investigate the development of shape perception in children [17, 28], to demonstrate that parts of the fusiform gyrus brain area specialize in faces [1, 9, 13], and to reveal synchronous activity of neurons across brain areas [7, 21].

Figure 1. Three Mooney face images – initially seem difficult to interpret, but eventually lead to a rich and stable percept of the objects in the image (a face) on the basis of very little visual detail.

The remarkable perception of Mooney images raises the question of whether their interpretation is primarily bottom-up, driven by image data and guided by generic assumptions, or alternatively whether it is essentially a top-down process driven by memory and attention and preceded by a preliminary recognition process. Psychological studies seem to suggest the latter. Among the evidence supporting this view is that people usually fail to perceive upside-down faces, arguably due to their unfamiliarity [9, 21], and that pre-exposure to original gray level (or color) image facilitates their recognition [7, 11]. In an extensive study Moore and Cavanagh [18] showed that shape primitives (*e.g.*, generalized cones) are rarely perceivable in two-tone images, both in isolation and in novel configuration with other primitives, even when the image contains explicit hints about the direction of the light source. These shapes, however, can readily be interpreted from gray level images and even from degraded line drawings. Familiar classes of objects, in contrast, are much more often perceivable in two-tone images. Even volumetric primitives of faces, if rearranged, cease to be perceived as coherent 3D objects. These findings strongly support the view that the interpretation of Mooney images is guided primarily by top-down processes.

In this paper we investigate this question from a mathe-

[1]The vision group at the Weizmann Institute is supported in part by the Moross Laboratory for Vision Research and Robotics. BN was supported by the Lord Sieff of Brimpton Memorial fund.

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

matical standpoint and ask whether a unique 3D shape can be recovered from a single Mooney image. As Mooney patterns reflect the interaction of lighting and shape in the image it is natural to approach this problem using Shape from Shading (SFS) techniques. In this view the main information available in a Mooney image is in the shape of the *Mooney transition curve*, i.e., the boundary between dark and bright regions in the image. This curve is in fact an isoluminance curve. We therefore investigate whether 3D shape can be recovered *along* the Mooney transition curve. By investigating the SFS equations we show that even along this isoluminance curve reconstruction is not unique. We consequently turn to algorithms that use prior knowledge of faces [6, 14] to reconstruct the 3D shape of faces from a single Mooney image. Our experiments show that these methods manage to work surprisingly well on these impoverished images. This is encouraging both from a cognitive perspective, and also from a practical standpoint, since Mooney images resemble images obtained under extreme lighting conditions (e.g., with saturated lighting in a dark environment).

Very few computational algorithms have been proposed to handle Mooney images. Most notably, Shashua [24] introduced a method for face recognition from a single Mooney image in a fixed pose and unknown lighting. Assuming the face is Lambertian and in the absence of attached shadow the set of images of the face under different lightings is 3-dimensional. A two-tone image provides at every pixel an inequality that is linear in both the components of the lighting vector and the threshold value, and so these parameters can be recovered by solving a linear program. This method, however, requires a 3D model of the specific individual to be identified in the image. (The method however can in principle be extended to any linear model.) A more recent method [16] too uses linear models for object recognition from Mooney images. This method finds initial values for the coefficients and the threshold by solving a linear system along the Mooney transition curve, and then uses an iterative computation to refine these values. None of these methods however addresses the problem of reconstructing a novel face from a Mooney image.

In the remaining sections we first investigate the uniqueness of reconstruction from a Mooney image (Section 2). Later on we introduce algorithms for reconstruction of 3D faces from single Mooney images and present experimental results (Section 3).

2. Non-Uniqueness of Reconstruction

Monocular gray level and color images contain various cues from which a 3D shape can be inferred. These include perspective cues, texture distribution, bounding silhouettes, and shading cues. As two-tone images are much more impoverished, most of these cues are rarely available, and the main cue that remains is shading. In two-tone images only a single bit of information is available at every pixel, marking whether the observed surface is either bright or dark due to either the color of the surface, whether the surface faces toward or away from the light, or whether the surface is in shadow. It is natural therefore to analyze the problem of reconstruction from a Mooney image as a shape from shading (SFS) problem with impoverished data.

The question of uniqueness arises in SFS also when gray level images are considered [5, 19]. In general, the same image can be produced by many different 3D shapes, and so typically SFS algorithms make the stringent assumptions that the lighting configuration is known, the reflectance properties of the object are specified, the color (albedo) of the object is known, and the appropriate boundary conditions (i.e., depth values at certain locations on the object) are given. Under these conditions SFS is unique up to translation in depth and reflection (convex/concave) ambiguity. Below we make similar assumptions. We assume a single point light source whose direction and magnitude are known, the observed surface is Lambertian, and albedo is uniform (or otherwise known). Our formulation also allows the introduction of boundary conditions. As a Mooney image is impoverished we focus on the Mooney transition curve, i.e., the boundary curve between black and white in the image, as we explain below.

As an example consider a two-tone image of a Lambertian sphere illuminated by a frontal light source seen on a bright background. It is easy to construct other simple surfaces that are consistent with this image, such as an axial ellipsoid with equal $x-$ and $y-$ axes, or a cone. Note however that both these surfaces are circularly symmetric about the optical axis, and so they all give the same z=const reconstruction along the Mooney transition curve. An interesting question, therefore, is whether reconstruction is still ambiguous if we restrict our attention to the Mooney transition curve. If such a reconstruction is unique, we might be able to derive shape invariants along the curve and use them to recognize the shape in a bottom-up fashion.

2.1. General treatment of ambiguity

Consider a gray level image $I(x, y)$ of a smooth Lambertian surface $z(x, y)$ with uniform albedo illuminated by a directional source $\mathbf{l} \in \Re^3$. The image intensities are given by $I = \mathbf{l}^T\mathbf{n}$, where $\mathbf{n}$ denotes the surface normal at each point, $\mathbf{n} = (1/\sqrt{z_x^2 + z_y^2 + 1})(-z_x, -z_y, 1)$. A two-tone image is obtained from I by applying a threshold $I \geq T$ for some constant $T > 0$. Below we assume without loss of generality that T is known, and that the light source coincides with the viewing direction, i.e., $\mathbf{l} = (0, 0, 1)$. Note however that our analysis can be applied to any directional source by a change of coordinates, as in [15], and the mag-

nitude of the light can be scaled by appropriately scaling T. With these assumptions we obtain

$$I(x,y) = \frac{1}{\sqrt{z_x^2 + z_y^2 + 1}}, \tag{1}$$

which can be rewritten in the form of an Eikonal Equation

$$|\nabla z|^2 = E(x,y) \tag{2}$$

on some closed domain $\Omega \subset \Re^2$, where $E = (1/I^2) - 1$.

In general, given a gray level image I, the corresponding eikonal equation can be solved by various methods, for example by upwind updates using a Dijkstra-like algorithm (*e.g.*, the fast marching method [15, 23, 27]). These solutions rely on appropriate boundary conditions. In general, upwind solutions require Dirichlet boundary conditions in which z is specified at every minimal point of E (maximal point of I) in Ω. These may include local minima, as well as any minimum point along the boundaries of Ω. Our analysis therefore will allow the introduction of boundary conditions.

Consider two surfaces z and z' that give rise respectively to two images I and I' (and as before respectively to E and E') which are "Mooney equivalent." By this we mean that $|\nabla z|^2 = E(x,y)$ and $|\nabla z'|^2 = E'(x,y)$ and $I = I'$ =const along an isoluminance curve γ. Some boundary conditions may also be specified, so that $z = z'$ (and at internal points also $|\nabla z|^2 = |\nabla z'|^2 = 0$) in some set $\mathcal{B} \subset \Omega$. Denote by $\alpha(x,y) = z' - z$, our goal is given z to find the possible assignments of α.

By subtracting the two eikonal equations for z and z' we obtain a new eikonal equation in α

$$|\nabla \alpha|^2 + 2\nabla\alpha \cdot \nabla z = E' - E. \tag{3}$$

To solve for α we introduce a change of coordinates $(x,y) \rightarrow (t,s)$ such that $\alpha_s = 0$ and $\alpha_t \neq 0$. t therefore points in the gradient direction of α, and this will be useful since the gradient direction is also the characteristic direction of (3). In this coordinate frame (3) is transformed to

$$\alpha_t^2 + 2z_t\alpha_t - (E' - E) = 0. \tag{4}$$

(4) is quadratic in α_t, and its solution is given by

$$\alpha_t = -z_t \pm \sqrt{z_t^2 + E' - E}. \tag{5}$$

From this equation we can derive a general solution for α in the entire domain Ω by integrating (5) with respect to t along the characteristic directions as follows

$$\begin{aligned}\alpha(t,s) = &- z(t,s) + z(t_0,s) \\ &\pm \int_{t_0}^{t} \sqrt{z_t^2 + E' - E}\, dt,\end{aligned} \tag{6}$$

where the point $(t_0, s) \in \mathcal{B}$. It can be readily verified that indeed $\alpha(t_0, s) = 0$.

Unfortunately, given a Mooney image we cannot use (6) to recover α since in general E and E' are unknown. However, along the transition curve, γ, we know that $E = E'$, and so (5) implies

$$\alpha_t|_\gamma = -z_t \pm z_t \in \{0,\ -2z_t\}. \tag{7}$$

To produce a non-trivial ambiguity we are generally only interested here in the negative solution $\alpha_t|_\gamma = -2z_t$. Since $\alpha_s = 0$, the positive solution $\alpha_t|_\gamma = 0$ if applied throughout γ leads to the trivial solution α =const, and so along γ: $z' = z$+const. As we are interested in smooth solutions, we can restrict our search to ambiguities that satisfy the negative solution $\alpha_t|_\gamma = -2z_t$.

Note that (5) and (7) can be used to explain uniqueness in the general SFS problem for gray level images. In that case $E = E'$ at every location in Ω, and the negative solution corresponds to the well-known convex-concave reflection ambiguity. In principle either of the two solutions can be selected independently at every point, but if we consider only smooth solutions such transitions can only occur at places in which $|\nabla z|^2 = 0$, and those transitions can be eliminated by supplying appropriate boundary conditions. In our case, however, we know only that $E = E'$ along γ.

While given a Mooney image we cannot compute α in the entire domain Ω, we can nevertheless still provide an explicit solution for α along the transition curve γ. Let σ be an arclength parameterization of γ, and let $\theta(\sigma)$ denote the angle between the tangent to γ and the t direction, we obtain

$$\alpha|_\gamma = -2\int_\gamma z_t \cos\theta d\sigma + \alpha(\sigma_0). \tag{8}$$

Eq. (8) implies that if we choose some characteristic directions for α along the transition curve γ then there will be exactly two shapes along this curve that will be consistent with the Mooney image, z and $z' = z + \alpha$. However, unlike in the general SFS problem, in the case of a Mooney image we in general are free to choose different characteristic directions along γ and thus produce many additional solutions. The constraints on α are therefore that it must be consistent with the boundary conditions in $\mathcal{B}$, if such conditions are given, and that its gradients coincide with *some* smoothly varying directional derivatives at points along γ. This implies that in general many ambiguities exist even if we restrict our attention to the Mooney transition curve only.

2.2. Ambiguity examples

In this section we demonstrate ambiguities for the simple Mooney pattern shown in Fig. 2 (bottom right). Consider the unit sphere $x^2 + y^2 + z^2 = 1$. We will use polar

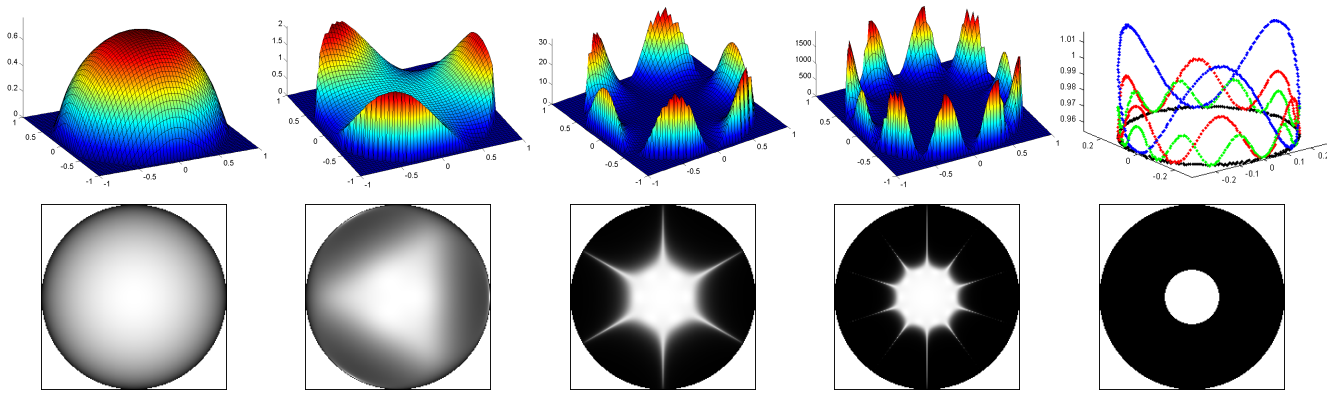

Figure 2. Surfaces whose Mooney image is identical to that of a sphere. Top: (from left to right) a sphere, three surfaces that produce the same Mooney pattern, and 3D reconstruction of the four corresponding γ curves. Bottom: original gray-level images of the four shapes and the common thresholded image. To produce the ambiguity surfaces we used $f(r) = r^q$ with $q \in \{3, 6, 10\}$ and $r_0 = 0.3$.

coordinates, $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$, so that $z = \sqrt{1 - r^2}$. The gradient of z is therefore given by

$$z_r = \frac{-r}{\sqrt{1 - r^2}} \qquad z_\theta = 0, \tag{9}$$

and its magnitude is $E = r^2/(1 - r^2)$. We further consider boundary conditions at the origin, with $z(0, \theta) = 1$ and $z_r(0, \theta) = 0$ for all θ. Let the curve γ consists of the circle $r = r_0$ and consider a point p on this circle. Let the t direction intersect p at some angle ϕ relative to the normal to γ. Then the directional derivative of z at p is

$$z_t(r_0, \theta) = \frac{-r_0}{\sqrt{1 - r_0^2}} \cos\phi. \tag{10}$$

Now, using (7) the gradient of α at p is

$$\alpha_t(r_0, \theta) = -2z_t(r_0, \theta) = \frac{2r_0}{\sqrt{1 - r_0^2}} \cos\phi. \tag{11}$$

Projecting the gradient onto the two polar directions yields

$$\alpha_r(r_0, \theta) = \frac{2r_0}{\sqrt{1 - r_0^2}} \cos^2\phi \tag{12}$$

$$\alpha_\theta(r_0, \theta) = \frac{2r_0^2}{\sqrt{1 - r_0^2}} \cos\phi \sin\phi, \tag{13}$$

where the factor r_0 is squared in (13) due to the use of polar coordinates.

Our objective is to find a function $\alpha(r, \theta)$ such that $\alpha(0, \theta) = \alpha_r(0, \theta) = 0$ at the origin and whose polar derivatives on γ assume the values above. We can achieve this by setting:

$$\alpha = cf(r)\cos^2(k\theta). \tag{14}$$

for some smooth function $f(r)$ and constants c and k. Consequently,

$$\alpha_r = cf_r \cos^2(k\theta) \tag{15}$$

$$\alpha_\theta = -2ckf(r)\cos(k\theta)\sin(k\theta). \tag{16}$$

The boundary conditions imply $f(0) = f_r(0) = 0$, and the conditions on γ imply

$$c = \frac{2r_0}{\sqrt{1 - r_0^2}} \frac{1}{f_r(r_0)} \tag{17}$$

$$k = -\frac{r_0^2}{\sqrt{1 - r_0^2}} \frac{1}{cf(r_0)} = -\frac{r_0 f_r(r_0)}{2f(r_0)}. \tag{18}$$

There are many ways to set $f(r)$ to satisfy these equations. One simple way is to set $f(r) = r^q$ for $q \geq 2$. In that case $c = 2/(qr_0^{q-2}\sqrt{1 - r_0^2})$ and $k = -q/2$. Fig. 2 shows several ambiguities from this family.

All the ambiguities in this family yield surfaces whose Mooney images are indistinguishable from that of the sphere, while their shape along γ oscillates at any desired frequency. However, these ambiguities in fact introduce new maximal values near the boundary of the sphere, and so in principle one may expect to obtain the height of these maximal points as part of the boundary conditions. Unfortunately, a local analysis of the type presented in Section 2.1 cannot prevent the emergence of new maxima. However, in general the constraints on α are such that in many cases it may still be possible to produce ambiguities that will not produce new maxima.

While it is possible to construct ambiguities that adhere to the supplied boundary conditions and perhaps even avoid introducing new extremal points it is worth noting that obtaining boundary values for Mooney images is arguably very difficult. While for gray level images we may simply detect the local maxima of the image, in Mooney images those maxima can lie anywhere within the bright regions in the image. Moreover, while in gray level images the surface boundaries can be identified as discontinuities, in Mooney images they may either blend with the background or be indistinguishable from internal transitions from black to white.

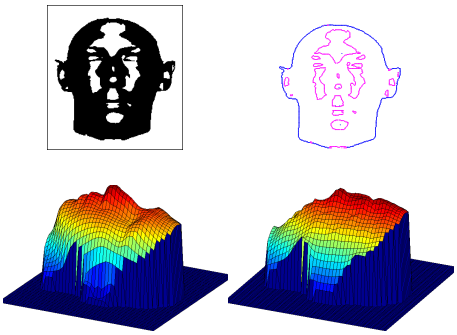

Figure 3. A smooth reconstruction of a face shape from a single Mooney image using extensive boundary conditions. The figure shows the mooney image (top left), a plot of the location of boundary conditions (in blue) and the transition curve (in magenta) contours (top right), initial surface (ground truth, bottom left) and the surface obtained with this reconstruction (bottom right).

Finally, we have sought to construct an ambiguity surface given a Mooney image of a face. To this end we have implemented an iterative method that constructs smooth shapes while adhering to boundary conditions and to constraints on the gradient magnitude. Our implementation does not restrict the direction of the gradient at γ, but only its magnitude. The result is shown in Figure 3. As expected, the overall shape of the face was preserved, due to the extensive use of boundary conditions, but most of the significant features were washed away.

3. 3D Reconstruction

In the previous section we showed that 3D reconstruction from single Mooney images is generally not unique even if we restrict our attention to the Mooney transition curve. This further supports psychological findings that the percept of shape from a Mooney image is possible only if prior knowledge is introduced. In the case of faces, such prior knowledge may include for example the 3D structure of known faces. In this section we describe how existing 3D face reconstruction methods that use prior knowledge of faces can reconstruct the shape of faces from single Mooney images. We consider two types of methods: a method for 3D face reconstruction that solves a shape from shading equation and uses a single reference model and a method that uses linear combinations of a database of faces.

3.1. Reconstruction with a single reference model

The first method, presented by Kemelmacher and Basri [14] (denoted KB), was designed to recover 3D shape and albedo of faces from single color or gray scale images using a single reference model of a different individual's face. The method allows for multiple unknown light sources and attached shadows by representing image intensities using a spherical harmonics approximation to reflectance ([3, 20]). Specifically it uses the first order of approximation such that image intensities can be represented as a linear combination of the components of the surface normals,

$$I(x,y) \approx \rho \mathbf{l}^T \mathbf{Y}(\mathbf{n}), \tag{19}$$

where $\mathbf{l} = (l_0, l_1, l_2, l_3)^T$ denote the harmonic coefficients of lighting, $\mathbf{Y}(\mathbf{n}) = (1, n_x, n_y, n_z)^T$ and n_x, n_y, n_z are the components of the surface normals $\mathbf{n}(x,y)$ at each point. Reconstruction is achieved by first roughly aligning the image with respect to the reference model by matching a small number of feature points and then optimizing the following functional

$$\min_{\mathbf{l},\rho,z} \int_{\Omega} \left(\left(I - \rho \mathbf{l}^T \mathbf{Y}(\mathbf{n}) \right)^2 + \lambda_1 \triangle g d_z + \lambda_2 \triangle g d_\rho \right) dxdy, \tag{20}$$

where the first term in the functional is the data term (Eq. 19) and $\triangle g d_z, \triangle g d_\rho$ are two regularization functions that maintain the difference between the reconstructed surface and the reference model smooth. This optimization is carried out in three steps in which lighting, shape, and albedo are recovered sequentially. In the initial step we solve for lighting $\mathbf{l}$ by fitting the reference model to the image. After that the main step, recovering $z(x,y)$, is carried out using the formula

$$I \approx \rho_{\text{ref}} (l_0 + \frac{1}{\sqrt{p_{\text{ref}}^2 + q_{\text{ref}}^2 + 1}} \tilde{\mathbf{l}}^T (p, q, -1)^T) \tag{21}$$

where $\tilde{\mathbf{l}}^T = (l_1, l_2, l_3)$, $p = z_x$ and $q = z_y$. This provides one equation for every unknown, and with appropriate regularization and boundary conditions it is solved directly using least squares. (See Kemelmacher and Basri [14] for further details).

A Mooney image is obtained from I by choosing some constant $T > 0$ and then by thresholding the image $I(x,y) \geq T$. In this case every equation in (21) becomes an inequality. Thus in the main step we should now solve a system of inequalities combined with regularization and boundary conditions. To avoid the instabilities that can arise from the use of inequalities we used the same least squares approach as in [14] to solve the optimization, with the Mooney image used for I. Constructing a smooth surface from a binary image requires regularization; otherwise the discontinuity at the transition curve may result in an unsmooth reconstruction. We address this by modifying the regularization weights λ_1, λ_2 along the Mooney transition curves to prevent discontinuities.

3.2. Reconstruction by statistical models

A fair number of recent approaches use statistical models for 3D face reconstruction. These methods learn the set of allowable reconstructions from a large dataset of

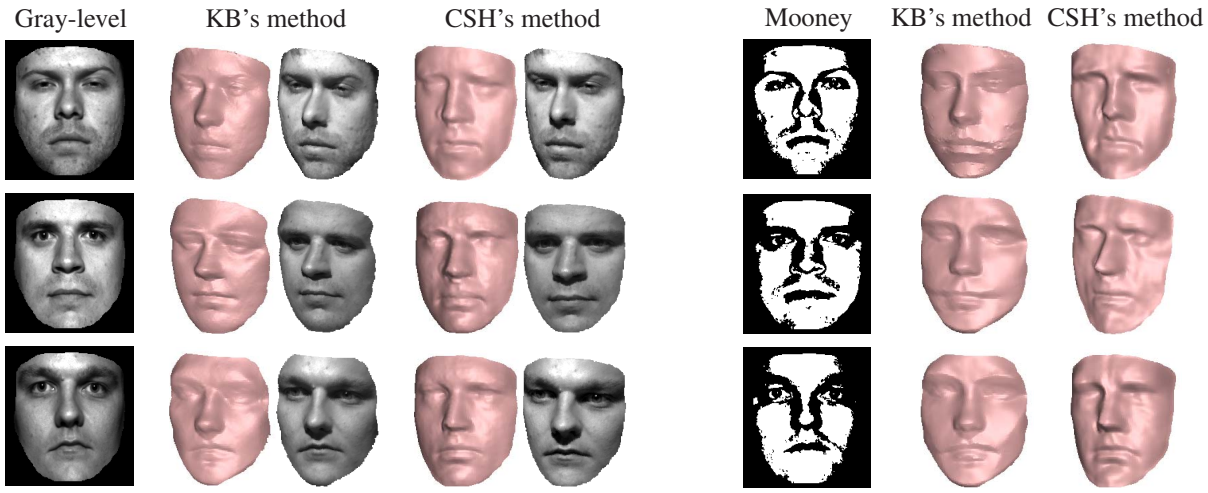

Figure 4. Reconstruction results with gray-level (columns $1-5$) and thresholded images (columns $6-8$) from the YaleB database. From left to right: gray-level input image, reconstruction results with KB's method (shape and shape+albedo) and with CSH's method, then we show Mooney input image, and again reconstruction results with both methods.

	KB's method	CSH's method
1	6.2°	12.1°
2	7.8°	11.8°
3	4.1°	5.6°

Table 1. Average angular error of shapes reconstructed from Mooney images in Fig. 4 relative to their reconstruction from gray scale images.

faces by either embedding all 3D faces in a linear space (e.g. [2, 4, 8, 26, 22, 30]) or by using a training set to determine a density function for faces [25, 29]. These methods can achieve accurate reconstruction, but they often require expensive alignment and parameter fitting. To simplify the process, Castelan *et al*. [6] (denoted CSH) proposed to combine surface shape and image brightness variations into a single, coupled statistical model. In the first step separate eigenspaces are constructed for image intensities and surface shape variations from the training data, and then these two spaces are combined into a single space. Then, given an image the coefficients that fit the model to the image is found and those coefficients, due to the coupling of brightness and shape in the model, can be used to recover shape parameters and consequently the shape of the face itself. The objective function obtained is linear and can be solved efficiently. Despite its simplicity, this model was shown to generally produce accurate surfaces from real images. This method however is limited in scope since it does not model lighting explicitly and cannot extrapolate to new lighting conditions. We will therefore apply this method to images of novel faces, but with lighting with which the method was trained.

One way to apply this model to Mooney images is to cast the problem as a system of linear inequalities (in the spirit of [24]), but due to the high dimensionality of the model this would allow a fair amount of slack. Instead, we again applied this method directly to the Mooney images. The shapes obtained with this method appeared bumpy therefore we applied gaussian smoothing to the outputs.

3.3. Results and discussion

For the CSH method we used for training 77 models from the USF database [12] acquired with a laser scanner. The database contains depth and texture maps of male and female adult faces with a mixture of race and age. For the KB method a mere one of the models in this training set was used as a reference model.

Figure 4 shows some reconstruction results. First we show reconstruction from gray level images taken from the YaleB database images [10]. Both methods produce nice reconstructions, with Castelan *et al*'s method producing some ringing effects near the nose (perhaps due to misalignments in the nose area). Then we show reconstructions

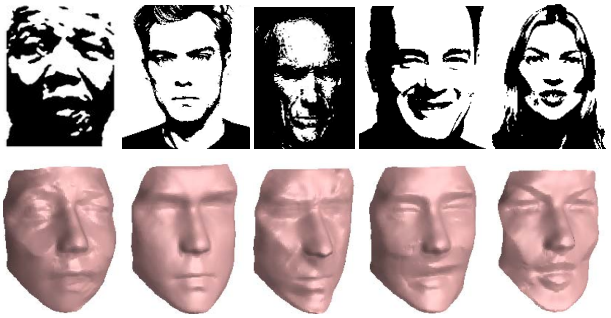

Figure 5. Reconstruction results. Top: input images. Bottom: face shape reconstruction using KB's method.

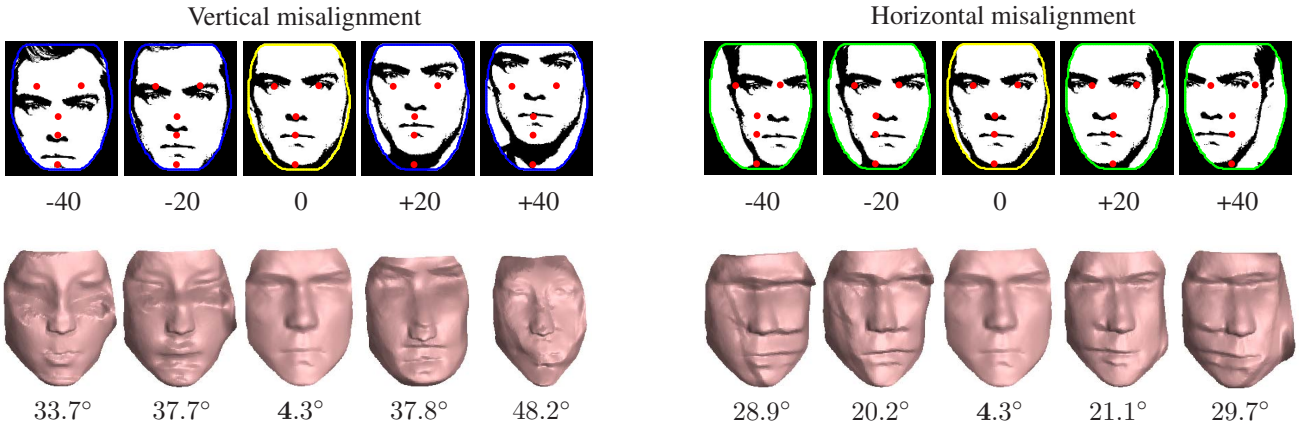

Figure 6. Reconstruction error with misplaced alignment points with KB's method. Top: masked two-tone images used for reconstruction. The red dots indicate the positions matched to five features of the reference model (eyes, nose, mouth, and chin). Vertical offsets are shown on the left and horizontal on the right. Numbers represent offsets in pixels. Second row: corresponding reconstructions. The numbers represent anglular error relative to reconstruction from a color image.

from Mooney images obtained by thresholding images from the YaleB dataset. To measure the error we compared the average of the angles between corresponding surface normals in each reconstruction relative to a reconstruction with the original gray scale image. Overall, both methods work surprisingly well producing reasonable reconstructions that have resemblance to their gray-level counterparts. This is further demonstrated by measuring the average angular error between the Mooney reconstructions and their gray-level counterparts (Table 1). We can observe that the KB's method is more consistent (the errors are smaller), however in both methods the error is relatively small (up to 12.1°).

Next, we applied KB's method to images downloaded from the Internet and thresholded by the same constant (except for the leftmost image, which was downloaded originally as a two-tone image). As CSH's method is not designed to work with lightings that are not available in the training dataset it did not perform well on these images and we do not include these results. Other PCA based methods might better deal with variations in lighting. Fig. 5 shows reconstruction results obtained with KB's method, these are very similar to reconstructions from color images. Note that the different features and facial expressions of each individual are clearly seen in these reconstructions.

When prior knowledge is used in reconstruction there is always the question of how dominant the prior knowledge is and how much it is modified by the data. This question is particularly relevant when impoverished data is used. This can be tested, for example, by applying reconstruction to non-face black and white patches. Here we implement this by applying KB's method to shifted copies of the image, i.e., to misplaced face patterns. Indeed, as we shift the position of the alignment features in vertical and horizontal directions the reconstruction results deteriorate fairly rapidly (Fig. 6). A plot of the error as a function of misalignment

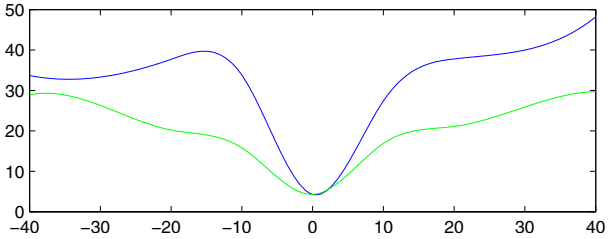

Figure 7. Angular error as a function of offset for vertical (blue curve) and horizontal displacements (green).

in pixels in presented in Fig. 7. A clear minimum in the error is obtained when the features are correctly aligned. The reconstruction quality slowly degrades with shifts of up to ± 5 pixels (10%) from correct alignment, which shows that the method is robust to small misalignments. Further, its

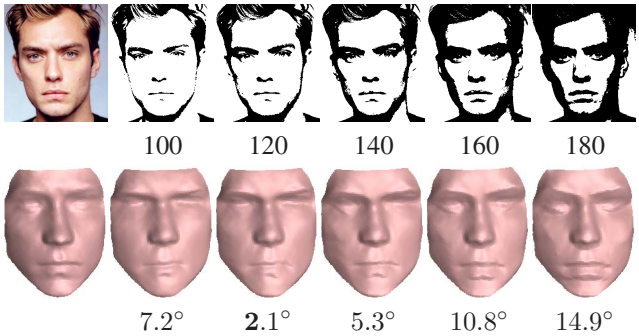

Figure 8. Reconstruction results and errors with different thresholds, obtained with KB's method. Original color image and two-tone images (top images) obtained with different threshold values (indicated below), and corresponding reconstructions (bottom images), along with angular errors relative to reconstructions from color image (shown on left).

degradation beyond that range indicates that the shape of the reference face is not imposed on the data. Finally, as Mooney faces can be produced from gray-scale images by using different thresholds, we demonstrate the effect of using different thresholds in Fig. 8.

4. Conclusion

Mooney images are fascinating testimonial to the ability of biological vision systems to accurately handle and interpret impoverished data. This paper provides additional support, from a mathematical perspective, to the widely held view in cognitive psychology that the perception of Mooney images is guided primarily by top-down processes in which prior knowledge plays a crucial role. Our main contribution in this paper is showing that shape reconstruction from a single Mooney image, under the standard Lambertian assumptions, is non-unique even if reconstruction is restricted to the Mooney transition curve alone. We have further demonstrated this by constructing families of ambiguities for simple shapes. Finally, we have discovered that recent face reconstruction approaches that use prior knowledge are capable of successfully recovering the 3D shape of faces from single Mooney images and showed novel results on real and synthetic data. These results may further encourage research in recognition and detection of Mooney faces in cluttered environments, provide insights on image quantization and may also suggest to psychophysicists concrete ways to test this fascinating human ability.

References

[1] T. Andrews and D. Schluppeck. Neural responses to mooney images reveal a modular representation of faces in human visual cortex. *Neuroimage*, 21:91–98, 2004. 1

[2] J. Atick, P. Griffin, and A. Redlich. Statistical approach to shape from shading: Reconstruction of 3d face surfaces from single 2d images. *Neural Comp.*, 8(6):1321–1340, 1996. 6

[3] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *PAMI*, 25(2):218–233, 2003. 5

[4] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. *SIGGRAPH*, I:187–194, 1999. 6

[5] A. Bruss. The eikonal equation: Some results applicable to computer vision. *J. of Math. Phys.*, 23(5):890–896, 1982. 2

[6] M. Castelan, W. Smith, and E. Hancock. A coupled statistical model for face shape recovery from brightness images. *Trans. on Image Processing*, 16(4):1139–1151, 2007. 2, 6

[7] R. Dolan, G. Fink, E. Rolls, M. Booth, A. Holmes, R. Frackowiak, and K. Friston. How the brain learns to see objects and faces in an impoverished context. *Nature*, 389:596–599, 1997. 1

[8] R. Dovgard and R. Basri. Statistical symmetric shape from shading for 3d structure recovery of faces. *ECCV*, 2004. 6

[9] N. George, B. Jemel, N. Fiori, and B. Renault. Face and shape repetition effects in humans: a spatio-temporal erp study. *Neuroreport*, 8(6):1417–1423, 1997. 1

[10] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *PAMI*, 23(6):643–660, 2001. 6

[11] J. Hegde, S. Thompson, and D. Kersten. Identifying faces in two-tone ('mooney') images: A psychophysical and fmri study. *J. of Vision*, 7(9):624, 2007. 1

[12] http://marthon.csee.usf.edu/HumanID/. 6

[13] N. Kanwisher, F. Tong, and K. Nakayama. The effect of face inversion on the human fusiform face area. *Cognition*, 68(1):B1–11, 1998. 1

[14] I. Kemelmacher and R. Basri. Molding face shapes by example. *ECCV, LNCS 3951*, I:277–288, 2006. 2, 5

[15] R. Kimmel and J. Sethian. Optimal algorithm for sfs and path planning. *J. of Math. Imaging and Vision*, 14(3):237–244, 2001. 2, 3

[16] J. Maver and A. Leonardis. Recognizing 2-tone images in grey-level parametric eigenspaces. *Pattern Recogn. Letters*, 23(14), 2002. 2

[17] C. M. Mooney. Age in the development of closure ability in children. *Canadian Journal of Psychology*, 11:219–226, 1957. 1

[18] C. Moore and P. Cavanagh. Recovery of 3d volume from 2-tone images of novel objects. *Cognition*, 67:45–71, 1998. 1

[19] J. Oliensis. Uniqueness in sfs. *IJCV*, 6(2):75–104, 1991. 2

[20] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object. *JOSA*, 18(10):2448–2459, 2001. 5

[21] E. Rodriguez, N. George, J. Lachaux, J. Martinerie, B. Renault, and F. Varela. Perception's shadow: long-distance synchronization of human brain activity. *Nature*, 397:430–433, 1999. 1

[22] S. Romdhani and T. Vetter. Efficient, robust and accurate fitting of a 3d morphable model. *ICCV*, 2003. 6

[23] J. Sethian. A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci. USA*, 93:1591–1595, 1996. 3

[24] A. Shashua. On photometric issues in 3d visual recognition from a single 2d image. *IJCV*, 21:99–122, 1997. 2, 6

[25] T. Sim and T. Kanade. Combining models and exemplars for face recognition: An illuminating example. *CVPR Workshop on Models versus Exemplars*, 2001. 6

[26] W. Smith and E. Hancock. Recovering facial shape and albedo using a statistical model of surface normal direction. *ICCV*, 2005. 6

[27] J. Tsitsiklis. Efficient algorithms for globally optimal trajectories. *Proc. Conf. on Decision and Control*, 1994. 3

[28] J. Yoon, J. Winawer, N. Wittoft, and E. Markman. Mooney image perception in preschool-aged children. *J. of Vision*, 7(9):548, 2007. 1

[29] L. Zhang and D. Samaras. Face recognition under variable lighting using harmonic image exemplars. *CVPR*, 2003. 6

[30] S. Zhou, R. Chellappa, and D. Jacobs. Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints. *ECCV*, 2004. 6