3D Shape Reconstruction from Autostereograms and Stereo

Ron Kimmel
Computer Science Department, Technion, Haifa 32000, Israel.

We study the problem of shape reconstruction from stereo images based on a weighted area minimization process of a depth function. As a simple example we present an efficient shape reconstruction from computer generated autostereograms. A minimal surface area based correlation is applied to accurately reconstruct the surface structure embedded first in one autostereogram image, and next in two or more 'stereo' images. The minimal area approach proved itself as a useful geometric measure in recent reconstruction and enhancement applications in computer vision and image processing. Here we develop a simplified version for the Faugeras and Keriven [5] stereo reconstruction model, and apply a weighted area measure as part of a solution to the correspondence extraction in the shape from stereo and the shape from autostereogram problems. The proposed schemes are computationally efficient and yield accurate 3D reconstructions for smooth as well as non-smooth surfaces.

1. INTRODUCTION

Shape from stereo is a classical computer vision problem in which we try to extract the '3D' shape of a scene from two or more pictures taken at two or more known camera positions. For almost three decades scientists have tried to develop automatic methods for three dimensional shape reconstructions from images. Shapes can be reconstructed from their shading image [6, 7], their texture, photometric stereo, structured light, defocus, motion sequences, zoom, etc. Some reconstruction approaches couple deterministic image formation models with stable numerical techniques, sometimes enforced by statistical properties. Early shape from stereo techniques in computer vision were tested on two random dot images, where the disparity was extracted by solving the correspondence problem between the specific points. Modern techniques try to balance between feature and point matching, where the feature scale depends on the locality of the feature, its validity, robustness to noise, and its similarity between the two images (robustness to deformations).

Before we explore the stereo problem and propose a solution, let us start with a 'toy' problem of shape construction from autostereogram that is similar to the

classical shape from stereo problem, yet simpler in some sense, Nevertheless, the expected quality of the results makes this problem a challenging task.

Autostereogram is a method for embedding 3D shapes in a single image that at first glance appears like a flat repeating pattern or noise. An observer can construct a three dimensional interpretation in his or her perception by matching picture elements along horizontal lines from the image plane. Autostereograms were invented by C. W. Tyler [15, 14, 16] as a tool for psychophysics research and the understanding of the human vision. Such images became part of popular art and appear in posters and books like 'The Magic Eye' [13].

There are simple algorithms for generating autostereograms as we present in Section 2. The basic idea is to scan a given hight array horizontally and duplicate pattern elements from the image plane that are located at a distance inversely proportional to the hight profile at the relevant point location [16]. An initial pattern is needed as initialization. Bruckstein et al. [3], controlled the spectral behavior of this initial pattern in order to generate autostereograms that are 'easy to interpret'.

Let us first deal with this simple problem, the problem of 3D shape reconstruction from computer generated autostereograms. The simpler the problem, the better the accuracy of the 3D reconstruction should be. Our goal is to show that geometric smoothness eliminates noise and echos, and allow an almost perfect three dimensional reconstruction. We start with a simple correlation procedure that solves the local correspondence problem of matching between distant points in the image domain. The local nature of this straightforward correlation yields noisy reconstructions with repeating echoes near the boundaries. Interestingly, we had similar visual misinterpretations that correspond to common perceptual mistakes of human observers. See [12] for discussions on geometric distortions like echoes formation, and incorrect convergence due to false fusion.

A minimal area based correlation was recently used by Faugeras and Keriven [5] to solve the stereo problem, in which two or more images from different given locations of the same object are used to reconstruct its 3D shape. A dynamic surface that serves as a virtual correlation interface between the different images locks onto the best correlation position that corresponds to the 3D reconstruction. Motivated by Faugeras and Keriven we first present a related, yet simpler, reconstruction algorithm for 3D shapes embedded in autostereograms. It is based on correlations between picture elements that result in a first estimation of the surface. A minimal area based approach is then applied to enhance and smooth the first naive correlation result. Next, we explore the shape from stereo problem and introduce an efficient iterative approaches to solve the weighted area formulation of the problem.

The minimal weighted area measure is chosen because of several reasons. 1. It is a geometric measure which is invariant to the choice of coordinates and is reparameterization invariant. 2. Most of the surfaces in nature tend to be minimal with respect to some geometric measure, so there is a connection to our natural physical world. 3. We can use implicit representations of the surface and design stable numerical schemes that work on fixed regular grids and efficiently find the surface of minimal weighted area.

This paper is organized as follows: In Section 2 we review a simple algorithm for generating random dot autostereograms. Next, Section 3 presents a minimal area

based reconstruction procedure that filters out the reconstruction noise. Section 4, generalizes the area idea and introduces the minimal weighted area framework for the stereo problem. The proposed methods are applied both to synthetic and real stereo data.

2. GENERATING AUTOSTEREOGRAMS

Given the 3D hight profile as a function $z(x, y)$ sampled on a discrete grid (e.g. grid size of $M \times M$), such that $0 < Z(i, j) < N$ (e.g. $M = 400, N = 70$) where $Z(i, j) = z(i\Delta x, j\Delta y)$, and a strip of random noise or a given pattern $R(i, j)$ of size $N \times M$, a simple autostereogram image synthesis algorithm is given as follows:

for $\{j = 0; j < M; j++\}$ for $\{i = 0; i < M; i++\}$

if $\{i < N\}$ $I(i, j) = R(i, j)$; else $I(i, j) = I(i - N + Z(i, j), j)$;

The artificial disparity between patterns is interpreted as depth, such that the larger the distance between repeating patterns the further a point is perceived in the eyes of the observer, see Figure 1. A symmetric algorithm, with somewhat better performances, was introduced by Thimbleby et al. in [12]. Most autostereograms are generated in a similar way, by swiping the disparity relation between pixels as a constraint. The reconstruction algorithm can be considered as one measure for the quality of the generated autostereogram.

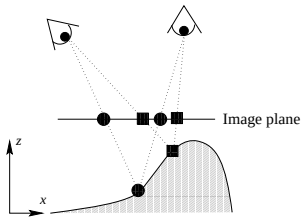

FIG. 1. The distance between repeating patterns is interpreted as depth.

3. SHAPE RECONSTRUCTION FROM AUTOSTEREOGRAM

In case the strip $R(i, j)$ does not include repeating colors for each line, i.e. $R(i, j) \neq R(i - k, j)$, $\forall i \neq k$, and there is a unique k such that $0 < k \leq N$ and $I(i, j) = I(i - k, j)$, then there are no ambiguities in the reconstruction. The shape reconstruction procedure then goes as follows:
For every (i, j), such that $0 \leq j \leq M - 1$, and $N \leq i \leq M - 1$, find the unique $k(i, j)$ such that $0 < k \leq N - 1$ and $I(i, j) = I(i - k(i, j), j)$, that yields the 3D shape given as the function $Z(i, j) = i - k(i, j)$.

When the random strip fails to satisfy the above demand, or as i increases and the number of random values decreases, more than one correlation may occur. In this case we define a vector of all possible Z values for each point (i, j) and initiate the reconstructed surface to be the largest Z at that point (determined by the distance to the closest pixel that correlates with the one under inspection). This value corresponds to the minimal k such that $0 < k$, and $I(i, j) = I(i - k, j)$. The surfaces reconstructed this way are noisy as shown in the middle column of Figure 4.

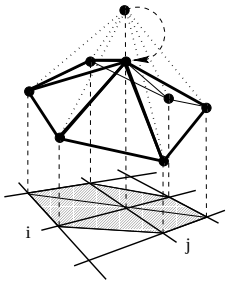

FIG. 2. The area of the six triangles is minimized by changing the $Z(i,j)$ candidate.

In order to overcome the noise problem we add the assumption that the surface is smooth. This assumption leads us to apply a minimal area rule that selects a smoother result subject to the correlation requirement. For each point we measure the total area of the six neighboring triangles as shown in Fig. 2. The Z value of the point is replaced by a different candidate from the vector of possible heights if the total area is reduced.

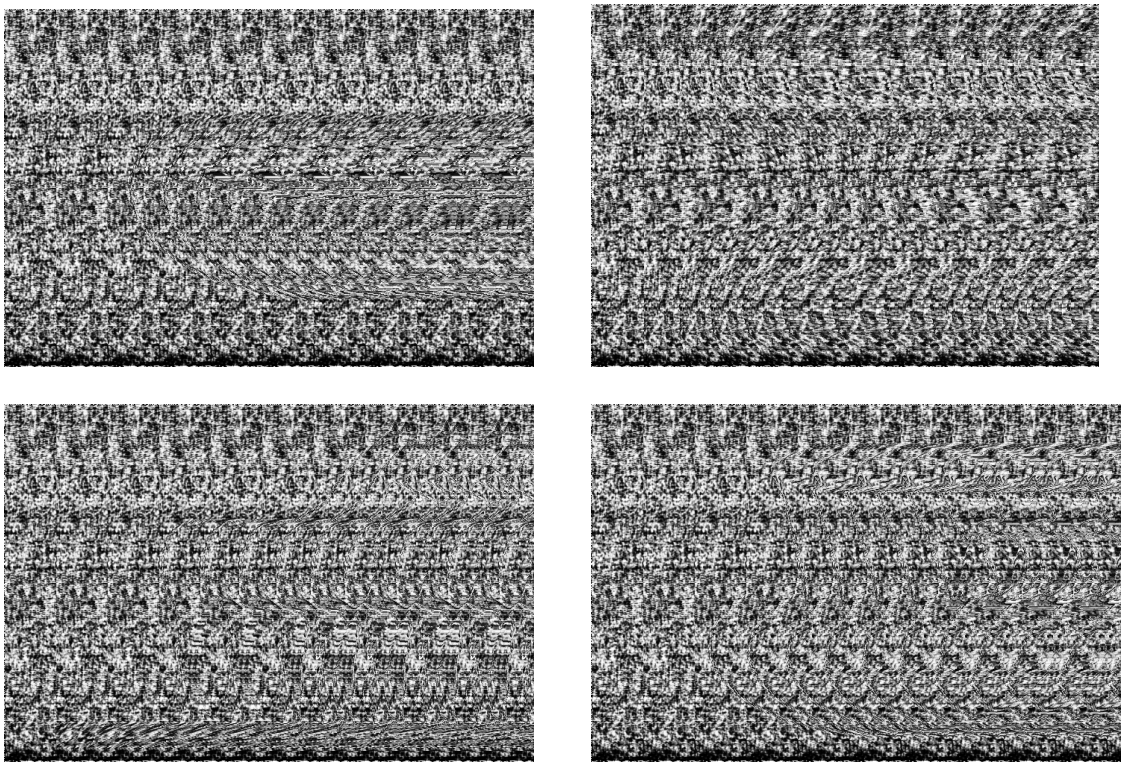

FIG. 3. Random dot autostereograms

This area minimization procedure causes the total area of all the triangles to strictly decrease at each update step. We also know that the total area is bounded

from below by the area of the flat image plane, and that each update step reduces the total area by more than a constant that can be a priorly determined. Thereby, we have convergence of our minimal area correlation algorithm in finite update steps. One scan over the domain is practically enough to result in a smooth reconstruction with the above initial guess.

Figure 3 shows the random dot autostereogram test images of six 3D profiles. In all cases the width of the pattern used for the initial random dot strip is $N = 70$, and its spectral behavior is $1/f^2$, see the 'easy patterns' in [3]. The hight resolution is set to 30.

Figure 4, left to right are: The original surface as gray levels (before embedding as an autostereogram), the reconstruction by the first correlated pixel, and the correction by minimal surface area correlation.

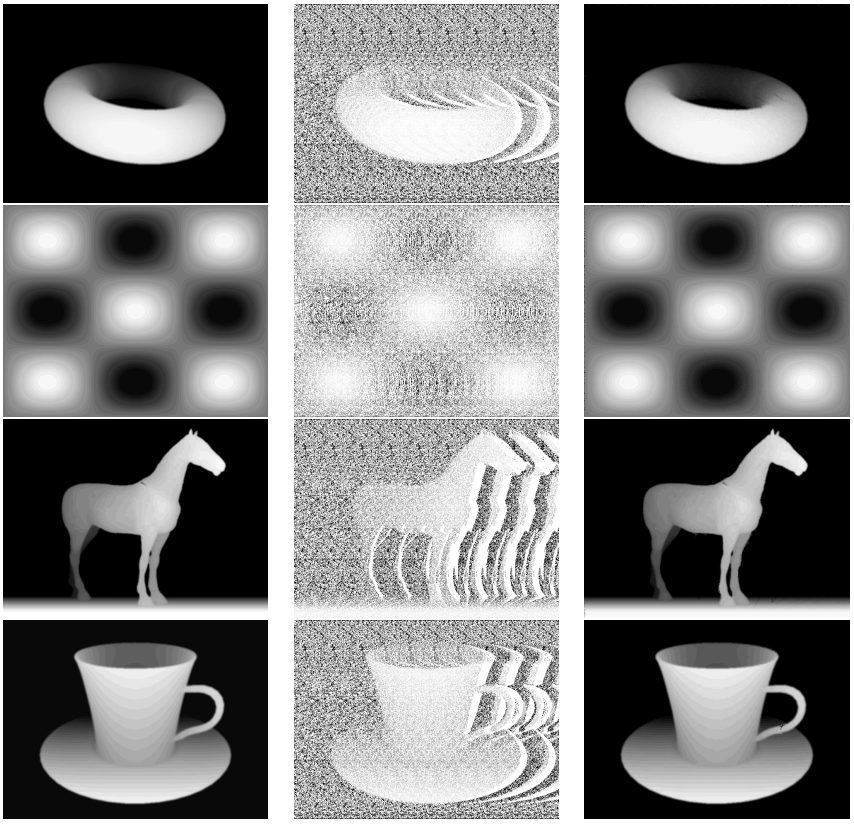

FIG. 4. Left to right: Original surface, first correlation, minimal surface refinement.

4. SHAPE FROM STEREO AS A SURFACE OF MINIMAL WEIGHTED AREA

Let us formulate the stereo problem as a weighted area minimization problem. Given the implicit surface representation $\mathcal{S} = \{(x, y, z)|\phi(x, y, z) = 0\}$, we have that $\min \int g(x, y, z)da$,where da is an area element, may be expressed as $\min \int g(x, y, z)|\nabla\phi|dxdydz$, (see the appendix for a proof) for which the Euler La-

grange (EL) equation is given by

$$\mathrm{div}\left(g(x,y,z)\frac{\nabla\phi}{|\nabla\phi|}\right)=0. \tag{1}$$

Where $\nabla\phi \equiv (\phi_x,\phi_y,\phi_z)$, and $\mathrm{div} \equiv (d/dx, d/dy, d/dz)$. Geometrically, we have the equivalent expression for the level sets of ϕ, given by

$$(g(x,y,z)H - \langle\nabla g(x,y,z),\vec{N}\rangle)\vec{N}=0, \tag{2}$$

where $\vec{N}=\frac{\nabla\phi}{|\nabla\phi|}$ is the normal of the level set surface of ϕ, and $H=\mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)$ is the mean curvature of the level set surface.

The main problem with this implicit representation is the relatively large numerical support needed to enjoy the fixed coordinates formulation. This usually leads to high computational cost when using the Euler Lagrange as a gradient descent procedure. In order to overcome the problem let us define the same geometric measure, yet now for the graph-surface $\mathcal{S}=(x,y,z(x,y))$. The weighted area minimization is given formally by $\min_z E(z)$ where,

$$E(z)=\int g(x,y,z(x,y))da=\int g(x,y,z(x,y))\sqrt{1+z_x^2+z_y^2}dxdy. \tag{3}$$

Taking the first variation with respect to z, yields

$$\frac{\delta E}{\delta z}=\mathrm{div}\left(g(x,y,z(x,y))\frac{\nabla z}{\sqrt{1+z_x^2+z_y^2}}\right)-g_z\sqrt{1+z_x^2+z_y^2}, \tag{4}$$

where now, $\mathrm{div} \equiv (d/dx, d/dy)$, and $\nabla z \equiv (z_x, z_y)$. This first variation equation should obviously has the same geometric meaning as above. It is given explicitly by the Euler Lagrange (EL) equation

$$0 \;=\; gH+\frac{g_xz_x+g_yz_y-g_z}{\sqrt{1+z_x^2+z_y^2}}, \tag{5}$$

where now the mean curvature of the graph surface is given by $H=\mathrm{div}\left(\frac{\nabla z}{\sqrt{1+|\nabla z|^2}}\right)$.

Next, we use our freedom of parameterization, multiply by $\sqrt{1+|\nabla z|^2}$, and end up with the same geometric equation $(gH-\langle\nabla g,\vec{N}\rangle)\vec{N}=0$, where the normal of the graph surface is $\vec{N}=\frac{(-z_x,-z_y,1)}{\sqrt{1+|\nabla z|^2}}$, and now $\nabla \equiv (d/dx, d/dy, d/dz)$ as before.

We can write the EL as a steepest descent flow that minimizes the weighted area by the flow equation

$$z_t=g\frac{(1+z_y^2)z_{xx}-2z_xz_yz_{xy}+(1+z_x^2)z_{yy}}{1+z_x^2+z_y^2}+g_xz_x+g_yz_y-g_z. \tag{6}$$

This formulation is inspired by Faugeras and Keriven level-set based stereo reconstruction model [5], however there are some fundamental differences. The above

method operates on a function and as such is restricted to be a function. It is more efficient from a computational point of view, since while working with fixed coordinate system we are processing a single surface function rather than the 3D space in which the level set surface is embedded or handling a narrow band around the surface with effective width around it. This formulation also allows us to introduce discontinuities of the reconstructed function by including a Mumford-Shah like penalty for discontinuous contours. It can be implemented by the Γ-convergence methods as in [1, 2, 10, 8]. The function restriction stabilizes the solution and guarantees uniqueness in case boundary conditions are provided. However, unlike Faugeras and Keriven, the surface is kept a graph and its topology is thereby restricted to be a function, which is a limitation. The correspondence function we use does not take into consideration the geometry of the reconstructed surface, as done in [5]. While it appears like a limitation, this simplification actually provides us with the powerful property of convergence into a stable solution as proved in [4], and practically we did not experience any dramatic drop in the performances.

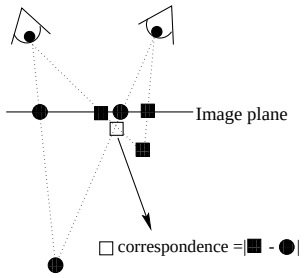

FIG. 5. The correspondence induces a position dependent scalar function in 3D space. The question is how to 'engrave' a minimal weighted area from this isotropic non-homogeneous space.

4.1. Reconstruction from stereo: Results

Figure 6 shows a reconstruction result from a stereo pair of a smooth synthetic hight profile. The initial conditions are set by a greedy process for each coordinate on the image plane. Next, a spatial median smoothing filter followed by a weighted minimal area minimization approach, brings the reconstructed surface very close to the original profile.

Next, we tested our scheme on real images. Here we applied a version of the correspondence measure discussed in [9, 11]. A correspondence disparity measure is assigned to each pixel (w.l.o.g) of the right image such that

$$\begin{aligned} SE(x,y,d) &= G_\sigma * (I_L(x-d,y)) - I_R(x,y))^2 \\ \bar{I}_L(x,y) &= G_\sigma * I_L(x,y) \\ Var_L(x,y) &= G_\sigma * (I_L(x,y))^2 - (\bar{I}_L(x,y))^2 \end{aligned}$$

$$Corr(x,y,d) = \frac{(SE(x,y,d))^2}{1 + Var_R(x,y) Var_L(x-d,y)}, \qquad (7)$$

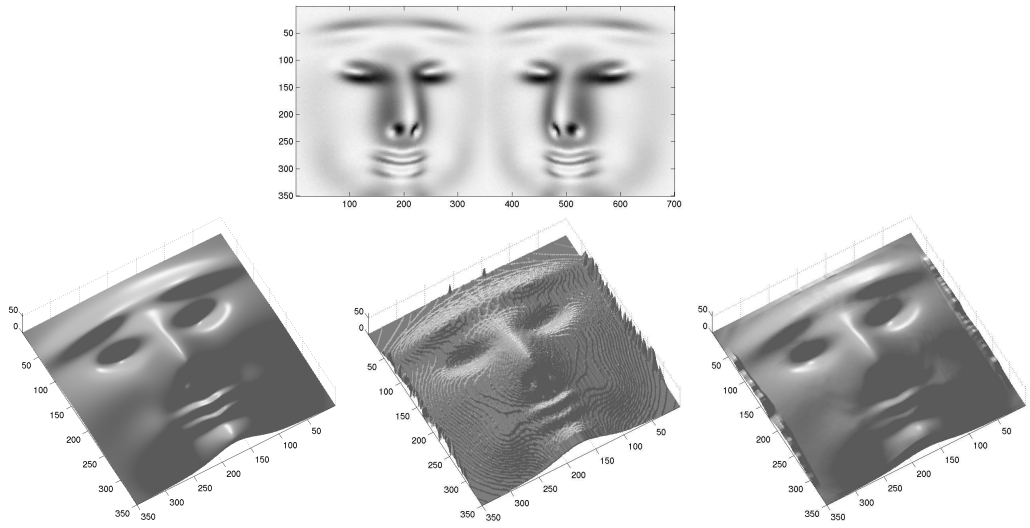

FIG. 6. Up: The two stereo images projected onto the same plane. Bottom Left: Original surface. Middle: Reconstruction by rough best match for the correspondence function from the stereo pair. Right: Reconstructed surface after smoothing via minimal weighted area.

where d is the disparity, G_σ is a normalized Gaussian kernel with variance σ, and the $*$ is a convolution operator. This measure could be easily extended to more than two images.

The goal is again to extract a surface with minimal weighted area as above, yet now we first search in the x, y, d discretized space explicitly. That is, instead of minimizing $\int g(x,y)da$, through its Euler Lagrange equation, we first search for the minimal weighted area by iteratively changing the values of $z_{ij} \equiv z(i\Delta x, j\Delta y)$ at each point, so that the total discrete area $\sum_{ij} g_{ij} da_{ij}$, is always decreasing by the update, as we did for the autostereogram case. This discrete step accelerates the convergence. Finally, we apply the differential minimization by the Euler Lagrange as a steepest descent flow Eq. (6).

Figure 7 presents perspective views of the same reconstruction from two images before and after the weighted area minimization phase. The surface presented in this example is the right image textured mapped onto the reconstructed surface.

5. CONCLUSIONS

We experimented with new solutions for the shape from stereo family of problems. As a first example we explored a minimal discrete area correlation approach to solve the shape from computer generated autostereograms problem. The minimal correlation area approach is one quality indicator for computer generated autostereograms, as well as an example for the discrete area as a promising measure for this correspondence problem.

Next, motivated by [5] we used the more general weighted area, were the weight is defined by an isotropic non-homogeneous correspondence function, to solve the classical shape from stereo problem. Our tests indicate that the modulated area is

FIG. 7. Shape from stereo: Top: The stereo pair and their normalized addition. Middle: The reconstructed surface with the right image textured mapped onto it before the weighted area minimization phase. Bottom: Reconstructed surface after the weighted area minimization.

indeed a very promising geometric measure for this family of shape reconstruction problems.

APPENDIX

Let us prove that, given the implicit surface representation $\mathcal{S} = \{(x, y, z) | \phi(x, y, z) = 0\}$, we have that $\min \int g(x, y, z) |\mathcal{S}_u \times \mathcal{S}_v| du dv$, may be expressed as

$$\min \int g(x, y, z) |\nabla \phi| dx dy dz.$$

Where $\mathcal{S}(u, v)$ is a parametric representation of the surface, and the second minimization is restricted to the zero level set of ϕ. We also assume that w.l.o.g. $\mathcal{S}(u, v) = \mathcal{S}(u, v, \phi)|_{\phi=0}$. That is, we can consider a family of surfaces $\mathcal{S}(u, v, \phi)$ in which our surface is embedded.

Proof. Recall that $\int\int\int du dv dw = \int\int\int |J|^{-1} dx dy dz$ where the Jacobian (the volume scaling ratio) is given by $J = \langle (x_u, y_u, z_u) \times (x_v, y_v, z_v), (x_w, y_w, z_w) \rangle$. In our case, for the general parameterized surface $\mathcal{S}(u, v)$, and its implicit representation $\phi(x, y, z) = 0$, we have that the normal is given by $\vec{N} = \frac{\mathcal{S}_u \times \mathcal{S}_v}{|\mathcal{S}_u \times \mathcal{S}_v|} = \frac{\nabla \phi}{|\nabla \phi|}$. We thereby have that $\mathcal{S}_u \times \mathcal{S}_v = |\mathcal{S}_u \times \mathcal{S}_v| \vec{N} = |\mathcal{S}_u \times \mathcal{S}_v| \frac{\nabla \phi}{|\nabla \phi|}$.

We also know that $1 = \frac{d\phi}{d\phi} = \langle \nabla\phi, (x_\phi, y_\phi, z_\phi) \rangle$. Thus, $\mathcal{S}_\phi = (x_\phi, y_\phi, z_\phi) = \frac{\nabla\phi}{|\nabla\phi|^2}$, and the Jacobian is given by

$$J = \langle \mathcal{S}_u \times \mathcal{S}_v, \mathcal{S}_\phi \rangle = \left\langle \mathcal{S}_u \times \mathcal{S}_v, \frac{\nabla\phi}{|\nabla\phi|^2} \right\rangle = \left\langle |\mathcal{S}_u \times \mathcal{S}_v| \frac{\nabla\phi}{|\nabla\phi|}, \frac{\nabla\phi}{|\nabla\phi|^2} \right\rangle = \frac{|\mathcal{S}_u \times \mathcal{S}_v|}{|\nabla\phi|}.$$

We can now conclude with

$$\begin{aligned} \int_\phi \int_u \int_v g(\mathcal{S}(u,v,\phi)) |\mathcal{S}_u \times \mathcal{S}_v| du dv d\phi &= \int_x \int_y \int_z g(x,y,z) |J|^{-1} |\mathcal{S}_u \times \mathcal{S}_v| dx dy dz \\ &= \int_x \int_y \int_z g(x,y,z) |\nabla\phi| dx dy dz. \quad \blacksquare \end{aligned}$$

REFERENCES

1. L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via γ-convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990.
2. L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. It.*, 7:105–123, 1992.
3. A M Bruckstein, R Onn, and T J Richardson. Improving the vision of magic eyes: A guide to better autostereograms. In K Bowyer and N Ahuja, editors, *Advances in Image Understanding, A Festschrift for Azriel Rosenfeld*. IEEE Computer Science, 1996.
4. V Caselles, R Kimmel, and G Sapiro. Geodesic active contours. *IJCV*, 22(1):61–79, 1997.
5. O Faugeras and R Keriven. Variational principles, surface evolution PFE's, level set methods, and the stereo problem. *IEEE Trans. on Image Processing*, 7(3):336–344, 1998.
6. B K P Horn. Obtaining shape from shading information. In P H Winston, editor, *The Psychology of Computer Vision*, pages 115–155. McGraw Hill, New York, 1975.
7. B K P Horn and M J Brooks, editors. *Shape from Shading*. MIT Press, Cambridge MA, 1989.
8. R Kimmel and N Sochen. Geometric-variational approach for color image enhancement and segmentation. In *Lecture Notes In CS: Scale-Space Theories in Computer Vision*, volume 1682. Springer, 1999.
9. R Klette, K Schluns, and A Koscan. *Computer Vision three dimentional data from images*. Springer, Singapore, 1998.
10. T Richardson and S Mitter. Approximation, computation, and distoration in the variational formulation. In B M ter Haar Romeny, editor, *Geometric–Driven Diffusion in Computer Vision*. Kluwer Academic Publishers, The Netherlands, 1994.
11. Y Shirai. *Three-dimentional Computer Vision*. Springer, Berlin, 1987.
12. H W Thimbleby, S Inglis, and I H Witten. Displaying 3D images: Algorithms for single-image random-dot stereograms. *IEEE Computer*, 37(10):38–48, 1994.
13. N E Thing Enterprises. *Magic Eye: A New Way of Looking at the World*. Michael Joseph Ltd, Penguin Group, 1993.
14. C W Tyler. Sensory processing of binocular disparity. In *Vergence Eye Movemets: Basic and Clinical Aspects*, pages 199–295. Butterworth, Boston, 1983.
15. C W Tyler and J J Chang. Visual echoes: the perception of repetition in random patters. *Vision Research*, 17(12):109–116, 1977.
16. C W Tyler and M B Clarke. The autostereogram. In *SPIE*, volume 1256, pages 182–197, 1990.