3D PASSIVE SHAPE RECOVERY FROM TEXTURE AND SILHOUETTE INFORMATION

L. Ballan*, N. Brusco*, G. M. Cortelazzo*

* Department of Information Engineering, University of Padova, v. Gradenigo 6A, 35100 Padova, Italy
e-mail: {ballanlu, brusco, corte}@dei.unipd.it
Fax: 00390498277699

Keywords: Stereo, Silhouette, Shape recovery, 3D Modeling, 3D Photography

Abstract

Recent efforts attempt to combine together informations of different passive methods. Critical issues in this research are the choice of data and how to combine such data in order to increase the overall information. The combination of stereo matching and silhouette information has recently received considerable attention both for obtaining high quality 3D models and for modelling 3D dynamic scenes, an application often referred to as 3D video. This work, building on recent results, reformulates 3D geometry recovery from stereo and silhouette informaton within a classical deformable model framework. Experimental verification shows that the theoretical advantages of the proposed solution improve its robustness and computational speed.

1 Introduction

The methods for recovering the 3D geometry of objects can be classified in various ways. A typical classification distinguishes passive from active methods. Passive sensing refers to the measurement of visible radiation which is already present in the scene; active sensing refers instead to the projection of structured light patterns onto the object or scene to be scanned.

Active sensing facilitates the computation of 3D structure by intrinsically solving the correspondence problem, a major issue with passive techniques. In general, active techniques such as those based on laser range scanning or light pattern projection tend to be more accurate but more expensive and slower than their passive counterparts. Furthermore, active sensing is not always feasible, e.g., for modeling distant or fast-moving objects. For these reasons and since passive techniques essentially require standard image capture devices such as photo-cameras or video-cameras, the interest towards passive 3D reconstruction techniques is bound to remain rather high. Historical passive sensing methods are stereo vision, structure from motion, shape from shading, space carving and shadow carving.

Recent efforts attempt to combine together informations of different passive methods. Critical issues in this research are what type of data to use and how to combine them, in order to actually increase the overall information. The combination of stereo matching and silhouette information has recently received considerable attention both for obtaining high quality 3D models [4] and for modelling 3D dynamic scenes [11], an application often referred to as 3D video.

This work re-addresses this approach within the classical deformable models framework, exploits silhouette information in a new way and proposes a solution which has some theoretical and computational advantages. This paper has six sections. The second section recalls the state-of-the-art in the formulation of 3D passive recovery from stereo and silhouette information and points out the most delicate issues. Sections 3 reformulates the problem within classical deformable models framework, defines a new force related to silhouette information, proves some theoretical advantages of the proposed reformulation and shows how to solve it. Section 4 adresses a fine grain improvement leading to a re-sampling more respectful of the geometrical quality of the mesh. Section 5 presents some experimental results. Section 6 draws the conclusions.

2 Shape recovery from stereo and silhouette information

The proposed 3D passive shape recovery procedure combines silhouette and stereo-matching information as schematically shown in Fig.1. The silhouettes are obtained by a segmentation algorithm [2] [8] from a sequence of photographs of the object taken from different positions depending on the characteristics of the scene. Stereo matching is also applied to the picture pairs of the sequence of photographs if the object is textured. Not all the pictures used for the silhouettes methods are used for stereo matching.

Silhouettes are first used by a shape-from-silhouette method [10] [7] [13] in order to obtain a coarse estimate of the surface; they are then used in order to correct for stereo-matching errors. The main advantages of shape-from-silhouette methods are that the obtained objects are well shaped and there are no problems with reflecting objects or objects without texture (if the segmentation algorithm is robust). The major drawback is that concavities cannot be modelled.

Texture is used by stereo matching methods [9], [16] which, differently from silhouette based techniques, can model

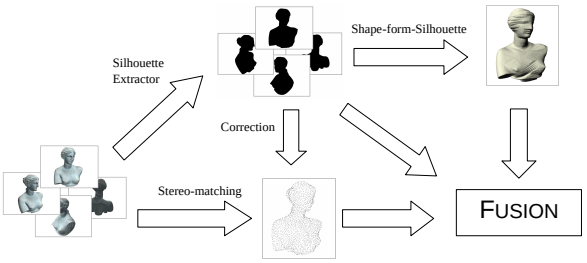

Figure 1: The proposed passive 3D modeling pipeline.

concavities. Stereo-matching does not work in regions without significant texture or where the available texture exhibits some periodicity. The latter can be partially avoided using a pyramidal approach [14]. 3D data near the silhouette edge are usually missing, since in these regions the object points can be easily mismatched with the background. Luckily, shape-from-silhouette methods can model these regions rather well.

In order to evaluate the quality of the obtained 3D mesh the quality parameters introduced in [5] can be very useful. Q_{equ} is the index of parametric regularity of a mesh face

$$Q_{equ}(f) = \frac{6}{\sqrt{3}} \frac{A}{s \cdot h} \in [0,1] \tag{1}$$

where A is the area of the face, s the semiperimeter and h the length of the longest edge. Q_{equ} is a value between 0 and 1, where 1 corresponds to a equilater triangle and thus to maximal regularity. Another quality index is Q_{plan}, which refers to the geometrical description of the mesh. Q_{plan} is defined as:

$$Q_{plan} = \frac{n \cdot n_1 + n \cdot n_2 + n \cdot n_3}{3} \in [0,1] \tag{2}$$

where n is the normal to the face and n_1, n_2, n_3 are the normals of the triangles which are adjacent to the three face edges. A good mesh must describe high curvature regions with a high sampling rate. Q_{plan} could be increased by sampling the mesh at high rate, but this would generate huge size models without enhancing the level of detail of low curvature regions. A sampling rate proportional to local curvature is therefore advisable.

Assuming the above 3D reconstruction process provides us with the coordinates of n points lying on a real surface Λ, the points could be expressed as $x_i = y_i + \varepsilon_i$, where y_i are the true values and ε_i are the measurement errors. We also have m views V_j of Λ, which can be considered as functions mapping $\Re^3$ in $\Re^2$ through projective transformations. For each view V_j we know the projection $P_j = V_j(\Lambda)$ of the original surface Λ, i.e. the set of points representing the silhouettes of Λ. The fusion problem of silhouettes and stereo matching information concerns the estimate of Λ from x_i and P_j.

Such a problem can be solved within a classical deformable model framework [4] [11]. Namely, a surface is made to evolve subject to three types of forces, an internal and two external ones. The first one, F_{int}, keeps the surface as smooth as possible, while the others, F_{tex} and F_{sil}, make it to converge to Λ. Formally, the evolution of the model at point P can be described as:

$$s(0) = s_0 \tag{3}$$

$$\frac{\partial s}{\partial t}(t)(P) = F_{int}(P,s) + F_{tex}(P) + F_{sil}(P,s) \tag{4}$$

where $s(t)$ is the estimate of Λ at iteration t, s_0 is the estimate obtained through the shape-from-silhouette method, and

$$F_{int}(P,s) = \nabla^2 s(P) - \nabla^4 s(P) \tag{5}$$

F_{tex} deformes the model in order to minimize its distance from cloud x_i; F_{sil} deformes the model in order to make it consistent with silhouette information i.e., F_{sil} tends to make the model silhouettes as similar as possible to the acquired ones.

In [4] F_{tex} is expressed as the Gradient Vector Flow (GVF) [17], obtained from point cloud x_i with the aim of eliminating the local minima arising when the surface reaches a boundary concavity. This vector field is irrotational and so it can't have a variational origin, but its convergence is proved in [17].

As for F_{sil}, in [11] it is defined as follows:

$$F_{sil}(P) = \sum_{j=1}^{m} f^{V_j}(P) \tag{6}$$

where $f^{V_j}(P)$ is nonzero iff $V_j(P)$ is external to P_j or is internal to P_j and on the boundary of $V_j(s)$ (that is, $V_j(s)$ is the silhouette of s viewed from V_j). In this case $f^{V_j}(P)$ is the back-proiection of the 2D vector joining $V_j(P)$ with the nearest point on $V_j(s)$ boundary. Hence the force is nonzero only along the curves obtained by sectioning s with the retinal planes relative to V_j and passing through P. The force field is therefore strongly discontinuous and can't have variational origin, i.e., it can't be derived from the Eulero-Lagrange equations of a minimum problem. As a consequence, convergence to a model consistent with silhouette information is not guarenteed. Moreover, the force is calculated as the sum of terms separately computed on each silhouette.

In [4] F_{sil} is defined as

$$F_{sil}(P) = \alpha(P) \cdot d_{vh}(P) \cdot n(P) \tag{7}$$

where $n(P)$ is the normal to the surface in P and $d_{vh}(P)$ is the signed distance from the visual hull defined as $d_{vh}(P) = \min_j d(V_j(P), \partial P_j)$, where ∂P_j is the boundary of P_j and $d(V_j(P), \partial P_j)$ is the signed distance between the projection v of P viewed from V_j and ∂P_j, positive when $v \in P_j$ and otherwise negative. $\alpha(P)$ can be expressed as

$$\alpha(P) = \begin{cases} 1 & \iff d_{vh}(P) \leqslant 0 \\ \frac{1}{(1+d(V_c(P),\partial V_c(s)))^k} & \iff d_{vh}(P) > 0 \end{cases} \tag{8}$$

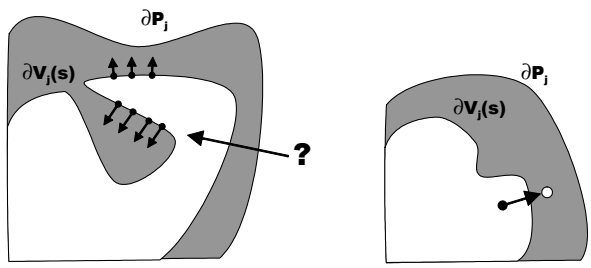

Figure 2: Left: F_{sil} of [4] can sometimes cause trouble; Right: F_{sil} should move the vertex to a point as distant from boundary ∂P_j as the vertex is distant from $\partial V_j(s)$.

where $c = \arg\min_j d(V_j(P), \partial P_j)$ and $\partial V_c(s)$ is the boundary of the silhouette of s viewed from V_c. The force field is continuous, defined over the entire model with the same direction and versus as the surface normal. Namely, the boundary points are subject to a force equal (both in magnitude and versus) to the distance vector from the visual hull. For the points internal to $V_j(s)$ the force intensity is modulated by $\alpha(P)$ so as to be inversally proportional to their distance from the boundary. In this way, the vertices can leave the visual hull and enter the concavity opposing only a force decreasing by $\alpha(P)$ with respect to the distance from the boundary. Eq.(7) avoids the difficulties arising from the sum of Eq.(6), however it is affected by three problems, namely:

1. $F_{sil}(P)$ always acts along direction and verse of the surface normal when $V_j(P) \in P_j$ for every j, even if this means moving P away from the silhouette. See for example the situation depicted in Fig.2Left, where F_{sil} causes the model to intersect itself.

2. F_{sil} is proportional to the distance between $V_j(P)$ and ∂P_j, even if $V_j(P)$ does not lie on the boundary of $V_j(s)$. In fact, in that case we would expect that F_{sil} moved P to a point as distant from boundary ∂P_j as the vertex is distant from $\partial V_j(s)$ (see Fig.2Right).

3. F_{sil} still doesn't have variational origin.

3 Problem reformulation and solution

Neither in [4] nor in [11] the proposed silhouette force has a variational origin. Nevertheless, as in the case of classical deformable models, we would like our model to be the minimum point of a functional ξ defined as:

$$\xi(s) = \int_s k_{int} \cdot \xi_{int} + k_{tex} \cdot d_\Sigma(P) + k_{sil} \cdot \xi_{sil} ds \quad (9)$$

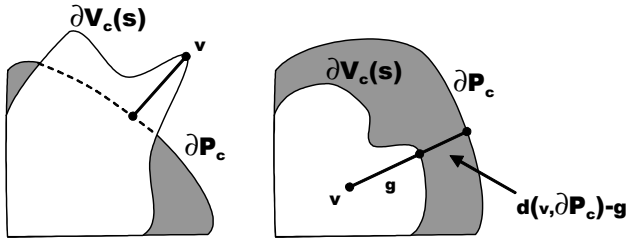

Figure 3: Representation of distances used to define ξ_{sil}.

ξ_{int} is the snake internal energy, which can be expressed as $\xi_{membrane} + \xi_{thin\ plate}$, with

$$\xi_{membrane} = \|\partial s_u\|^2 + \|\partial s_v\|^2 \quad (10)$$

$$\xi_{thin\ plate} = \left\|\partial^2 s_{uu}\right\|^2 + \left\|\partial^2 s_{vv}\right\|^2 + 2\left\|\partial^2 s_{vu}\right\|^2 \quad (11)$$

$d_\Sigma(P)$ is the distance between P and point cloud Σ defined as [12]

$$d_\Sigma(P) = \min\{d(P,x) \mid \forall x \in \Sigma\} \quad (12)$$

Finally we define ξ_{sil} as

$$\xi_{sil} = S_c(V_c(P)) \quad (13)$$

where $V_c(P)$ is the projection of P onto the image plane of view V_c, $c = \arg\min_j d(V_j(P), \partial P_j)$; $d(v, \partial P_j)$ is the signed distance between v and ∂P_j, which is positive if $v \in P_j$ or else negative; $S_c(v)$ is defined as:

$$S_c(v) = \begin{cases} \frac{1}{2} d(v, \partial P_c)^2 & v \notin P_c \\ (d(v, \partial P_c) - g) \cdot h(g) & v \in P_c \end{cases} \quad (14)$$

with g distance between v and the point where $V_c(s)$ intersects the segment joining v and its nearest point on ∂P_c (the one used to compute distance $d(v, \partial P_c)$) (see Fig.3); Function $(d(v, \partial P_c) - g)$ is the distance between $\partial V_c(s)$ and ∂P_c along the above mentioned segment. This value is constant for all the points lying on the extension of such segment to the nearest boundary. Weight $h(g)$ is a nondecreasing function assigning greater weights to the points internal to $V_c(s)$ but external to P_c:

$$h(x) = \begin{cases} -\frac{1}{2p} x^2 + x + 1 & 0 \leqslant x \leqslant p \\ \frac{p}{2} + 1 & x \geqslant p \end{cases} \quad (15)$$

The cost of the points whose projection lies on the boundary of $V_c(s)$ is porportional to its distance from the boundary of P_c: the deeper we get in $V_c(s)$, the greater the point costs, until an upper bound set by $h(x)$ is reached. This procedure penalizes all the surfaces with projections within the projections of Λ. The points of s with projection along V_c external to P_c have a cost which is half their squared distance from boundary ∂P_c. In this way all the surfaces with one projection including the projections of Λ are penalized.

Let's now prove that

Theorem 1 $\xi_{sil}(s) = 0$ *iff* $V_j(s) = P_j,\ \forall j = 1, \dots m.$

Proof. ($\Leftarrow$) If $V_j(s) = P_j,\ \forall j = 1, \dots m$ then $\partial V_j(s) = \partial P_j$, so $(d(v, \partial P_c) - g) = 0$, that is the distance between the boundaries is zero. Moreover, for each $v \in V_c(s)$ we have that $v \in P_c$ and therefore $S_c(v) = (d(v, \partial P_c) - g) \cdot h(g) = 0$; as a consequence, $\xi_{sil}(s) = 0$. ($\Rightarrow$) If $\xi_{sil}(s) = 0$, being the integral of a positive function $S_c(V_c(P))$ has to be null for each $P \in s$. Therefore, $v = V_c(P)$ must belong to P_c, otherwise this would imply $S_c(V_c(P)) \neq 0$ for one P. Being $v \in P_c$, we have $(d(v, \partial P_c) - g) \cdot h(g) = 0$. $h(g)$ is always strictly positive, so $(d(v, \partial P_c) - g)$ must be 0. This means the distance between the boundaries is zero, i.e. $\partial V_c(s) = \partial P_c$. Having chosen $c = \arg\min_j d(V_j(P), \partial P_j)$, we obtain $\partial V_j(s) = \partial P_j\ \forall j = 1, \dots m$. If the boundaries are equal, then their relative surfaces are equal too and therefore $V_j(s) = P_j$, $\forall j = 1, \dots m$. ■

A solution of Eq.(9) can be found by first computing the Eulero-Lagrange equations for $\xi(s)$ and then by solving them through a gradient descent method. Obviously the opposite of the gradient represents the forces deforming the surface. Namely, the internal forces will be calculated as the laplacian and bilaplacian of s. We have preferred the solution proposed in [4] for the texture forces because it solves the problems concerning the local minima.

As for the computation of the silhouette force applied to the vertex v of the mesh, two cases must be distinguished. When $V_c(v) \notin P_c$, the force is a vector applied to v and directed towards the nearest point on ∂P_c with intensity $d(v, \partial P_c)$. As for every other silhouette force, this force lies on the retinal plane passing through v. On the contrary, if $V_c(v) \in P_c$ then

$$\nabla S_c(v) = (d(v, \partial P_c) - g) \cdot h'(g) \tag{16}$$

This is because $\nabla(d(v, \partial P_c) - g) = \nabla d(v, \partial P_c) - \nabla g$ is null as $\nabla d(v, \partial P_c)$ and ∇g are parallel versors. Weight $h'(x)$ can be written as

$$h'(x) = \begin{cases} -\frac{1}{p}x + 1 & 0 \leqslant x \leqslant p \\ 0 & x \geqslant p \end{cases} \tag{17}$$

In this way the force is defined just for the points with distance from the boundary $\partial V_c(s)$ at most equal to p. Moreover, it is directed towards the nearest point of ∂P_c with intensity

$$(d(v, \partial P_c) - g)\left(1 - \frac{1}{p}g\right) \tag{18}$$

proportional to the distance between the two boundaries. The force decreases with the distance from the boundary of $V_c(s)$, until it becomes zero at distance p. The intensity of the points belonging to $\partial V_c(P)$ is instead always equal to $(d(v, \partial P_c) - g)$.

In our approach F_{sil} is not bound to have the same verse and direction as the surface normal, instead it moves as the silhouette would have to do in order to fit the data. This property is guarenteed because the force is generated as a gradient. Moreover, F_{sil} tends to move v to a point as distant from the boundary ∂P_j as the vertex is distant from $\partial V_j(s)$ (see Fig.2Right).

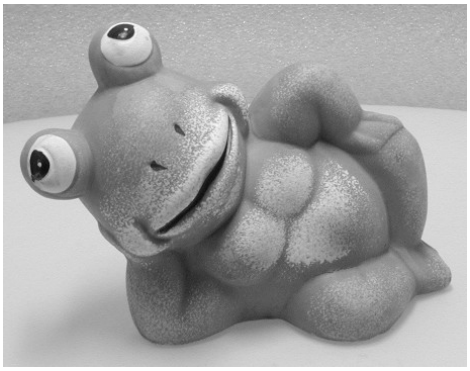

Figure 4: Picture of a frog.

Furthermore it is worth nothing that the computation is simpler than in the case of [4], [11]. This is rather relevant since the calculation of the silhouette force is a bottle-neck during model evolution. In order to speed up the computation of $d(v, \partial P_c)$ we use a kd-tree [1] a priori built on every ∂P_j. In order to calculate $V_j(s)$, s was rendered by means of *OpenGl* libraries [6]. Finally, the computation of g is trivial as it amounts to counting the number of pixels internal to $V_c(s)$ along the segment joining $V_c(v)$ to its nearest point on ∂P_c. Therefore one just needs to draw a segment and to find its intersection with $V_c(s)$.

Let's note that in [4], [11] after every rendering a further distance map needs to be calculated in order to compute the distance $d(v, \partial V_c(s))$, consequently increasing the time complexity of the algorithm.

4 Refinement

The results obtained by the method proposed in section 3 are rather good as concerns parametric and geometric quality, as shown by Fig.5. However, in this formulation the final model is bound to have a parametrization similar to an isometry, i.e., to be uniformly sampled. Namely, this is due to the use of functional $\xi_{membrane}$. However, as mentioned before, a good quality mesh should be sampled proportionally to its local curvature while in deformable models the sampling rate is fixed. In this case some regions could be undersampled and others oversampled, with consequent poor mesh quality or too large model sizes, respectively.

In order to solve this problem we defined a second phase of evolution, to be started after the model reaches adequate

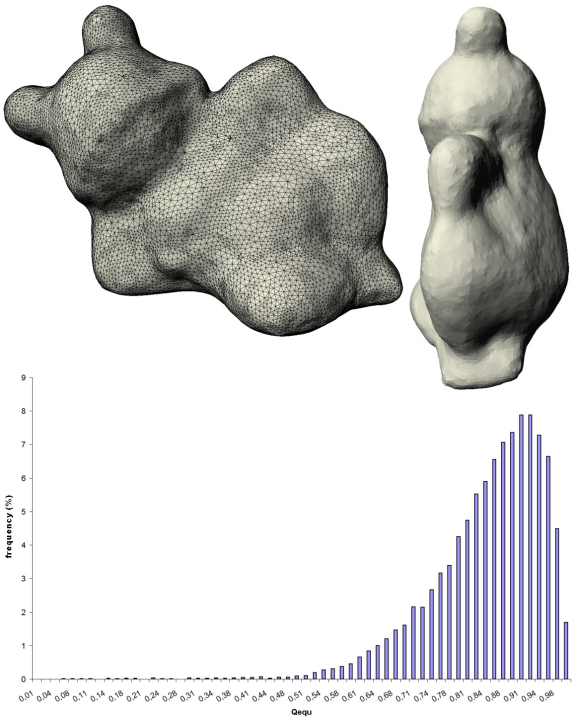

Figure 5: Frog model reconstructed using snakes and relative histogram of Q_{equ}.

convergence to Λ. A selective subdivision of the model based on local curvature is first needed. The model will then evolve so as to minimize a second functional ξ' equal to ξ exept for its internal energy. Intuitively, we could think of setting the latter equal to $\xi_{thin\ plate}$; unfortunately, as shown in Fig.6, this would not lead to a correct description of the curvature of a multiresolution mesh, which is our case. Therefore, we choose to use another functional

$$\xi_{int} = \overline{\kappa} \tag{19}$$

where $\overline{\kappa}$ represents the mean curvature of the surface. As shown in Fig.6, ξ_{int} is independent of the chosen parametrization. As stated in [3], the relative force F_{int} is equal to $-\overline{\kappa}n$, where n is the surface normal; see [3] for a numerical implementation.

In this way only the surface geometrical properties are changed, while the parametrization chosen through subdivision is untouched. We thus obtain a sampling which respects compactness and geometrical quality of the mesh. The resolution enhancement allows us to capture details of the original model in high curvature regions (see Fig.9). In the formulation of [4] this would have been impossible, as it would imply a prohibitive model size.

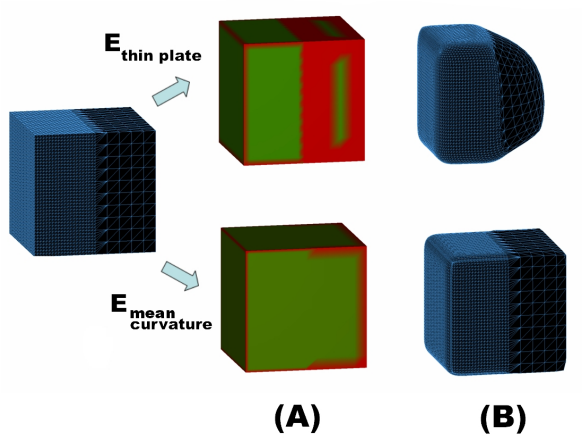

Figure 6: Comparison between $E_{thin\ plate}$ and mean curvature: (A) intensity; (B) effects on the model.

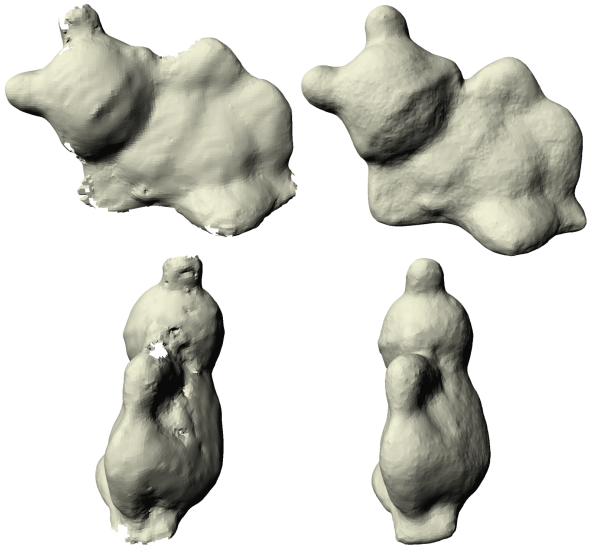

Figure 7: Comparison between the proposed method (Right) and a stereo-based method (Left).

Figure 8: Picture of Buddha.

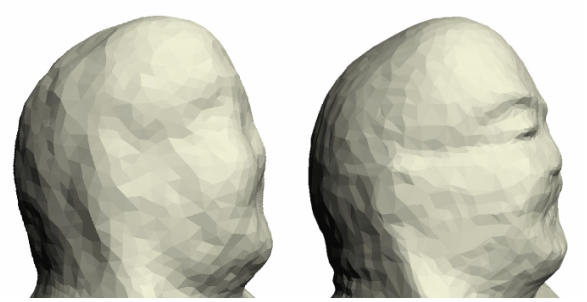

Figure 9: The resolution enhancement allows us to capture details of the original model in high curvature regions. Left: model of Buddha of Fig.8 obtained using classical deformable models. Right: the same model obtained using variable sampling rate.

5 Experimental results

Tests were performed both on real and synthetic models. The acquisition system of the first ones consisted in a rotating table and a fixed camera. The object was positioned on the table while the software controlled table rotation and picture shooting. The pictures of the synthetic models were generated by a rendering software. The model was framed by n 43mm target cameras. The rendering was supervisioned by a script which also calculated the relative projection matrices.

We compared the results obtained by the proposed method with those obtained by a stereo-based methods. We recall that the latter utilizes only information coming from a stereo matching algorithm, building a 3D view from each pair of images. The final mesh is then generated by aligning the 3D views and by averaging the overlapping regions. Differently from stereo-based procedures, our method produces closed surfaces which are manifolds and which can be shown to be both geometrically and parametrically regular (see Fig.5). Parametric regularity can be appreciated by comparing histograms of Q_{equ} frequencies as in Fig.10. Moreover, we observed that silhouette information compensates for the lack of texture information and enhances the level of details in regions where texture is present. This property comes from the fact that intrinsic error of silhouette information is remarkably smaller than texture error. We also found that silhouette information corrects matching errors.

Reconstruction error was evaluated on a synthetic 120x200x260 mm^3 model shown in Fig.11. The model was acquired at $50cm$ distance with a 1024x768 $pixel$ spatial resolution and a field of view of $35°$. Having both the original and the reconstructed model, we finally estimated the average and maximun distance between the two surfaces. In our case we obtained $d_{average} = 0,82\ mm$ ($0,2$ on the diameter) with a $0,62\ mm^2$ variance and $d_{max} = 9,6mm$.

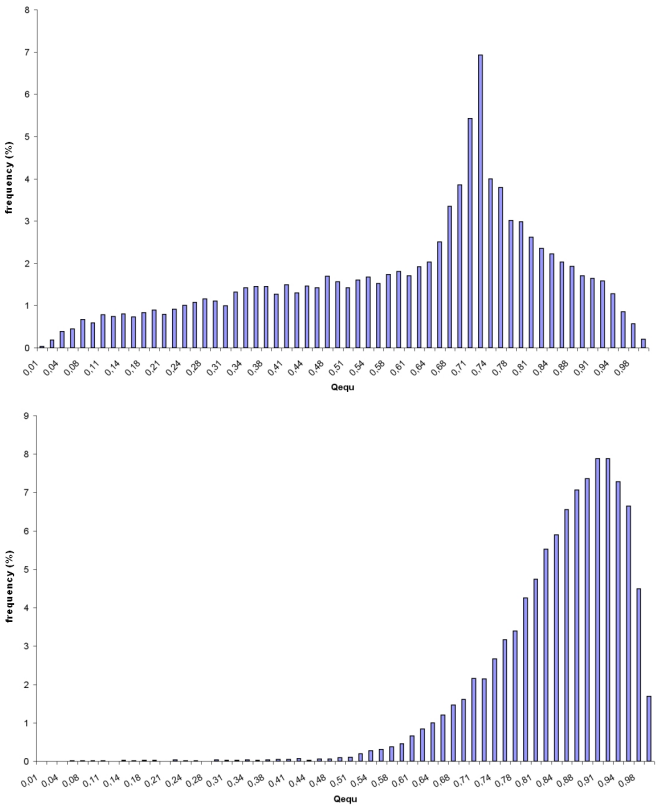

Figure 10: Comparison between histograms of Q_{equ} for the proposed method (Below) and a stereo-based method (Above).

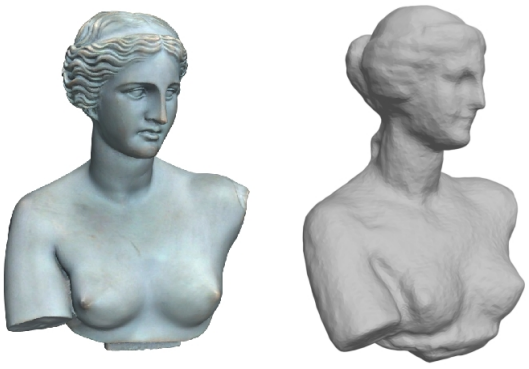

Figure 11: Sinthetic model used to evaluate the reconstruction error of our system.

6 Conclusions

This paper reformulates a 3D passive multimodal digitization scheme using both texture and silhouette information improving upon its formulation. As demonstrated by tests, the proposed system maintains the properties typical of both passive techniques. Indeed, it can be proved to be resilient to measurement errors and capable of reconstructing a wide range of objects such as those featuring:

- Surfaces characterized by good quality texture, sufficient lighting and not too high a specular reflectance (we recall that stereo-based methods completely fail to acquire object even with minimal specular reflectance);
- Specular surfaces without texture or with a periodical texture, provided that the pictures take the profile of such surfaces (information in this case comes from the silhouettes);
- Concavities characterized by good texture and sufficient lighting.

The proposed method still doesn't allow the reconstruction of reflecting or transparent regions, nor the modelling of objects not exhibiting the above mentioned features.

The obtained results prove that the deformable model approach leads to a feasible minimum problem with respect to mesh quality. Indeed, the proposed approach gives a *closed regular manifold* with a *regular parametrization*, unlike stereo-based methods where neither the manifold nor the closure hypothesis generally hold.

Furthermore, the reconstruction error is rather satisfactory. For instance, the surface reconstructed from 1024x768 pictures taken from 50cm distance is affected by an average error of 0,8mm. Such an error can be remarkably reduced using digital cameras of higher resolution.

The main contributions of this paper can be summarized as follows: a reformulation of the silhouette-stereo fusion problem first proposed in [4]; the definition of a new silhouette force of variational origin which, differently from the one defined in [4] and [11], proves to be coherent with the adopted problem formulation and of faster and simpler computability (moreover, the force represents an index of convergence of the model to the original surfaces and it enhances the method robustness); the definition of a second stage of evolution involving only the surface geometrical properties in order to minimize the mean curvature $\overline{\kappa} ds$, which eliminates the fixed sampling bound implied by the use of deformable models and obtains good quality meshes without affecting the model size.

Further research will concern the combination of other passive methods together with silhouettes and stereo. Current work attempts to incorporate in the method the *shadow-carving* [15].

References

[1] A. Atramentov and S. M. LaValle. Efficient nearest neighbor searching for motion planning. *ICRA*, pages 632–637, 2002.

[2] J. Canny. Computational approach to edge detection. *PAMI*, 8:679–698, 1986.

[3] M. Desbrun, M. Meyer, P. Schroder, and A. H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. *International Conference on Computer Graphics and Interactive Techniques*, pages 317–324, 1999.

[4] C. H. Esteban and F. Schmitt. Silhouette and stereo fusion for 3d object modeling. *Computer Vision and Image Understanding*, 96(3):367–392, 2004.

[5] P. J. Frey and H. Borouchaki. Surface mesh quality evaluation. *International journal for numerical methods in engineering*, 45:101–118, 1999.

[6] J. Fuller. *OpenGL Programming Guide, Second Edition*. Addison-Wesley Publishing Company, 1997.

[7] A. Laurentini. The visual hull concept for silhouette based image understanding. *IEEE PAMI*, 16(2):150–162, 1994.

[8] L. Lucchese and S. K. Mitra. Color image segmentation: A state of the art approach. *Proc. of the Indian National Science Academy*, 67(2):207–221, march 2001.

[9] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proc. Royal Society of London*, 204:301328, 1979.

[10] Y. Matsumoto, K. Fujimura, and T. Kitamura. Shape-from-silhouette/stereo and its application to 3-d digitizer. *Proceedings of Discrete Geometry for Computing Imagery*, pages 177–190, 1999.

[11] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara. Real-time 3d shape reconstruction, dynamic 3d mesh deformation, and high fidelity visualiziation for 3d video. *Computer Vision and Image Understanding*, 96(3):393–434, 2004.

[12] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*, volume 153 of *Applied Mathematical Sciences*. Springer, 2003.

[13] M. Potmesil. Generating octree models of 3d objects from their silhouettes in a sequence of images. *Computer Vision, Graphics, and Image Processing*, 40:1–29, 1987.

[14] N. Roma, J. Santos-Victor, and J. Tom. A comparative analysis of cross-correlation matching algorithms using a pyramidal resolution approach. *2nd Workshop on Empirical Evaluation Methods in Computer Vision, Dublin*, 2000.

[15] S. Savarese, H. E. Rushmeier, F. Bernardini, and P. Perona. Shadow carving. *ICCV*, pages 190–197, 2001.

[16] R. Szeliski and D. Scharstein. Symmetric sub-pixel stereo matching. *Seventh European Conference on Computer Vision*, 2:525–540, 2002.

[17] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, pages 359–369, 1998.