Image-Based Multiresolution Implicit Object Modeling

Aurelio Colosimo, Augusto Sarti, Stefano Tubaro

Dipartimento di Elettronica e Informazione
Politecnico di Milano
Piazza L. Da Vinci 32, 20133 Milano
Italy

ABSTRACT

In this paper we propose an image-based 3D modeling method based on a multi-resolution evolution of a level-set of a volumetric function, steered by the texture mismatch between views. A key feature of this approach is that it operates in an adaptive multi-resolution fashion, which boosts up the computational efficiency.

1. INTRODUCTION

A 3D manifold can be generally defined and represented either *explicitly* as an atlas (juxtaposition of partially overlapping local charts), or *implicitly* as the set of points that satisfy a non-linear constraint in the 3D space (level set of a volumetric function). Similarly, image-based modeling of 3D objects can be envisioned as based on either one of the above two representations. In the former case, a global object model is obtained as a complex "patchworking" of simple local reconstructions (typically depth maps), while in the latter the object surface is described as a level set of an appropriate volumetric function.

As we may expect, an atlas-based 3D modeling method deals with topological complexity with a "divide-and-conquer" strategy, which simplifies the local shape estimation process. The price to pay for this simplification, however, is in the steps that are necessary to combine the local reconstruction into a global closed one (registration, fusion and hole-mending), which are usually quite a burden. An implicit surface representation, on the other hand, tends to be quite insensitive to topological complexity, as it may accommodate self-occluding surfaces, concavities, surfaces of volumes with holes (e.g. doughnuts, objects with handles, etc.), or even multiple objects. However, an implicit surface representation is intrinsically more redundant than an atlas-like manifold, as we need to specify a volumetric function (whose zero level set is the desired surface) rather than a collection of depth maps.

An image-based 3D modeling method that uses an implicit representation of surfaces was recently proposed by Faugeras and Keriven [1]. This modeling approach is based on the temporal evolution of a volumetric function whose zero level-set is a closed surface that represents the surface model as it tends to approximate the imaged object. This surface, which initially contains the imaged object, evolves by following a motion that is always locally normal to the surface, with a speed that depends on the local surface curvature and to a measurement of the local "texture mismatch" between imaged and "transferred" textures. Transferring an imaged texture onto another view means back-projecting it onto the model and re-projecting it onto the other view. In order to keep the computational complexity at a manageable level, the updating of the volumetric function is only performed within a "narrow band" [3] around the current surface. Our solution, however, significantly generalizes this approach, as it operates in an adaptive multi-resolution fashion, which boosts up the computational efficiency. Multi-resolution, in fact, enables us to quickly obtain a rough approximation of the objects in the scene at the lowest possible voxset resolution. Successive resolution increments allow us to progressively refine the model and add details. In order to do so, we introduce "inertia" in the level-set evolution, which tends to favor topological changes (e.g. the creation of doughnut-like holes in the structure). Finally, through a careful control of the components that steer the level-set evolution (hysteresis, biased quantization, etc.), we are able to recuperate details that were lost at lower resolution levels (surface creases, ridges, etc.).

2. OVERVIEW OF THE MODELING STRATEGY

One of the terms that contribute to steering the level-set evolution is the "texture mismatch" between imaged and "transferred" textures [2], which is a function of the correlation between homologous luminance profiles [1]. The texture mismatch is

$$\begin{aligned} C(S,\mathbf{n}) &= \iint_S \Phi(S,\mathbf{n},v,w)|\mathbf{s}_v \times \mathbf{s}_w| dv dw \\ \Phi &= 1 - \sum_{i,j=1;i\neq j}^{n} \frac{1}{|I_i||I_j|}\langle I_i, I_j\rangle \ . \end{aligned} \quad (1)$$

where $\mathbf{d}\sigma = |\mathbf{s}_v \times \mathbf{s}_w| dv dw$ is the infinitesimal area element of the surface S, associated to the local surface parametrization (v,w) induced by the image coordinate chart; $\mathbf{N}$ is the surface normal; and $I_i(\mathbf{m}_i)$ is the luminance of pixel $\mathbf{m}_i$ in the i-th image. This definition of $\mathbf{d}\sigma$ guarantees that the surface representation will be independent on the variables (u,w).

The inner product (correlation) between the pair of subimages I_i and I_j is defined as follows:

$$\begin{aligned} \langle I_i, I_j\rangle &= \frac{1}{4pq}\int_{-p}^{p}\int_{-q}^{q} \left(I_i(\mathbf{m}_1+\mathbf{m}) - \overline{I_i}(\mathbf{m}_1)\right) \\ &\quad \cdot \left(I_j(\mathbf{m}_2+\mathbf{m}) - \overline{I_j}(\mathbf{m}_2)\right) \mathbf{dm}, \end{aligned} \quad (2)$$

where $\mathbf{m}_1$ e $\mathbf{m}_2$ are homologous image points (i.e. image points that correspond to the same point of the surface model), and

$$\overline{I_k}(\mathbf{m}_k) = \frac{1}{4pq}\int_{-p}^{p}\int_{-q}^{q} I_k(\mathbf{m}_k+\mathbf{m})\mathbf{dm} \qquad k = 1,2$$

Although the correlation could be computed between all the viewpoints where there is visibility, only the pair of views with the best visibility is considered. Visibility can be easily checked through a ray-tracing algorithm and measured as a function of the angle between visual ray and surface normal. Notice that normalizing the correlation has a twofold purpose: to limit its range between 0 and 2; and to guarantee that low-energy areas (smooth texture) will have the same range of high-energy (rough texture) areas. Finally, subtracting the average from a luminance profile tends to compensate a non-lambertian behavior of the imaged surfaces.

The model surface at time t (propagation front) is described as the zero level-set of a volumetric function $u(\mathbf{x},t)$, $\mathbf{x} = \left[\begin{array}{ccc} x_1 & x_2 & x_3 \end{array}\right]^T$, i.e. as the set of points $\mathbf{x}$ that satisfy the implicit equation $u(\mathbf{x},t) = 0$. The goal is to characterize a temporal evolution for u such that that its zero level-set will tend to approximate the viewed object. We define this evolution through the front propagation equation $u(\mathbf{x}(t+\triangle t)) = u(\mathbf{x}(t)) - |\nabla u(\mathbf{x}(t))|\,\beta(\mathbf{x}(t))\triangle t$, where $\mathbf{x}(t)$ describes the trajectory of the point $\mathbf{x}$ as time passes. In other words, the level-set evolves by following a motion that is always normal to the surface, whose velocity is given by $\beta(\mathbf{x})$. Our choice of this velocity wase made with the twofold need of guaranteeing surface smoothess and consistency between images and final model

$$\beta(\mathbf{x}) = \Phi div(\mathbf{N}) + \nabla\Phi\cdot\mathbf{N} + k\Phi \ . \quad (3)$$

The first term $\Phi div(\mathbf{N})$ tends to favor the motion by curvature. The presence of the texture mismatch Φ, however, tends to slow down areas with modest cost, and speed up areas of high cost. Ideally, one would be lead to think that the first term is sufficient correctly steer the model's evolution, as correct surface regions should have a zero cost, while other reagions are left free to evolve. This, however, is not really true as the cost is rarely equal to zero due to a non-perfect luminance transferral and a non-lambertian radiometric behavior. This causes the front propagation to "trespass" the correct surface. The second term of eq. (3) will tend to contrast this behavior. Finally, the third element of eq. (3) acts an "inertial" term in order to favor concavities in the final model, provided that a proper dynamic adaptation of k is performed.

If the volumetric function that characterizes the level-set is defined on a static voxset of N voxels, the computational complexity of each front propagation step is proportional to N^2, as it is proportional to the surface of the level-set (narrow-band computation). Furthermore, since the velocity β is multiplied by $|\nabla u|$ (equal to the sampling step), the number of iterations turns out to be proportional to N, with a resulting algorithimc complexity that is proportional to N^3. In order to dramatically reduce this complexity, we developed a multi-resolution approach to level-set evolution. The algorithm starts with a very low resolution level (a voxset of 10–15 voxel per side). When the propagation front converges, the resolution increases and the front resumes its propagation. The process is iterated until we reach the desired resolution. A progressive resolution increment has the desired result of minimizing the amount of changes that each propagation step will introduce in the model, with the result of achieving a better global minimum of the cost function. Furthermore, the number of iterations will be dramatically reduced with respect to a fixedresolution approach, as it will turn out to be proportional to N^2.

Indeed, starting from a low-resolution voxset, we need to prevent the algorithm from losing details at that resolution or to make sure that the algorithm will be able to recover the lost details. In fact, one has to keep in mind that the motion by curvature tends to dominate over the other terms, therefore some of the details of the object may totally disappear. In order to prevent this from happening, we use a method based on thresholding the local curvature with a hyperbolic tangent function. This guarantees a smoother behavior than a simpler clipping function.

In spite of this smooth thresholding mechanism, in some cases it is not possible to prevent some of the smaller details from disappearing. For this reason, we developed a technique that enables the recovery of lost details before the resolution is increased, which is based on a mechanism of hysteresis in the surface implosion. The idea is to keep track of all the voxels on the zero level-set whose cost Φ is below a certain threshold. After the "implosion" of the level-set, we let the propagation front evolve while driven by a different cost function that depends on the distance between the surface and such points. This operation makes the surface litterally "climb up" the lost details. As an example of applications, see Fig. 1.

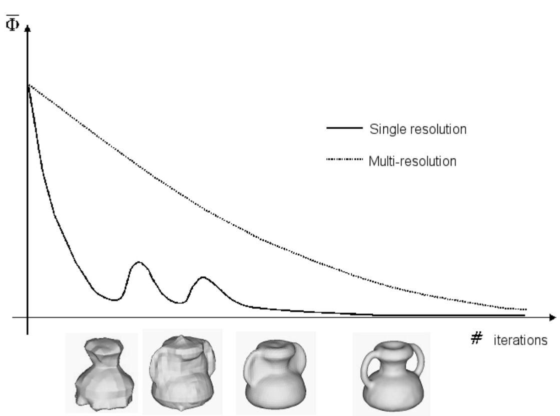

Figure 1: Illustration of the temporal evolution of the cost function (texture mismatch) and of the model. Notice that the cost value suddenly increases at every resolution change, due to the mechanism of recovery of details.

3. EXAMPLES OF APPLICATION

We tested our approach with several real subjects acquired with a camera moving around them. The method exhibited a remarkable robustness against topological complexity and segmentation problems. The results are shown in Figs. 2 to 8.

Figure 2: Sequence of original views.

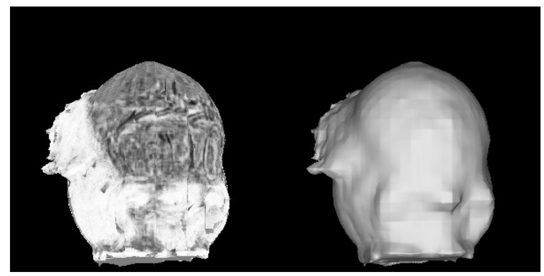

Figure 3: A view of the cost function mapped onto the propagation front. The darker the texture, the heavier the mismatch.

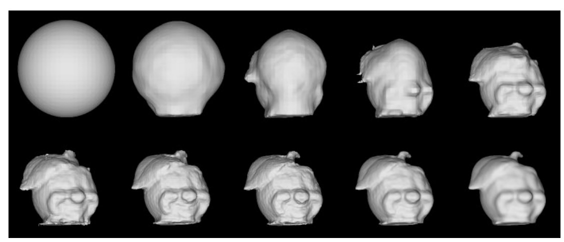

Figure 4: Temporal evolution of the propagation front. The initial volumetric resolution is very modest (in this case the voxset size is 20x20x20), and is not able to account for some topologically complex details of the surface (the fifth frame in lexicographic order is the best one can do at this resolution). As the resolution increases, more details begin to appear, such as the stem of the apple.

4. CONCLUSIONS

In this paper we proposed an image-based 3D modeling method based on a multi-resolution evolution of a level-set of a volumetric function, steered by the texture mismatch between views. The proposed solution operates in an adaptive

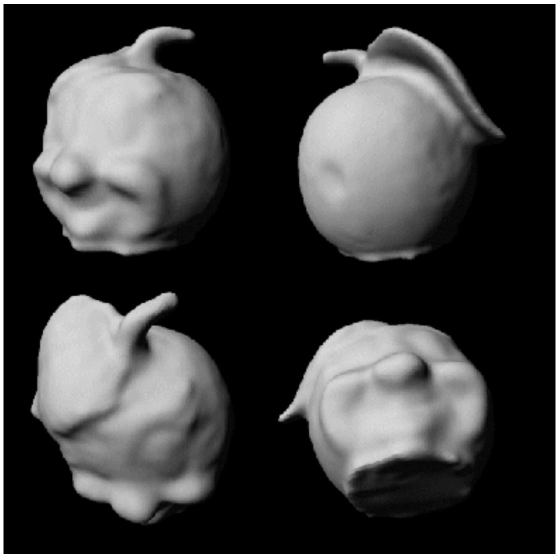

Figure 5: Final model.

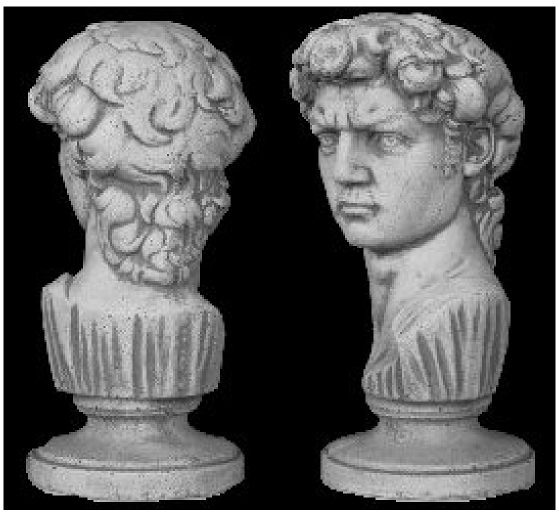

Figure 6: One of the original views of the subject.

multi-resolution fashion, which brings its computational efficiency to practical usability.

5. REFERENCES

[1] O. Faugeras, R. Keriven. "Variational Principles, Surface Evolution, PDE's, level set Methods, and the Stereo Problem". *IEEE Tr. on Image Processing*, Vol. 7, No. 3, March 1998.

[2] P. Pigazzini, F. Pedersini, A. Sarti, S. Tubaro: "3D Area Matching with Arbitrary Multiview Geometry". *Signal Processing: Image Communications.* Vol. 14, Nos. 1-2, 1998, pp. 71-94.

[3] R. Malladi, J.A. Sethian, B.C. Vemuri, "Shape Modeling with Front Propagation: A Level Set

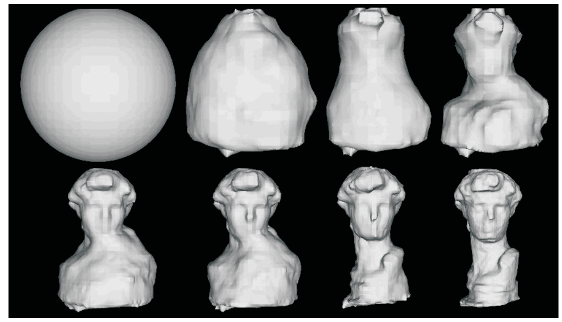

Figure 7: Temporal evolution of the propagation front.

Figure 8: Final model.

Approach". *IEEE Tr. on PAMI*, Vol. 17, No. 2, pp.158-175, Feb. 1995.