Multiview normal field integration using level set methods

Ju Yong Chang
School of Electrical Eng.
Seoul National University
151-600, Seoul, Korea
jangbon@snu.ac.kr

Kyoung Mu Lee
School of Electrical Eng.
Seoul National University
151-600, Seoul, Korea
kyoungmu@snu.ac.kr

Sang Uk Lee
School of Electrical Eng.
Seoul National University
151-600, Seoul, Korea
sanguk@sting.snu.ac.kr

Abstract

In this paper, we propose a new method to integrate multiview normal fields using level sets. In contrast with conventional normal integration algorithms used in shape from shading and photometric stereo that reconstruct a 2.5D surface using a single-view normal field, our algorithm can combine multiview normal fields simultaneously and recover the full 3D shape of a target object. We formulate this multiview normal integration problem by an energy minimization framework and find an optimal solution in a least square sense using a variational technique. A level set method is applied to solve the resultant geometric PDE that minimizes the proposed error functional. It is shown that the resultant flow is composed of the well known mean curvature and flux maximizing flows. In particular, we apply the proposed algorithm to the problem of 3D shape modelling in a multiview photometric stereo setting. Experimental results for various synthetic data show the validity of our approach.

1. Introduction

Integrating a normal (gradient) field is the essential step in many computer vision problems including the shape from shading [11], photometric stereo [26], shape from texture [6], and so on. And most of algorithms construct 2.5D shapes from a single normal field. However, even with a single normal field, normal integration is not an easy task due to the integrability problem caused by the inherent photometric noise and numerical error. So far many works have been proposed for the robust integration of a normal field to noise, and these methods can be categorized into two groups: The first approach is to map the obtained noisy non-integrable normal field into an integrable space. Then, a simple path integration is applied to the transformed normal field so that a desired shape is reconstructed. Since Frankot and Chellappa [7] have proposed to project the possibly non-integrable gradient field onto the nearest integrable slopes using Fourier basis functions, many similar approaches have been proposed such as the methods using cosine basis functions [8], localized wavelet basis functions [13], and a redundant non-orthogonal set of basis functions (shapelets) [15]. Recently, Petrovic *et al*. [21] proposed to use the belief propagation algorithm in graphical networks to enforce integrability under the assumption that the noise is Gaussian. The second approach is to reconstruct the desired shape directly. In this approach, some error functional involving the depth and normal field is defined and minimized to produce a surface. Horn and Brooks [12] used a variational technique to derive the Euler equation of an error functional, and solved it by the Gauss-Seidel relaxation method. In contrast, Simchony *et al*. [22] proposed a direct analytical method by solving a Poisson equation.

However, note that all the algorithms mentioned above are devised to integrate a single normal field, and they can not be applied directly to the multiview normal integration problem. So, in this paper, we address a novel *multiview normal fields integration* problem and propose an efficient algorithm based on a variational energy minimization framework. In contrast with conventional single-view normal integration algorithms that reconstruct only 2.5D structures, the proposed algorithm can recover the full 3D shape of an object. To the best of our knowledge, there has not been reported any algorithm that integrates multiview normal fields in the computer vision literature before.

Our approach is as follows. We formulate the problem by using a variational framework with a new error functional obtained by generalizing the conventional error functional in [12] in the multiview setting. Then we derive a geometric PDE that minimizes the proposed functional, which is proven to be composed of both the well known mean curvature flow [4] and flux maximizing flow [24]. The obtained geometric PDE is solved by a level set method.

In this paper, we also show that the proposed multiview normal integration algorithm can be applied efficiently to the problem of reconstructing complete 3D shape in a *multi-*

view photometric stereo setting, in which multiview normal fields are estimated from a number of photometric stereo image sets captured under varying light conditions and camera viewpoints. Note that although few works have been reported on reconstructing shape using multiview photometric stereo images, their problem statements and approaches are different from ours. Lim *et al*. [17] proposed an iterative algorithm for shape reconstruction from multiple images of a moving Lambertian object. The algorithm starts with an initial piecewise linear surface. And, in the first step of each iteration, the surface normals are estimated through matrix factorizations of intensities gathered from the previous surface estimate and images. In the second step, they are integrated through the minimization using gradient descent, and a new surface is obtained. However, since only a single-view normal field can be handled by their normal integration algorithm, their result is not a complete 3D model but a 2.5D depth map. Instead of estimating and integrating the normal fields from photometric stereo images, Vogiatzis *et al*. [25] proposed to use a mesh model for the reconstruction of a full 3D shape of a Lambertian and textureless object. Their algorithm initializes the object as a visual hull, and uses the bootstrapping technique that is similar to the work of [17]. Although the full 3D reconstruction is possible due to the mesh based volumetric model, there still remain limitations of assuming the Lambertian and textureless object.

We have conducted experiments on the performance of the proposed algorithm on various test sets, and demonstrated that our algorithm could reconstruct the shapes of arbitrary closed objects accurately with the multiple normal fields obtained from different camera views. Moreover, an additional experiment showed that the proposed algorithm could be applied to the multiview photometric stereo problem successfully.

The rest of this paper is organized as follows. In the next section, we address a brief description and notations of the multiview normal field integration problem. In Section 3, we describe the proposed algorithm in detail. And, experimental results are shown in Section 4. Finally, conclusions are drawn in Section 5.

2. Problem statement

Let S be a closed, twice differentiable surface in $\mathbb{R}^3$, P be a generic point on S, and $\mathbf{N}$ be an unit inward normal at P. And, we assume that there are n calibrated cameras. So, we know the extrinsic and intrinsic parameters of each camera. The extrinsic parameters can be represented by a rigid body transformation $g_i \in SE(3)$, $i = 1, 2, ..., n$ which is composed of a rotation matrix $R_i \in SO(3)$ and a translation vector $T_i \in \mathbb{R}^3$. Then, if we consider $\mathbf{X}$ to be the coordinates of P with respect to the inertial reference frame, the coordinates of P in each camera coordinate frame can be written as $g_i(\mathbf{X}) = R_i\mathbf{X} + T_i$. For the intrinsic parameters, we assume the canonical perspective projection $\pi : \mathbb{R}^3 \longrightarrow \mathbb{P}^2$; $\mathbf{X} = [X_1, X_2, X_3]^T \longmapsto \mathbf{X}/X_3$. Extension to the case of general intrinsic parameters can be easily derived. Therefore, the projection of a point P onto the i-th camera has the coordinates of $\mathbf{x}_i = \pi_i(\mathbf{X}) \doteq \pi(g_i(\mathbf{X}))$.

Let the normal field $\mathbf{v}_i : \Omega \subset \mathbb{R}^2 \longmapsto \mathbb{R}^3$, be formed by projecting the normal vectors of the surface S onto the image plane of the i-th camera under certain geometric conditions. By considering some additive noise in the projection process, and the visibility constraint of surface points for the camera, a more realistic normal field formation model can be obtained as follows:

$$\begin{cases} \mathbf{v}_i(\mathbf{x}_i) = \mathbf{N}(\mathbf{X}) + \mathbf{n}_i, \\ \forall \mathbf{X} \in \{\mathbf{X} | \psi_{i,S}(\mathbf{X}) = 1\}, \end{cases} \qquad i = 1, 2, ..., n, \qquad (1)$$

where $\mathbf{n}_i$ is a noise vector that quantifies the deviation from the true normal vector, and

$$\psi_{i,S}(\mathbf{X}) \doteq \begin{cases} 1, & \text{if } P \text{ is visible from the } i\text{-th camera;} \\ 0, & \text{otherwise} \end{cases}$$

is the visibility function that indicates whether the point P on the current surface S is visible or not from the i-th camera.

Now, the problem of multiview normal field integration can be stated as inferring the coordinates $\mathbf{X}$'s of all surface points (or simply S) with given estimated or observed normal fields $\mathbf{v}_i$, $i = 1, 2, ..., n$ using (1).

3. Proposed algorithm

3.1. Proposed error functional

Energy minimization has been one of the main approaches to solve the traditional single-view normal integration problem. Let $Z : \Omega \subset \mathbb{R}^2 \longmapsto \mathbb{R}$ denotes a target surface, and $[p, q]^T : \Omega \subset \mathbb{R}^2 \longmapsto \mathbb{R}^2$ to be a given noisy gradient field. Then, the single-view normal integration problem is usually formulated as finding the surface Z that minimizes the following least square error functional [11, 22] given by

$$J_1(Z) \doteq \int\int ((Z_x - p)^2 + (Z_y - q)^2) dxdy, \qquad (2)$$

where $[Z_x, Z_y]^T$ is the gradient field of Z. This error functional integrates the discrepancy between the observed gradient field and the surface gradient field on the image domain.

In this paper, we also employ an energy minimization approach to solve the multiview normal integration problem in (1) by generalizing the error functional in (2). Note however that in our multiview normal field integration problem, since there is no particular viewpoint that can be used

as a domain of integration, the target object cannot be represented as a graph on a planar domain as in the single-view case. So, in our new error functional, the discrepancy between the observed normal fields and the desired surface normal should be integrated on the surface domain. Moreover, we have to consider the visibility of each surface point P for each camera, and normalize the summation of the discrepancy measure by the number of cameras from which the point is visible. Thus, we propose the following error functional for the integration of multiview normal fields:

$$J_2(S) \doteq \int_S \frac{1}{N_S(\mathbf{X})} \sum_{i=1}^{n} \psi_{i,S}(\mathbf{X}) \|\mathbf{N}(\mathbf{X}) - \mathbf{v}_i(\pi_i(\mathbf{X}))\|^2 dA, \tag{3}$$

where dA is an area form and

$$N_S(\mathbf{X}) \doteq \sum_{i=1}^{n} \psi_{i,S}(\mathbf{X}) \tag{4}$$

denotes the number of cameras from which the point P is visible.

By using (4) and the fact that $\mathbf{N}(\mathbf{X})$ and $\mathbf{v}_i(\mathbf{X})$ are unit normal vectors ($\|\mathbf{N}(\mathbf{X})\| = \|\mathbf{v}_i(\mathbf{X})\| = 1$), the proposed functional in (3) can be rewritten as follows (for the simplicity of the notation, N_S, $\psi_{i,S}$, $\mathbf{N}$, $\mathbf{v}_i$, $\mathbf{V}$ are assumed to be functions of $\mathbf{X}$):

$$\begin{aligned}
J_2(S) &= \int_S \frac{1}{N_S} \sum_i \psi_{i,S} [\|\mathbf{N}\|^2 + \|\mathbf{v}_i\|^2 - 2\mathbf{N} \cdot \mathbf{v}_i] dA \\
&= 2 \int_S \frac{1}{N_S} \sum_i \psi_{i,S} [1 - \mathbf{N} \cdot \mathbf{v}_i] dA \\
&= 2 \int_S [\frac{\sum_i \psi_{i,S}}{N_S} - \frac{\sum_i \psi_{i,S} \mathbf{N} \cdot \mathbf{v}_i}{N_S}] dA \\
&= 2 \int_S [1 - \mathbf{N} \cdot \frac{\sum_i \psi_{i,S} \mathbf{v}_i}{N_S}] dA \\
&= 2 \int_S dA - 2 \int_S \mathbf{N} \cdot \mathbf{V} dA,
\end{aligned} \tag{5}$$

where

$$\mathbf{V}(\mathbf{X}) \doteq \frac{\sum_i \psi_{i,S}(\mathbf{X}) \mathbf{v}_i(\pi_i(\mathbf{X}))}{N_S(\mathbf{X})} \tag{6}$$

is a vector field denoting the average of the projected visible normal vectors. If we ignore the proportional constant, the first term represents the Euclidean area of the surface S, and the second term is equal to the flux of a given vector field $\mathbf{V}$ through the surface. The optimal solution of this functional can be obtained by simultaneously performing the minimization of the area term and the maximization of the flux term.

Euclidean or more general Riemannian area (length) have been applied to many fields in computer vision such as object segmentation [3, 4] and multiview stereo [5]. By minimizing these area (length) functionals, minimal surfaces (curves) are obtained. Note that in the proposed functional, regularization is implicitly enforced by the minimization of the Euclidean area term. On the other hand, flux has also been considered in many image segmentation problems [24]. Flux of a vector field provides a measure of how well the surface (curve) is aligned with the direction perpendicular to the vector field. Thus, the flux term in the proposed functional encourages the surface normal to align with the vectors in the field $\mathbf{V}$.

3.2. Optimization by variational method

The multiview normal integration problem can be solved by minimizing the proposed functional in (5) using various optimization techniques including variational method and graph cuts. The variational method has been used as a typical approach for energy minimization in computer vision. In the variational method, the local minimum of the energy functional can be obtained by solving the corresponding Euler-Lagrange equation. And the Euler-Lagrange equation gives the geometric flow which describes the evolution of the surface. Recently, graph cuts also has attracted much attention as an energy minimization method due to its good performance. And it has been applied to optimize the geometric functional composed of area (length) and flux so that the global minimum was obtained [2, 14]. In this paper, we adopt the variational method as an optimization tool, and leave the using of graph cuts as a future work.

Let S_t be a smooth family of closed surfaces with time index t, and S_0 be an initial surface. This family of surfaces describes a surface evolution. Then, the local minimum of the proposed functional can be obtained by evolving the surface through a particular geometric PDE. The following theorem gives us such an energy minimizing flow. (The proof is omitted due to lack of space. It can be easily proved by using theorems in [4] and [24].)

Theorem 1. *The gradient flow that minimizes the functional (5) is given by the following equation with* $S(t = 0) = S_0$

$$S_t = 2H\mathbf{N} - div(\mathbf{V})\mathbf{N}, \tag{7}$$

where H denotes the mean curvature of S.

In other words, the gradient flow that minimizes (5) is obtained by moving each point of the surface in the direction of the inward normal by an amount proportional to the mean curvature minus the divergence of the vector field $\mathbf{V}$. This can be easily deduced from the formation of the functional (5) which is composed of the area and flux terms. It is well known that the mean curvature flow ($S_t = 2H\mathbf{N}$) [4] minimizes the Euclidean area and the flux maximizing flow ($S_t = div(\mathbf{V})\mathbf{N}$) [24] maximizes the flux term.

By using (6), the divergence in (7) can be rewritten as

$$\begin{aligned} div(\mathbf{V}) &= div(\frac{\sum_i \psi_{i,S}\mathbf{v}_i}{N_S}) \\ &= \sum_i div(\frac{\psi_{i,S}}{N_S})\mathbf{v}_i + \sum_i \frac{\psi_{i,S}}{N_S} div(\mathbf{v}_i). \end{aligned}$$

Since $\frac{\psi_{i,S}}{N_S}$ is constant (0 or 1) at most regions on the surface S except some points where visibility from the i-th camera changes, $div(\frac{\psi_{i,S}}{N_S})$ can be approximated to be 0, and the above equation becomes

$$div(\mathbf{V}) \approx \sum_i \frac{\psi_{i,S}}{N_S} div(\mathbf{v}_i) = 2\sum_i \frac{\psi_{i,S}}{N_S} H_i,$$

where H_i denotes the observed mean curvature from the i-th camera. Now, the gradient flow in (7) can be approximated by

$$S_t \approx 2H\mathbf{N} - 2\overline{H}\mathbf{N}, \tag{8}$$

where $\overline{H} \doteq \sum_i \frac{\psi_{i,S}}{N_S} H_i$ denotes the average of the observed mean curvature values. From (8), we can see that the flow converges when the mean curvature of the current surface is equal to the average of the observed mean curvatures. This means that the geometry (mean curvature) of the surface is determined by multiple cues (mean curvature) of multiview inputs. So our algorithm can reconstruct the 3D shape more robustly than single-view normal integration algorithms that use only one cue.

3.3. Implementation using level set

In order to implement the gradient flow in (7), we use the level set method [19, 20]. In the level set method, a surface is implicitly represented by a zero level set of a Lipschitz continuous function $\phi : \mathbb{R}^3 \longrightarrow \mathbb{R}$. This representation enables embedded surfaces to have different topologies so that it allows topological changes during surface evolution. The gradient flow in (7) gives the following level set equation for the function ϕ:

$$\phi_t = 2H\|\nabla\phi\| - div(\mathbf{V})\|\nabla\phi\|. \tag{9}$$

Specifically, the above level set equation is composed of a parabolic term ($2H\|\nabla\phi\|$) and a hyperbolic term ($div(\mathbf{V})\|\nabla\phi\|$). The parabolic term can be discretized using the central difference method, while the hyperbolic term can be independently discretized by upwinding in the proper direction.

The surface evolution through the gradient flow in (7) requires an initial surface S_0 and the computation of the vector field $\mathbf{V}$. In this paper, we initialize the surface evolution by using a visual hull [16]. We assume that object's silhouette information from each viewpoint is also given in addition to the normal fields. Then, by using such silhouettes, we can obtain the visual hull of the object, that is defined as the maximal object that produces same silhouettes. In our level set based implementation of (9), ϕ_0 is defined as the signed distance function of which the zero level set is equal to S_0. For the computation of the vector $\mathbf{V}$ of each point, we need visibility of that point from each camera, that depends on the surface S. Thus, at each iteration, the visibility of every point for each camera is determined first using the camera geometry and the current surface S. In order to compute the visibility efficiently, we employ the implicit ray tracing algorithm by Tsai *et al*. [23] that exploits the level set representation.

3.4. Application to multiview photometric stereo

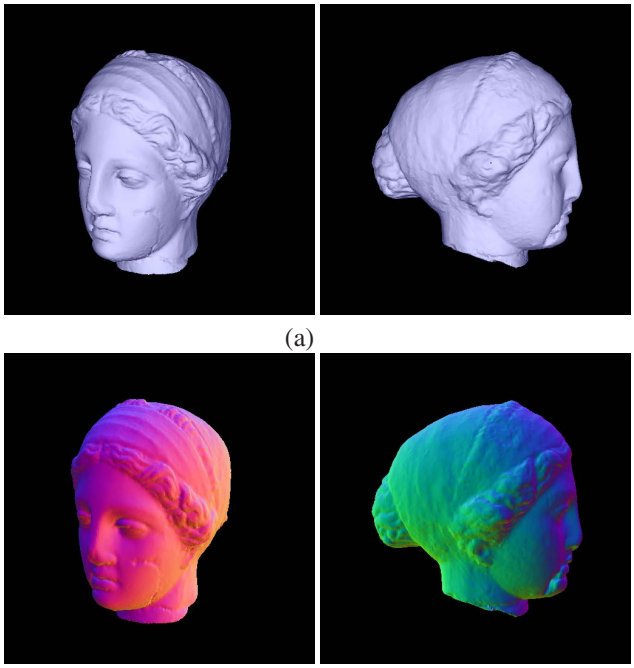
In this section, we apply our multiview normal integration algorithm to the multiview photometric stereo problem. In conventional single-(camera) view photometric stereo, multiple images at a fixed viewpoint are taken under different light directions. By using such images, it generally produces a gradient (normal) field, that is subsequently integrated for the reconstruction of a 2.5D surface. In contrast, multiview photometric stereo can be defined as inferring the shape from images under varying both lighting conditions and viewpoints, rather than a fixed viewpoint.

Our algorithm can be stated as follows. Let I_{ij}, $i = 1, 2, ..., n$, $j = 1, 2, ..., l$ denote input images taken at n camera viewpoints with l unknown light direction vectors. In this situation, it is known that the shape of a Lambertian object with a variable albedo can be reconstructed up to an unknown Generalized Bas-Relief (GBR) ambiguity [1]. In this paper, we additionally assume that the light direction vectors $\mathbf{L}_j$ are known, so that the GBR ambiguity can be resolved. By using a conventional single-view photometric stereo technique, each normal field corresponding to each viewpoint can be computed independently through a matrix factorization using Singular Value Decomposition. Note that each normal field $\mathbf{v}_i$ is the optimal solution of the following least square energy function:

$$\sum_{j=1}^{l} \sum_{(x,y)} (I_{ij}(x,y) - a_i(x,y)\mathbf{L}_j \cdot \mathbf{v}_i(x,y))^2,$$

where $a_i(x,y)$, $i = 1, 2, ..., n$ denotes the surface albedo. Details on computing $\mathbf{v}_i$ can be found in [26].

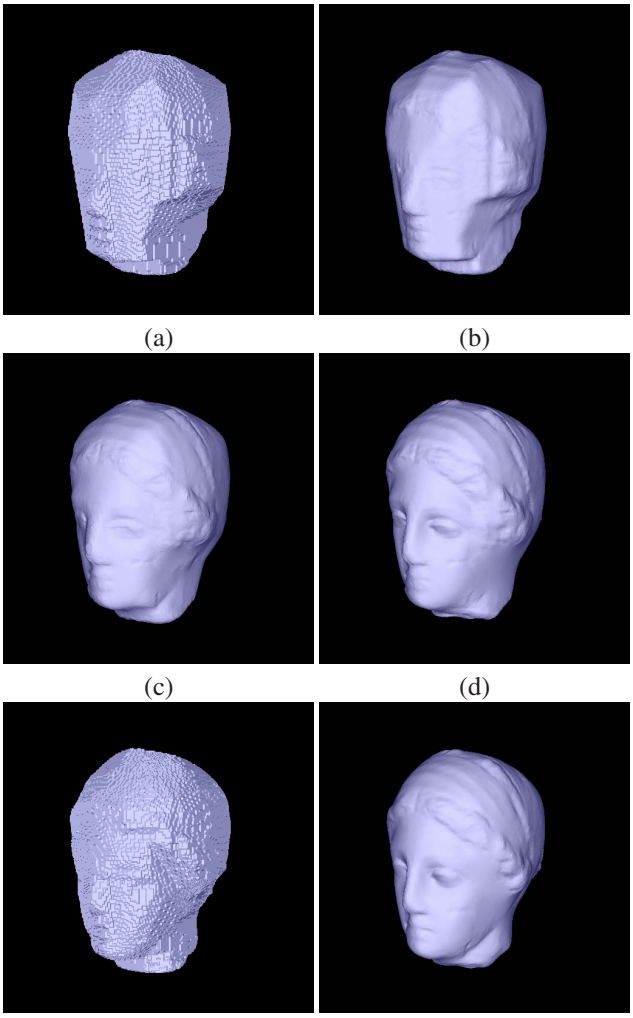
Now, we can apply our multiview normal integration algorithm to the obtained multiple normal fields $\mathbf{v}_i$, $i = 1, 2, ..., n$, so that a 3D shape of the object can be reconstructed. Note that our approach consists of two steps, that is, the estimation of each normal field by conventional single-view photometric stereo, and the multiview integration of such normal fields. Thus, although the Lambertian model is assumed in our current implementation, it can be relaxed by employing more general photometric stereo



(a)

(b)

Figure 1. The igea model is rendered from several views in (a). In order to get the normal fields, the normal vectors are transformed to the RGB color vectors, which are used for rendering the model. Then the captured normal fields are displayed in (b).



(a) (b)

(c) (d)

(e) (f)

Figure 2. From 6 input normal fields, the visual hull is computed and displayed in (a). Our algorithm is initialized by the visual hull, and produces the surface evolution shown in (b)-(d). The visual hull and the converged surface generated from 18 normal fields is displayed in (e) and (f), respectively.

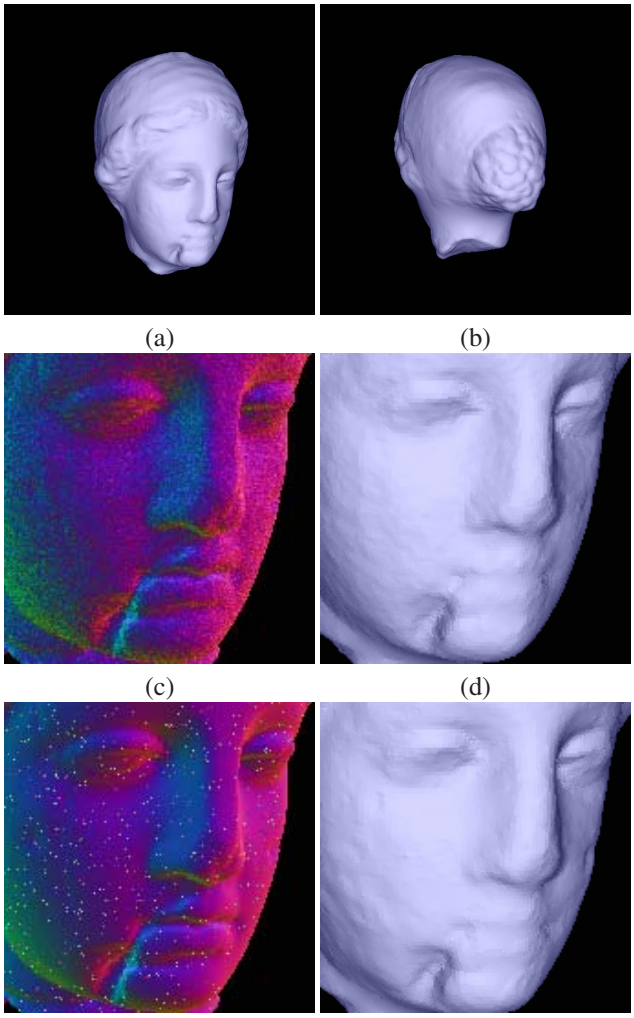
algorithms for the first step of our algorithm. Then normal fields of an object with almost arbitrary reflectance model can be obtained, and its 3D shape model can be also reconstructed. For example, by adopting recent algorithms [9, 10], we can reconstruct the full 3D shape of an object with a completely arbitrary reflectance model. And, in the recent work of Lim *et al*. [17], not only images of varying illumination at a viewpoint but also those from different viewpoints are used to estimate the normal field. But, as mentioned in their work, because there was not a multiview normal integration algorithm, multiview normal information was not fully utilized and only 2.5D reconstruction was possible. In this case also, our multiview normal integration algorithm can make it possible to recover the full 3D shape.

4. Experimental results

In this section, we present an experimental results of our algorithm. Firstly, we test our multiview normal field integration algorithm, and then application results on multiview photometric stereo are presented.

4.1. Multiview normal field integration

To test our multiview normal integration algorithm, we used the *igea* and *rabbit* models obtained from the Cyberware homepage (http://www.cyberware.com/). They are mesh based synthetic models. The igea model is composed of 134345 vertices and 268686 triangles, and the rabbit model has 67038 vertices and 134074 triangles. The normal vector of each triangle was computed and transformed to the RGB color vector as $\mathbf{C} = [0.5, 0.5, 0.5]^T + 0.5 \cdot \mathbf{N}$, where $\mathbf{N}$ is a normal vector and $\mathbf{C}$ is a RGB color vector of which components are ranged between 0.0 and 1.0. Then a normal field from one view was constructed by rendering the mesh model with the normal based color vectors. We used an openGL rendering function. 512×512 normal fields from several views were captured synthetically, and some noise was added. Then the normal fields and the corresponding camera parameters were used as inputs to our

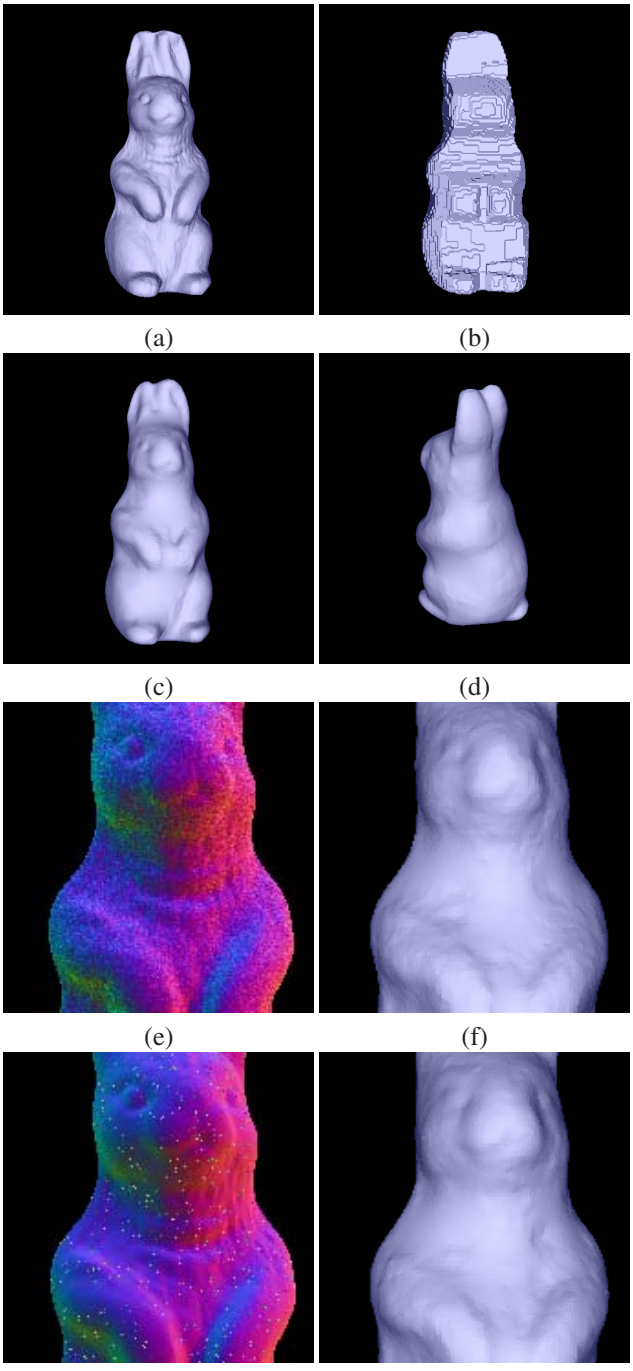


(a) (b) (c) (d) (e) (f)

Figure 3. The reconstructed shape of the igea model is rendered from two viewpoints in (a), (b). Serious gaussian and random noises are added to the input normal fields, and the resultant noisy normal fields are displayed in (c), (e). Reconstruction results from those noisy inputs are shown in (d) and (f).

algorithm. The mesh model and the captured noiseless normal fields of the igea model are illustrated in Figure 1.

Our algorithm was implemented in C++ and executed on a PC with Pentium IV 3.4GHz CPU and 2.0GB RAM. A $200 \times 200 \times 200$ grid was used to implement the level set based iteration, and the narrow band implementation with 16 voxels in width was adopted for efficiency. Note that except these grid size and band width, there is no other parameter in our algorithm. Typically, our algorithm was converged within 1 hour, and the final converged 3D array was obtained. By using the marching cubes algorithm [18], the zero level set of the converged 3D array was converted into the mesh model, which was subsequently rendered for the evaluation of the performance qualitatively.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 4. The rabbit model is rendered in (a). From 6 normal fields, the visual hull is computed and displayed in (b). The reconstructed shape by our algorithm is rendered from two different viewpoints in (c), (d). Noisy normal fields corrupted by gaussian and random noises are displayed in (e), (g). Then reconstruction results from those noisy inputs are shown in (f) and (h).

We illustrate the surface evolution result of our algorithm for the igea model in Figure 2(a)-(d). The surface was ini-

tialized from the visual hull generated by 6 input normal fields, and evolved by using the level set method. We can see that our surface evolution algorithm converged well to the original mesh model. To investigate the robustness of our algorithm to the number of input normal fields, another experiment using 18 inputs was performed. The initial visual hull and the converged surface are shown in Figure 2(e) and (f). Although the number of inputs was considerably smaller in the former experiment, the converged surfaces of the two experiments look similar. Thus we can conclude that our algorithm is robust to the number of views and only a few views are sufficient to recover the full shape correctly.

To evaluate the performance of our algorithm qualitatively, some more experiments were performed for the igea and rabbit models. The results are illustrated in Figure 3 and 4. Firstly, the 3D shapes of both models were reconstructed by our algorithm, and rendered from several views as shown in Figure 3(a)-(b) and Figure 4(c)-(d), respectively. We can see that every detail of the original shape has been well recovered. And, obviously our algorithm produced the full 3D shape reconstruction as contrasted with conventional normal integration algorithms. Secondly, in order to investigate the robustness of the proposed algorithm to noise, we added gaussian and random noises to the input normal fields. Gaussian noise with mean 0 and standard deviation 0.2 was added to each component of the normal vector, that was subsequently normalized as a new noisy normal vector. Figure 3(c) and Figure 4(e) show the Gaussian noise added normal fields of the igea and rabbit models, respectively. For the random noise test, we generated noisy normal fields by changing the original normal vector to a random unit vector with 0.03 probability as shown in Figure 3(e) and Figure 4(g). The final shape reconstruction results are presented in Figure 3(d) and (f), and Figure 4(f) and (h) for each model and noise type, respectively. We observe that the shapes have been reconstructed accurately and robustly enough despite of such noisy inputs. In fact, the robustness of our algorithm to noise can be readily expected by the least square formulation and the area minimizing regularization of the proposed functional. Moreover, as indicated in (8), using multiple cues from multiview inputs improves the robustness to noise. Finally, note that the normal fields in Figure 3(c) and (e) have serious discontinuities in the regions of nose and chin. It is still difficult for many conventional normal integration algorithms to reconstruct the shapes in such challenging regions. However, by virtue of using multiview inputs, our algorithm could reconstruct them quite successfully.

4.2. Multiview photometric stereo

For the test of our multiview photometric stereo algorithm, the shape and texture information of a real clay doll was acquired by using a 3D laser scanner, so that a 3D mesh model was constructed. Then, by assuming the Lambertian reflectance model and using the obtained texture as the surface albedo, multiview photometric stereo images were synthesized by rendering the mesh model under varying light directions and camera viewpoints. The shape of the acquired dancer model and the synthesized images from a fixed viewpoint are illustrated in Figure 5(a), and Figure 5(b)-(d), respectively. For qualitative performance evaluation, synthesized images under 13 camera viewpoints and 13 light directions were used as inputs of our algorithm. Reconstruction results are shown in Figure 5(e) and (f). We can see that the 3D shape of the original model has been well recovered, and conclude that our algorithm is effective for shape reconstruction in the multiview photometric

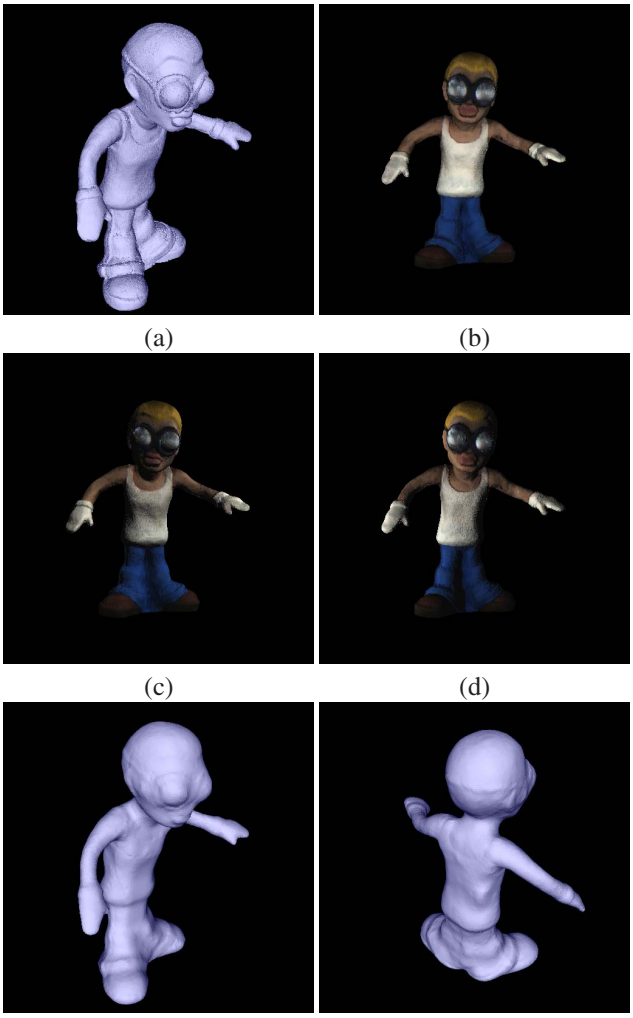


Figure 5. The ground truth shape of the dancer model is rendered in (a). Three images captured under different light directions, a fixed viewpoint, and the Lambertian reflectance assumption, are shown in (b)-(d). The reconstructed shape by our multiview photometric stereo algorithm is displayed in (e) and (f).

stereo setting.

5. Conclusion

In this paper, we have presented a novel multiview normal fields integration algorithm. It can combine normal fields from multiple viewpoints and generate a full 3D shape of a target object. To the best of our knowledge, this is the first algorithm that can integrate multiview normal fields. Our algorithm has been applied to reconstruct a shape of a Lambertian object in a multiview photometric stereo setting, so that a new multiview photometric stereo algorithm has been presented. It is efficient and flexible, so that can be easily extended to cover objects with a more general reflectance model.

In the optimization step of our algorithm, a variational technique has been adopted for minimization of the proposed functional, and the resultant geometric PDE has been solved via a level set method. However this approach has a limitation that it can find only some local minimum, and the initial surface has to be close to the solution. So, one of our future works will be applying other optimization techniques including the graph cuts to our algorithm. In our current implementation of multiview photometric stereo algorithm, we assumed Lambertian reflectance model. Thus another direction of our future work will be employing more general photometric stereo algorithms as the normal fields estimation stage of our algorithm, so that we can reconstruct the shape of objects with arbitrary reflectance models.

References

[1] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. *Int'l J. Computer Vision*, 35(1):33–44, Nov. 1999.

[2] Y. Boykov and V. Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *Proc. Int'l Conf. Computer Vision*, volume I, pages 26–33, 2003.

[3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Int'l J. Computer Vision*, 22(1):61–79, Feb. 1997.

[4] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. Minimal surfaces: A three dimensional segmentation approach. *Numerische Methematik*, 77(4):423–451, 1997.

[5] O. Faugeras and R. Keriven. Variational principles, surface evolution, pde's, level set methods, and the stereo problem. *IEEE Trans. Image Processing*, 7(3):336–344, March 1998.

[6] D. Forsyth. Shape from texture and integrability. In *Proc. Int'l Conf. Computer Vision*, pages 447–452, 2001.

[7] R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading. *IEEE Trans. Pattern Anal. Machine Intell.*, 10(4):439–451, July 1988.

[8] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(6):643–660, June 2001.

[9] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proc. Int'l Conf. Computer Vision*, pages 341–348, 2005.

[10] A. Hertzmann and S. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Trans. Pattern Anal. Machine Intell.*, 27(8):1254–1264, August 2005.

[11] B. Horn. Height and gradient from shading. *Int'l J. Computer Vision*, 5(1):37–75, August 1990.

[12] B. K. P. Horn and M. J. Brooks. The variational approach to shape from shading. *Comput. Vision, Graphics, Image Processing*, 33:174–208, Feb. 1986.

[13] J. Hsieh, H. Liao, M. Ko, and K. Fan. Wavelet-based shape from shading. *Graphical Models and Image Processing*, 57(4):343–362, July 1995.

[14] V. Kolmogorov and Y. Boykov. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In *Proc. Int'l Conf. Computer Vision*, volume I, pages 564–571, 2005.

[15] P. Kovesi. Shapelets correlated with surface normals produce surfaces. In *Proc. Int'l Conf. Computer Vision*, pages 994–1001, 2005.

[16] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(2):150–162, Feb. 1994.

[17] J. Lim, J. Ho, M. Yang, and D. Kriegman. Passive photometric stereo from motion. In *Proc. Int'l Conf. Computer Vision*, pages 1635–1642, 2005.

[18] W. Lorensen and H. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *Computer Graphics*, 21(4):163–169, 1987.

[19] S. Osher and R. Fedkiw. Level set methods and dynamic implicit surfaces, 1st edition, 2002.

[20] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *J. Computational Physics*, 79:12–49, 1988.

[21] N. Petrovic, I. Cohen, B. Frey, R. Koetter, and T. Huang. Enforcing integrability for surface reconstruction algorithms using belief propagation in graphical models. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 743–748, 2001.

[22] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(5):435–446, May 1990.

[23] Y. Tsai, L. Cheng, S. Osher, P. Burchard, and G. Sapiro. Visibility and its dynamics in a pde based implicit framework. *J. Computational Physics*, 199:260–290, 2004.

[24] A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *IEEE Trans. Pattern Anal. Machine Intell.*, 24(12):1565–1578, Dec. 2002.

[25] G. Vogiatzis, C. Hernandez, and R. Cipolla. Reconstruction in the round using photometric normals and silhouettes. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 1847–1854, 2006.

[26] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 158–164, 1997.