HAL open science

Minimizing the Multi-view Stereo Reprojection Error for Triangular Surface Meshes

Amael Delaunoy, Emmanuel Prados, Pau Gargallo I Piracés, Jean-Philippe Pons, Peter Sturm

To cite this version:

Amael Delaunoy, Emmanuel Prados, Pau Gargallo I Piracés, Jean-Philippe Pons, Peter Sturm. Minimizing the Multi-view Stereo Reprojection Error for Triangular Surface Meshes. BMVC 2008 - British Machine Vision Conference, Sep 2008, Leeds, United Kingdom. pp.1-10, 10.5244/C.22.58 . inria-00308944

HAL Id: inria-00308944

https://inria.hal.science/inria-00308944

Submitted on 4 Aug 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Minimizing the Multi-view Stereo Reprojection Error for Triangular Surface Meshes

A. Delaunoy[1*], E. Prados[1*], P. Gargallo[1], J.-P. Pons[2*] and P. Sturm[1]
[1] INRIA Rhône-Alpes / PERCEPTION / LJK, Grenoble, France
[2] CERTIS / École des Ponts, Paris, France

Abstract

This article proposes a variational multi-view stereo vision method based on meshes for recovering 3D scenes (shape and radiance) from images. Our method is based on generative models and minimizes the reprojection error (difference between the observed images and the images synthesized from the reconstruction). Our contributions are twofold. 1) For the first time, we rigorously compute the gradient of the reprojection error for non smooth surfaces defined by discrete triangular meshes. The gradient correctly takes into account the visibility changes that occur when a surface moves; this forces the contours generated by the reconstructed surface to perfectly match with the apparent contours in the input images. 2) We propose an original modification of the Lambertian model to take into account deviations from the constant brightness assumption without explicitly modelling the reflectance properties of the scene or other photometric phenomena involved by the camera model. Our method is thus able to recover the shape and the diffuse radiance of non Lambertian scenes.

1 Introduction

Multi-view stereo is the problem of recovering the shape of scenes using cameras. Given a set of images of a scene taken from different camera positions, the goal is to reconstruct the shape $\mathcal{S}$, and optionally the appearance, of the object. Since it is the inverse problem of image rendering, this problem can be modelled in a Bayesian framework by minimizing the difference between the images of the reconstructed model and the observed ones, i.e. the reprojection error. The correct variational interpretation of the Bayesian analysis yields to the minimization of some energy functional defined on the images [6, 10, 12, 16, 20] and requires the visibility of the surface points to be carefully accounted for. Handling visibility is one of the major difficulties in the stereo vision problem. Previous work copes with this difficulty more or less elegantly. Most often the authors approximate the visibility in some pre-processing steps which can be completely prior to the whole algorithm or else inside the iterations of the minimization process [1, 5, 8, 9, 12, 16, 21].

Only recently, some authors [6, 20] manage to rigorously and fully account for visibility in the optimization process. In this recent work, Yezzi and Soatto [20] (for convex surfaces) and Gargallo et al. [6] (in the general case) provide the exact gradient of the reprojection error. This computation is done in the mathematical framework of continuous (smooth) surfaces as used in [17]. This work has shown that the correct computation of the reprojection error involves a new term in the gradient which forces the apparent contours

*Research was supported by the Flamenco project ANR-06-MDCA-007.

generated by the reconstructed surface to perfectly match with the apparent contours in the observed images. This makes the use of additional energy terms like ballooning [18], visual hull [7, 14] or contours unnecessary and completely removes the minimal surface bias present in many other problem formulations.

Our work can be then seen as an extension of the previous work [6, 20]. One of the limitations of the gradient computed in the mathematical framework of the continuous surfaces is that, in practice we mainly deal with discrete representations of the surface. So whichever discrete representation one uses, it may be a discrete level set representation or a mesh, one has to discretize the computed continuous gradient (i.e. [1, 3, 22] for meshes). Also, this unavoidably introduces errors and gives rise to significant numerical difficulties. Computing the correct gradient flows for triangulated meshes is the work of [4, 15]. Nevertheless, previous authors did not use energy functionals depending on visibility (3D image segmentation [15] and 3D shape matching [4]). In this paper, we give the exact gradient of the reprojection error for the case of discrete surfaces represented by triangulated meshes.

Another limitation of the algorithms implemented in [6, 20] is due to the fact that the authors assume that the scene is perfectly Lambertian. In real life, such surfaces are rare and, therefore, multi-view stereo algorithms have to be robust to non-Lambertian reflections. To improve the robustness, a common strategy is to modify input images in order to remove specular highlights [23]. However, these methods are strongly limited by the specific lighting configuration. Some authors do not consider the image pixels that potentially have specular reflection components: they treat these points as outliers [1, 8]. To compensate for the loss of information they then need a lot of images. Other authors try to improve the robustness to non-Lambertian effects by directly incorporating a specular reflectance model in the mathematical formulation of the problem [9, 21], nevertheless until now, these formulations lead to very complex and consuming optimization processes which return rather inaccurate results. Also, this kind of formulation is necessarily limited to specific materials (having the same reflectance as the one used in the modelling). Another widespread idea is to use robust similarity measures [10, 12, 19]. However, these are rather complex to implement and cannot deal with strongly erroneous information as in the case of camera occlusions. So the second contribution of this paper is a modification of the Lambertian model in order to take into account the deviations from the constant brightness assumption based on a smooth image-based component. We do not explicitly model the reflectance properties of the scene (contrary to [9, 21]) and other complex photometric phenomena involved by the cameras. Thus the method is fully generic, simple, easy to implement and very efficient. It recovers quite robustly the shape and the diffuse radiance of non Lambertian scenes viewed from flawed cameras.

2 Problem Formulation

Inferring the 3D shape and photometric information (such as radiance) of a scene from images of that scene is the inverse problem of image synthesis. Also if one is able to generate images from a candidate model, then, the comparison of these generated images and the observed images (i.e. the data) provides a natural reprojection error measure. In fact, this approach corresponds exactly to the Bayesian formulation of the problem which consists in maximizing the likelihood (i.e. here the reprojection error) and the prior probability (priors on the model) instead of maximizing the posterior probability directly. The formulation in terms of energies corresponds to the minimization of the sum of the data term (the reprojection error) and a prior term.

2.1 Modelling of the Reprojection Error

In order to be able to compare the whole observed images (data) with the images generated by the model, it is crucial to define and model the background. Also, as shown by [6, 20], this allows us to be sure that the estimated foreground surface does not shrink to an empty set (which is the global optimum for most cost functionals used in other work) and this allows that contours generated by the recovered object perfectly match the apparent contours of the input images. Whereas most of the previous work assumes that the background is known (e.g. simply modelling it as uniformly black, or by exploiting given silhouette images), here we also estimate the background images $B_i : \mathcal{I}_i \rightarrow \mathbb{R}^3$, under the single assumption that these images are smooth, similarly as [20]. Here i corresponds to the index of the camera and $\mathcal{I}_i$ is its image domain.

Now, let us assume that the scene surface $\mathcal{S}$ is Lambertian and the illumination static. Let $C : \mathcal{S} \rightarrow \mathbb{R}^3$ be the radiance function that associates colors to the points on the surface. Ideally, the color $I_i(\mathbf{u})$ observed at pixel $\mathbf{u}$ of image I_i should be equal to the color $C(\pi_{i,\mathcal{S}}^{-1}(\mathbf{u}))$ of its backprojection $\pi_{i,\mathcal{S}}^{-1}(\mathbf{u})$ onto the surface or, in the case where $\mathbf{u} \notin \pi_i(S)$, to the color $B_i(\mathbf{u})$ of the same pixel on the background images (π_i denoting the projection associated with camera i). Thus, the reprojection error of the surface is

$$E_{data} = \frac{1}{2} \sum_i \left[\int_{\pi_i(\mathcal{S})} \left(I_i(\mathbf{u}) - C(\pi_{i,\mathcal{S}}^{-1}(\mathbf{u})) \right)^2 d\mathbf{u} \; + \int_{\mathcal{I}_i - \pi_i(\mathcal{S})} \left(I_i(\mathbf{u}) - B_i(\mathbf{u}) \right)^2 d\mathbf{u} \right] . \tag{1}$$

Finally, in order to well pose the problem, we use as a prior on $\mathcal{S}$ an additional smoothing area energy $E_{R\mathcal{S}} = \int_{\mathcal{S}} d\sigma$. We also assume that the background images are smooth by adding the term $E_{RB} = \sum_i \int_{\mathcal{I}_i} |\nabla B_i(\mathbf{u})|^2 d\mathbf{u}$.

2.2 Breaking Down the Constant Brightness Assumption

Few, if any, real world scenes are truly Lambertian. In this section, we present a modification of the Lambertian model (1) to take into account deviations from the constant brightness assumption without explicitly modelling the reflectance properties of the scene. We consider predicted images that would be produced by a perfect Lambertian object, plus an additional corrective component. This additional term contains the non-Lambertian component of the radiance. Also, we estimate both the Lambertian color C and the corrective component $F_i : \mathcal{I}_i \rightarrow \mathbb{R}^3$. In order to constrain the problem, we assume the corrective component to be smooth by adding the energy term $E_{RF} = \sum_i \int_{\mathcal{I}_i} |\nabla F_i(\mathbf{u})|^2 d\mathbf{u}$. The surface radiance is estimated using the corrected image $R_i : \mathcal{I}_i \rightarrow \mathbb{R}^3$ such that $R_i = I_i - F_i$. In other words, in equation (1) we replace I_i by R_i.

Additionally to correct non-Lambertian reflectance effects, images F_i can also absorb other complex photometric phenomena neglected in the camera model, which makes our method able to recover quite robustly the shape and the diffuse radiance C of non-Lambertian scenes viewed from flawed cameras. Contrary to other methods which try to be robust to specular effects by modelling the reflectance using parametric models [21], the algorithm we propose here is quite generic and can deal with many materials without any modification. Also, it does not need complex similarity measures such as normalized cross-correlation or mutual information as [12], is simpler (at least from the implementation point of view), and naturally allows us to deal with significant degradations of some parts of images, as for example can be the case with occlusions of some cameras.

3 Minimization of the Total Energy

For optimizing our total energy $E_{total} = E_{data} + \lambda_{\mathcal{S}} E_{R\mathcal{S}} + \lambda_B E_{RB} + \lambda_F E_{RF}$ we perform gradient descents alternatively with respect to F_i, B_i and $\mathcal{S}$. The computation of the gradients with respect to F_i and B_i is classical since they are image-based [20]. For a fixed shape $\mathcal{S}$ and a fixed C, we have

$$\begin{aligned} \nabla E_{total}(F) &= -\sum_i \left[(I_i - C \circ \pi_{i,\mathcal{S}}^{-1} - F_i)h + \lambda_F \Delta F_i \right] \\ \nabla E_{total}(B) &= -\sum_i \left[(I_i - B_i)(1-h) + \lambda_B \Delta B_i \right] \ , \end{aligned} \tag{2}$$

where h is the characteristic function that indicates if $\mathbf{u}$ is covered by the projection of the surface $\mathcal{S}$ ($h(\mathbf{u}) = 1$) or not ($h(\mathbf{u}) = 0$, $\mathbf{u}$ being explained by the background); λ is the smoothness parameter; $C(\mathbf{x})$ is computed taking the mean color of the projection in the corrected image R_i where $\mathbf{x}$ is visible. In contrast, the computation of the gradient of E_{total} with respect to $\mathcal{S}$ (more exactly of E_{data} since E_{RB} and E_{RF} do not depend on $\mathcal{S}$; $E_{R\mathcal{S}}$ is classical [2, 20]) is difficult and requires some work. Actually, the computation of the exact gradient of this kind of functional is the goal of [6, 20], where the computation is done in the purely theoretical framework of the continuous (smooth) surfaces as the one described by [17] using Gâteaux derivatives.

Here, contrary to [6, 20], we are going to directly compute the exact gradient of a surface represented by a mesh (of course the surface being non smooth, the results of [6, 20] do not apply). Similarly as [6], we first rewrite the energy as an integral over the visible volume. Then in section 4 we will compute the exact gradient of E_{data} with respect to the *discrete* surface mesh. For simplicity of notation, we are going to compute the gradient for a single camera and so we remove the dependency on i. For several cameras, the gradient will be the sum of the gradients associated with each camera.

Rewriting the Energy as an Integral over the Visible Volume

The first step to be able to optimize our energy functional with respect to the shape $\mathcal{S}$ is to rewrite the energy as an integral over the surface instead of the image. This change of variables involves adapting the measure on the surface [16] and in counting only the visible points [6, 12, 20]. This can be achieved by $d\mathbf{u} = -\frac{\mathbf{x}\cdot\mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3}\nu_{\mathcal{S}}(\mathbf{x})d\sigma$ where $d\sigma$ is the classical surface area measure, $\nu_{\mathcal{S}}$ is the visibility function (giving 1 for an $\mathbf{x}$ that is visible from the camera and 0 otherwise), $\mathbf{n}(\mathbf{x})$ is the unit normal vector to the surface $\mathcal{S}$ at $\mathbf{x}$, and $\mathbf{x}_z$ is the depth of $\mathbf{x}$. Thus, by using the separation technique proposed by [20], the energy functional becomes (for a *single* image):

$$\begin{aligned} E_{data}(\mathcal{S}) &= -\int_{\mathcal{S}} g_R(\mathbf{x}) \frac{\mathbf{x}\cdot\mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3} \nu_{\mathcal{S}}(\mathbf{x}) d\sigma \ + \int_{\mathcal{I}-\pi(\mathcal{S})} g_B(\mathbf{u}) d\mathbf{u} \ , \\ &= -\int_{\mathcal{S}} [g_R(\mathbf{x}) - g_B(\pi(\mathbf{x}))] \frac{\mathbf{x}\cdot\mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3} \nu_{\mathcal{S}}(\mathbf{x}) d\sigma \ + \int_{\mathcal{I}} g_B(\mathbf{u}) d\mathbf{u} \ , \end{aligned} \tag{3}$$

where $g_R(\mathbf{x})$ is $1/2\,[R(\pi(\mathbf{x})) - C(\mathbf{x})]^2$ and $g_B(\mathbf{u})$ is $1/2\,[R(\mathbf{u}) - B(\mathbf{u}))]^2$. The right-hand term of equation (3) does not depend on $\mathcal{S}$, so in the following we intentionally omitted it as it does not contribute to the gradient expression (with respect to $\mathcal{S}$). Hence, denoting

$g(\mathbf{x}) = g_R(\mathbf{x}) - g_B(\pi(\mathbf{x}))$ for convenience, the energy to minimize with respect to $\mathcal{S}$ is

$$E(\mathcal{S}) = -\int_{\mathcal{S}} g(\mathbf{x}) \frac{\mathbf{x} \cdot \mathbf{n}(\mathbf{x})}{\mathbf{x}_z^3} \nu_{\mathcal{S}}(\mathbf{x}) d\sigma. \tag{4}$$

Finally, by using Gauss' divergence theorem which states that for all domains $\Omega \in \mathbb{R}^3$ and for all vector fields $\mathbf{w} : \mathbb{R}^3 \to \mathbb{R}^3$, $\int_\Omega \nabla \cdot \mathbf{w} d\mathbf{x} = \int_{\partial\Omega} \mathbf{w} \cdot \mathbf{n} d\sigma$, the error functional (4) can be rewritten as :

$$E(\mathcal{S}) = -\int_{\mathbb{R}^3} \left(\nabla g(\mathbf{x}) \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3} \right) \nu_{\mathcal{S}}(\mathbf{x})\, d\mathbf{x} = -\int_{\mathbb{R}^3} f(\mathbf{x})\, \nu_{\mathcal{S}}(\mathbf{x})\, d\mathbf{x}\,. \tag{5}$$

For convenience, we have denoted $f(\mathbf{x}) = \nabla g(\mathbf{x}) \cdot \frac{\mathbf{x}}{\mathbf{x}_z^3}$. Let us note that above results hold for any piecewise continuous (smooth) surface.

4 Gradient of the Reprojection Error for Meshes

In this section we compute the discrete gradient of $E(\mathcal{S})$ (and then of E_{data}) with respects to the surface shape, in the case where the surface $\mathcal{S}$ is represented as a triangulated mesh $\mathbf{X}$. Vertices of $\mathbf{X}$ are denoted by $\mathbf{x}_k$ and $\mathcal{S}$ is deformed by moving vertices $\mathbf{x}_k$ according to some vectors $\mathbf{V}_k$. The method for computing the gradient of $E(\mathcal{S})$ with respect to $\mathbf{x}_k$ consists in computing the directional derivative of $E(\mathcal{S}[t])$ for any deformation $\mathbf{x}_k[t] = \mathbf{x}_k^0 + t\mathbf{V}_k$ and then in rewriting it as a scalar product of $\mathbf{V}_k$, i.e. as $\langle \mathbf{V}_k, G \rangle$. The obtained vector G is called the *gradient* and the energy necessarily decreases when deforming the surface according to its opposite direction $-G$. Indeed, for $\mathbf{x}_k[t] = \mathbf{x}_k^0 - tG$, we have $[E \circ \mathcal{S}]'(0) = -\langle G, G \rangle \leq 0$, see [4].

To compute the gradient, we have to understand the variations of our energy when $\mathcal{S}$ is deforming. Since our energy $E(\mathcal{S})$ is an integral over the visible volume, its variations are only due to the variations of the visible volume. Also, as illustrated by Figure 1 (reduced to the 2D case for simplicity), when the vertex $\mathbf{x}_k$ is moving according to $\mathbf{V}_k$ we have to separate two cases: 1) when all the triangles adjacent to $\mathbf{x}_k$ are visible, the variation of the visible volume is just the sum of the tetrahedra formed by the adjacent triangles and the moved point $\mathbf{x}_k + \mathbf{V}_k$ (see Fig. 1, left). The corresponding gradient computation is detailed in Sec. 4.1. 2) when $\mathbf{x}_k$ is generating occluding contours in images, i.e. when it is a *horizon* point, its movement affects the visibility of other points located behind it (called *terminator* points). So the variation of the visible volume is the sum of the first case term, plus the volume swept out by the crepuscular rays generated by the horizon movement (see Fig. 1, right). The corresponding gradient computation is detailed in Sec. 4.2.

4.1 Term Due to the Tetrahedra of the Visible Adjacent Triangles

In the sequel, we denote $\mathcal{S}_j$ the jth triangle of the mesh. Following the variation of the energy caused by the visible adjacent triangles of $\mathbf{x}_k$, we have

$$E(\mathcal{S}[t]) - E(\mathcal{S}[0]) = \sum_j \varpi_{j,\mathbf{V}_k} \int_{Vol[j,\mathbf{V},t]} f(\mathbf{x})\, d\mathbf{x} \; + B(t), \tag{6}$$

where $Vol[j, \mathbf{V}, t]$ is the volume of the tetrahedron formed by the vertices of the *visible* triangle $\mathcal{S}_j$ and the point $\mathbf{x}_k + \mathbf{V}_k$. The sign $\varpi_{j,\mathbf{V}_k}$ specifies if matter has been added to

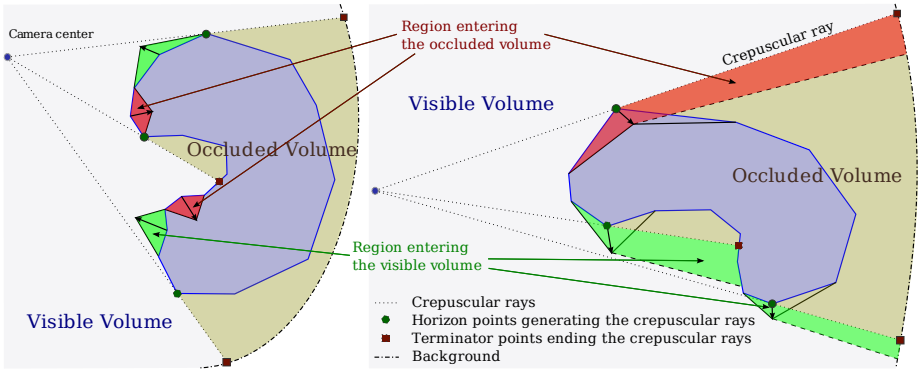

Figure 1: Geometric representation of the change of visibility when moving the mesh. Contrary to the interior (left), movements of the horizon (right) strongly affect the movement of the visible interface between visible and occluded volumes by creating a movement of the crepuscular rays.

or removed from the object volume and it is equal to $N_j \cdot \frac{\mathbf{V}_k}{|\mathbf{V}_k|}$, where N_j is the outward surface normal of triangle $\mathcal{S}_j$. $B(t)$ on the right of equation (6) is null except if the vertex $\mathbf{x}_k$ is a horizon point; this additional part will be detailed in the next paragraph.

Now, we parameterize the volume $Vol[j, \mathbf{V}, t]$ by $\mathbf{x}(u, v, w) = \mathbf{x}(u, v) + w\ t\mathbf{V}_k$, where $\mathbf{x}(u, v) = \mathbf{x}_k + u\ \overrightarrow{\mathbf{x}_k\mathbf{x}_{k1}} + v\ \overrightarrow{\mathbf{x}_k\mathbf{x}_{k2}}$ parametrizes the triangle $\mathcal{S}_j$; $\mathbf{x}_{k1}$ and $\mathbf{x}_{k2}$ being the other two vertices of $\mathcal{S}_j$; $(u, v) \in T = \{(u, v) \mid u \in [0, 1] \text{ and } v \in [0, 1 - u]\}$ and $w \in [0, 1 - u - v]$. By a change of variables, (6) becomes

$$\sum_j \varpi_{j,\mathbf{V}_k} \int_T \int_0^{1-u-v} f(\mathbf{x}(u, v) + wt\mathbf{V}_k) |det(\overrightarrow{\mathbf{x}_k\mathbf{x}_{k1}}, \overrightarrow{\mathbf{x}_k\mathbf{x}_{k2}}, t\mathbf{V}_k)|\, dwdudv + B(t). \tag{7}$$

Let A_i be the area of the triangle $\mathcal{S}_j$. Since $\varpi_{j,\mathbf{V}_k} |det(\overrightarrow{\mathbf{x}_k\mathbf{x}_{k1}}, \overrightarrow{\mathbf{x}_k\mathbf{x}_{k2}}, t\mathbf{V}_k)| = -2t\ A_j \mathbf{V}_k \cdot N_j$, equation (7) becomes

$$-2t\ \mathbf{V}_k \cdot \left[\sum_j A_j N_j \int_T \int_0^{1-u-v} f(\mathbf{x}(u, v) + wt\mathbf{V}_k)\, dwdudv \right] + B(t). \tag{8}$$

It follows that the limit of $\frac{E(\mathcal{S}[t]) - E(\mathcal{S}[0])}{t}$ when t tends to zero is

$$\mathbf{V}_k \cdot \left[-2 \sum_j A_j N_j \int_T \nabla g(\mathbf{x}(u, v)) \frac{\mathbf{x}(u, v)}{\mathbf{x}(u, v)_z^3} (1 - u - v)\, dudv \right] + \lim_{t \to 0} B(t), \tag{9}$$

the part in square brackets being the considered first part of the gradient.

4.2 Term Due to the Movement of the Horizon

Let $H_{k,j}$ be the vector such that $[\mathbf{x}_k, \mathbf{x}_k + H_{k,j}]$ is the edge of the triangle $\mathcal{S}_j$ generating the horizon. The volume corresponding to the movement of the horizon can be parameterized by the points $\mathbf{y}(u, v)$ of the triangle $\{\mathbf{x}_k, \mathbf{x}_k + H_{k,j}, \mathbf{x}_k + v\ t\mathbf{V}_k\}$ generated by the movement of the horizon. More rigorously, it can be parameterized as the set of points

$\mathbf{x}(u,v,r) = r\ \mathbf{y}(u,v)$ where $\mathbf{y}(u,v) = \mathbf{x}_k + u\ H_{k,j} + v\ t\mathbf{V}_k$; r corresponds to the depth of $\mathbf{x}$ in the view point direction; $r \in [1, \mathcal{T}_{(u,v)}]$.
Let us note that $\mathbf{y}(u,v)$ depends on t; we emphasize this dependency by denoting $\mathbf{y}_t(u,v)$. By a change of variable, we get $E(\mathcal{S}[t]) - E(\mathcal{S}[0]) =$

$$A(t) + 1/2 \int_{(u,v)\in T} \int_{r\in[1,\mathcal{T}_{(u,v)}]} f(r\ \mathbf{y}_t(u,v))\ (H_{k,j} \wedge\ t\mathbf{V}_k) \cdot \frac{\mathbf{y}_t(u,v)}{|\mathbf{y}_t(u,v)|} r^2 dudvdr\ ,$$

where $A(t)$ corresponds to the part described in the previous paragraph. It follows that the limit of $\frac{E(\mathcal{S}[t])-E(\mathcal{S}[0])}{t}$ when t tends to zero is

$$\lim_{t\to 0} A(t) + 1/2 \int_{(u,v)\in T} \int_{[1,\mathcal{T}_u]} f(r\ \mathbf{y}(u))\ (H_{k,j} \wedge \mathbf{V}_k) \cdot \frac{\mathbf{y}(u)}{|\mathbf{y}(u)|} r^2 dudvdr\ , \qquad (10)$$

where $\mathbf{y}(u) = \mathbf{x}_k + u\ H_{k,j}$ and $\mathcal{T}_u = \mathcal{T}_{(u,0)}$. Denoting $L(u) = \int_{1,\mathcal{T}_u} f(r\ \mathbf{y}(u)) r^2 dr$, the right-hand part of equation (10) can be rewritten as

$$\mathbf{V}_k \cdot \frac{1}{2} \int_{u\in[0,1]} L(u) \left(\frac{\mathbf{y}(u)}{|\mathbf{y}(u)|} \wedge H_{k,j} \right) (1-u) du\ . \qquad (11)$$

Finally, remembering that $f(\mathbf{x}) = -\nabla g(\mathbf{x}) \frac{\mathbf{x}}{\mathbf{x}_z^3}$, we can explicit

$$L(u) = \left[g(T(\mathbf{y}(u))) - g(\mathbf{y}(u)) \right] \frac{1}{[\mathbf{y}(u)]_z^3}\ , \qquad (12)$$

where $T(\mathbf{x})$ is the terminator of $\mathbf{x}$. Then the complete gradient flow optimizing E_{data} (with respect to $\mathbf{x}_k$) is the term in brackets of $\lim_{t\to 0} A(t)$ in (9) obtained previously, plus the following term if $\mathbf{x}_k$ belongs to the horizon :

$$\sum_{\text{horizon edges } H_{k,j}} \frac{1}{2} \int_{u\in[0,1]} \left(g(T(\mathbf{y}(u))) - g(\mathbf{y}(u)) \right) \frac{\mathbf{y}(u) \wedge H_{k,j}}{|\mathbf{y}(u)|\ [\mathbf{y}(u)]_z^3} (1-u) du\ . \qquad (13)$$

5 Experiments

We have implemented our algorithm using the Delaunay topology-adaptive meshes proposed by [11]. The visibility is computed using OpenGL Z-buffer. The evolution is done using a multi-resolution scheme and starting from the visual hull. Horizons are located using the changes of signs of the dot products of facet normals and viewpoint directions.

Tests on the Discrete Gradient

As in [6], we first reconstruct three uniformly colored balls arranged on a plane (20 640×480 images), see Fig. 2. This way we ensure the importance of the horizon term as the color gradient is null over the surface except at interfaces between objects in images. With the only use of the interior term (section 4.1), the surface shrinks. By using only the horizon term (given in section 4.2), we correctly reconstruct and separate the balls, and apparent contours correctly reproject in images. Then we tested our algorithm on synthetic Lambertian data for the Stanford dragon mesh (Fig. 3) composed of 32 640×480

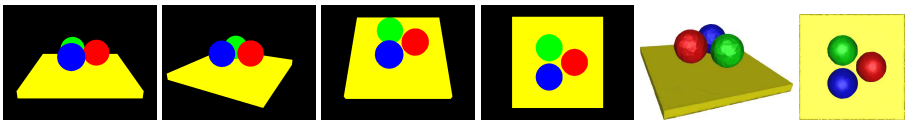

Figure 2: The balls sequence. From left to right : 4 of 20 input images; results with the horizon term computed in section 4.2 from different viewpoints.

Figure 3: Synthesized dragon sequence. From left to right : 1 of 32 diffuse input images; ground truth model; recovered shape by our algorithm;

Figure 4: Synthesized dragon sequence. From left to right : 2 of 32 original images; components added for simulating the photometric camera effects; input images; estimated corrective component (scaled by 2 for visualization); estimated background images; estimated shape and radiance.

images. The result shows the correct reconstruction of the dragon, even though the texture is smooth and some parts in shadow are dark (See Tab. 1).

Non-Lambertian surfaces with photometric camera problems

In order to test the corrective model of Sec. 2.2, we modify some of the input images used in the previous tests to simulate photometric camera problems. 8 of the 32 input images have been changed by adding a various smooth colored gradient. We not only recover the shape, but we also correctly estimate the radiance of the object, cf. Fig. 4 and Tab. 1. The corrective images are recovered on the foreground region. As expected from Sec. 3, the recovered background images corresponds to the one of the input images on the complementary regions. In a second experiment, we add a specular reflectance model to the dragon images and modify 3 of the 32 images such that half of these images is black. This could simulate occluding objects in front of the camera. Results obtained with the perfect Lambertian model of Sec. 2.1 are much less accurate than the ones obtained with the corrective model of Sec. 2.2. Also, the corrective model allows to correctly separate and recover the diffuse and specular components of the images (see Fig. 5 and Tab. 1).

Finally, we tested our method on the classical Dino (16 640×480 images) and Temple (16 640×480 images) datasets (Fig. 6). We refer to the Middlebury benchmark website [13] for evaluation. Even if these datasets were intentionally chosen for being Lambertian, they exhibit a certain amount of specular reflection that can be interpreted in the resulting corrective component. This component also shows a greenish upper part and a magenta

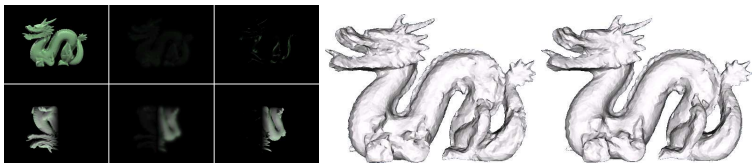

Figure 5: Dragon sequence synthesized with specular reflectance model and occlusions. From left to right : 2 of 32 input images; recovered corrective images; ground truth corrective images; recovered shape without the corrective model; recovered shape with the corrective model.

	Diffuse images (Fig. 3)	Diffuse images Corrective model (Fig. 4)	Specular images Lambertian model (Fig. 5)	Specular images Corrective model (Fig. 5)
Acc. 95%	0.241mm	0.199mm	0.410mm	0.300mm
Comp. 0.5mm	98.3%	98.7%	97.2%	97.6%

Table 1: Numerical evaluation of the proposed method for the dragon sequence that shows accuracy at 95% and completeness at 0.5mm following [13]. (compared to ground truth)

Figure 6: Dino and Temple sequence. From left to right: 1 of 16 input images; estimated non-Lambertian component (scaled by 8 for visualization); estimated background images (scaled by 2 for visualization); estimated radiance; estimated mesh seen from a different viewpoint.

lower part which might be due to some photometric camera bias as it appears in all images.

6 Conclusion

In this paper we have presented a novel deformable surface approach for the multiview stereo problem based on a *discrete* representation. Contrary to previous works, during the evolution, *we correctly deal with visibility changes* by expressing the exact gradient of the reprojection error functional. In particular, exactly as in the continuous case [6], this forces the contour generators of the surface to appear at their correct location in the images and removes the minimal surface bias from which most previous methods suffer. In

addition, our method is robust to *smooth non-Lambertian effects* (such as smooth specular reflections) and *photometric camera defects* (such as photometric calibration errors and partial occlusions) by introducing an original and simple *image-based corrective model* which has proved to be effective on the tests we have made on synthetic and real images.

References

[1] N. Birkbeck, D. Cobzas, P. Sturm, and M. Jägersand. Variational shape and reflectance estimation under changing light and viewpoints. In *ECCV*, volume 1, pages 536–549, 2006.

[2] M. Desbrun, M. Meyerm, P. Schröder, and A. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *SIGGRAPH '99*, pages 317–324, 1999.

[3] Y. Duan, L. Yang, H. Qin, and D. Samaras. Shape reconstruction from 3d and 2d data using pde-based deformable surfaces. In *ECCV*, pages 238–251, Prague, Czech Republic, 2004.

[4] I. Eckstein, J.-P. Pons, Y. Tong, C.C. Jay Kuo, and M. Desbrun. Generalized surface flows for mesh processing. In *Symp. on Geometry Processing*, pages 183–192, 2007.

[5] O.D. Faugeras and R. Keriven. Variational-principles, surface evolution, pdes, level set methods, and the stereo problem. *IEEE TIP*, 7(3):336–344, March 1998.

[6] P. Gargallo, E. Prados, and P. Sturm. Minimizing the reprojection error in surface reconstruction from images. In *IEEE ICCV*, pages 1–8, Rio de Janeiro, Brazil, 2007.

[7] C. Hernández and F. Schmitt. Silhouette and stereo fusion for 3d object modeling. *Comput. Vis. Image Underst.*, 96(3):367–392, 2004.

[8] C. Hernandez, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE TPAMI*, 30(3):548–554, 2008.

[9] H. Jin, S. Soatto, and A. J. Yezzi. Multi-view stereo reconstruction of dense shape and complex appearance. *IJCV*, 63(3):175–189, 2005.

[10] H. Jin, A.J. Yezzi, and S. Soatto. Variational multiframe stereo in the presence of specular reflections. In *3DPVT*, pages 626–630, 2002.

[11] J.-P. Pons and J.-D. Boissonnat. Delaunay deformable models: Topology-adaptive meshes based on the restricted delaunay triangulation. In *IEEE CVPR*, Minneapolis, USA, Jun 2007.

[12] J.-P. Pons, R. Keriven, and O. Faugeras. Modelling dynamic scenes by registering multi-view image sequences. In *IEEE CVPR*, pages 822–827, San Diego, USA, Jun 2005.

[13] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *IEEE CVPR*, pages 519–528, 2006.

[14] S. N. Sinha and M. Pollefeys. Multi-view reconstruction using photo-consistency and exact silhouette constraints: a maximum-flow formulation. In *IEEE ICCV*, pages 349–356, 2005.

[15] G. Slabaugh and G. Unal. Active polyhedron: surface evolution theory applied to deformable meshes. *IEEE CVPR*, 2:84–91 vol. 2, 2005.

[16] S. Soatto, A. J. Yezzi, and H. Jin. Tales of shape and radiance in multi-view stereo. In *IEEE ICCV*, pages 974–981, 2003.

[17] J. E. Solem and N. Chr. Overgaard. A geometric formulation of gradient descent for variational problems with moving surfaces. In *Scale-Space*, pages 419–430, 2005.

[18] G. Vogiatzis, C. Hernández Esteban, P. H. S. Torr, and R. Cipolla. Multiview stereo via volumetric graph-cuts and occlusion robust photo-consistency. *PAMI*, 29(12):2241–2246, 2007.

[19] R. Yang, M. Pollefeys, and G. Welch. Dealing with textureless regions and specular highlights-a progressive space carving scheme using a novel photo-consistency measure. In *IEEE ICCV*, pages 576–583, 2003.

[20] A. Yezzi and S. Soatto. Stereoscopic segmentation. *IJCV*, 53(1):31–43, 2003.

[21] T. Yu, N. Xu, and N. Ahuja. Shape and view independent reflectance map from multiple views. *IJCV*, 73(2):123–138, 2007.

[22] A. Zaharescu, E. Boyer, and R. P. Horaud. Transformesh: a topology-adaptive mesh-based approach to surface evolution. In *ACCV*, pages 166–175, Tokyo, Japan, Nov 2007. Springer.

[23] T. Zickler, S. P. Mallick, D. J. Kriegman, and P. Belhumeur. Color subspaces as photometric invariants. *To appear in IJCV*, 2008.