# Towards Skeleton based Reconstruction : From Projective Skeletonization to Canal Surface Estimation

Bastien Durix, Géraldine Morin, Sylvie Chambon, Céline Roudet, Lionel Garnier

# Towards Skeleton based Reconstruction: From Projective Skeletonization to Canal Surface Estimation

Bastien Durix
Géraldine Morin
Sylvie Chambon
IRIT, UMR CNRS 5505
Paul Sabatier University, 31000 Toulouse
firstname.lastname@enseeiht.fr

Céline Roudet
Lionel Garnier
LE2I, UMR CNRS 6306
Burgundy University, 21000 Dijon
firstname.lastname@u-bourgogne.fr

## Abstract

*We present a novel approach to reconstruct a 3D object from images corresponding to two different viewpoints: we estimate the skeleton of the object instead of its surface. The originality of the method is to be able to reconstruct a complete tubular 3D object from only two input images. Unlike classical reconstruction methods like multiview stereo, this approach does not rely on interest points but estimates the topology of the object and derives its surface. Our contributions are twofold. First, given two perspective images of the 3D shape, the projection of the skeleton is computed in 2D. Second the 3D skeleton is reconstructed from the two projections using triangulation and matching. A mesh is finally derived for each skeleton branch.*

## 1. Introduction

Nowadays, 3D content is more and more present in media like movies, or video games. Furthermore, many 3D models have to be animated, like characters in a movie. As computers can display more and more polygons, 3D content becomes long and cumbersome to create and to animate; thus automated methods were established to ease the creation of 3D objects. For example, sketching is used to ease the creation of 3D content [14]. In [7], Chen *et al.* elaborated a method based on generalized cylinders, to reconstruct cylindrical objects from a single orthographic image.

Another way to create 3D content is to reproduce objects that physically exist, which is the goal of 3D reconstruction, a field of computer vision. For example, photometric stereo [3] or multiview stereo [11] have been created to estimate geometric information on images. Photometric stereo uses light variation to compute a normal field of the object and then compute the height field from a given viewpoint. However it requires a lot of control over the lighting. Multiview stereo uses several calibrated viewpoints, and matches interest points to triangulate a 3D point cloud. This method is accurate, and outputs a point cloud. Topological information (*i.e.* a mesh) may be recovered (*e.g.* [13]); topology is often needed for editing and animating the 3D content. However it can only reconstruct textured objects (the mouse toy in Fig. 1 can not be reconstructed using multiview stereo). Furthermore, numerous viewpoints are required for recovering a dense point cloud. In particular, pictures around the whole object are required, to get a full model, and not only a face of it.

In this paper, we introduce an approach of 3D reconstruction using curve skeletons, from two calibrated images (*cf.* Fig. 1). We suppose that the shape can be represented by a curve skeleton. The pipeline is the following: first, a binary mask is extracted for each image, and a 2D skeleton is computed on each mask. Then a 3D skeleton is triangulated from the 2D skeletons. The final step computes a mesh associated with the 3D skeleton.

In practice, the proposed reconstruction generates a complete 3D model, with topology, from only two calibrated inputs which is less constrained than for classical reconstruction methods. This method can be used to ease the work of a graphic designer: from a real object, it reconstructs a 3D model that can later be edited, to add details. As skeletons are useful for animation, recovering a skeleton of the generated 3D model can ease animation.

In Section 2, we first present skeletons and their uses. Section 3 presents the projection of a skeleton, first in the orthographic setting, second in the perspective setting. Our first contribution is an algorithm to compute the perspective skeleton from a 2D perspective image. Once the perspective skeletons are computed for both images, Section 4 presents the triangulation of the 3D skeleton, our second contribution. Finally, in Section 5, results are presented and compared to multiview stereo reconstruction.

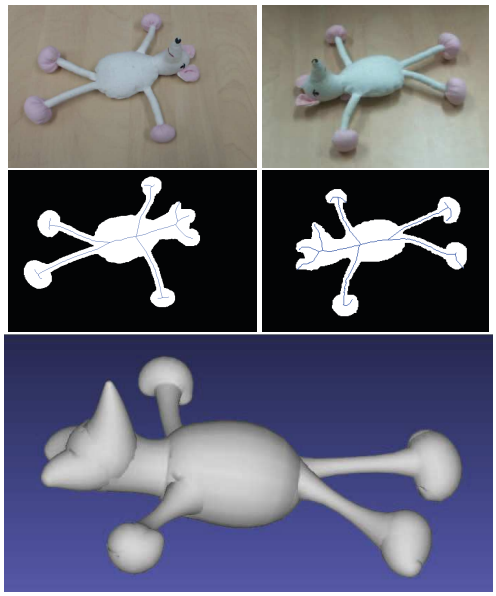

Figure 1. Illustration of the skeleton reconstruction pipeline. First, we take two calibrated acquisitions of an object. Then we extract the perspective skeleton for each image. After a semi-automatic match of the different parts of the skeleton, we can finally recover the complete 3D shape. We can see that tubular shapes (arms, legs, nose, body) are well reconstructed, whereas flat shapes (ears) are approximated by a rounded shape.

## 2. Related work

### Skeleton

Skeletons have been introduced by Blum in [4]. They provide an intuitive model of a shape by a lower dimension shape-centered structure. The **medial axis** is the locus of maximal ball centers. A maximal ball in a given shape is a ball inside the shape, which can not be strictly included in any other ball inside the shape. Maximal balls are also balls inside the shape, which share at least two points with the boundary. Combining each center of the medial axis and its maximal ball radius gives the skeleton of a shape. The skeleton provides enough information to reconstruct the entire shape (*cf.* Fig. 2). Note that what we call the medial axis is indeed the internal medial axis; the external medial axis is defined by maximal balls outside the shape.

The extraction of a skeleton from a shape is called the skeletonization operation. Originally, Blum suggested a grassfire algorithm: the idea is to "burn" the shape from its boundaries and propagate fire at constant speed: the propagation fronts meet on the medial axis. The grassfire algorithm also inspired discrete methods called thinning which consists in successive erosion of the shape *e.g.* [23, 16].

Opposed to iterative thinning methods, one of the most popular skeletonization method is the Voronoï skeletonization [19]. The principle is to sample the boundary of the shape, and to compute the Voronoï diagram of these sample points. The skeleton is defined by the internal Voronoï edges in 2D [9, 2]. This method has been also used in 3D to find the mesh associated to a point cloud [1].

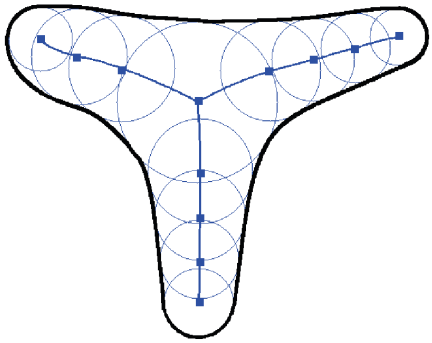

Figure 2. The skeleton of a shape (drawn in blue here) is defined by the set of all maximal ball centers inside the shape, combined to their radii (only a subset is shown here).

The reverse operation of skeletonization is finding a mesh from a skeleton. Various methods exist to solve this problem in 3D, depending on the way to describe the skeleton. Some of them are based on the implicit representation of the shape, from the skeleton [21, 26]. Other methods [10, 15, 8] reconstruct the mesh from a discrete skeleton. Here, we describe skeletons in a continuous way, as canal surfaces.

### Canal Surfaces

We consider continuous curve skeletons. Consequently, we suppose that the manipulated shapes are in a subset of canal surfaces that are defined as the continuous envelopes of a family of spheres, each described by a $\mathcal{C}^1$ center function $C(t)$ and a $\mathcal{C}^1$ radius function $r(t)$ [20]. Each sphere in the family is tangent to the canal surface along a characteristic circle. One can assume that 3D curve-skeletons are canal surfaces without singularities (that can be automatically detected [24]).

Using canal surfaces has an advantage: they can be explicitly characterized from the skeleton. Indeed, for each point $P$ on the surface

$$\min_t \|\overrightarrow{C(t)P}\| = r(t) \ .$$

By squaring then deriving this expression, this implies that a **characteristic circle**, which is the contribution of a sphere to the surface, is the intersection of the sphere with the plane $\Pi(t)$ (*cf.* Fig. 3), defined as

$$\Pi(t) = \left\{ P \in \mathbb{R}^3, \overrightarrow{C(t)P} \cdot \overrightarrow{C'(t)} = r(t)r'(t) \right\} \ , \qquad (1)$$

where $\overrightarrow{C'(t)}$ is the tangent to the curve at $C(t)$.

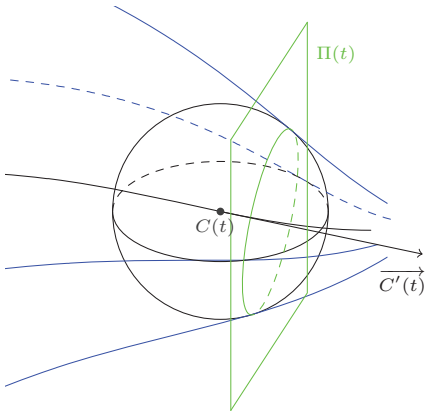


Figure 3. A canal surface is the envelope of a one-dimensional family of spheres: each parameter $t$ defines the sphere of center $C(t)$ and radius $r(t)$. The surface can be computed explicitly: for each $t$, the plane $\Pi(t)$, orthogonal to $\overrightarrow{C'(t)}$, intersects the sphere corresponding to $t$ in the characteristic circle (in green). On that figure, the plane contains the center of the sphere since here $r'(t) = 0$.

Canal surfaces have already been used in the context of 3D reconstruction. Caglioti *et al.* show that it is possible to reconstruct tubular objects [5], or a ball trajectory [6] from a single image (a long exposure photo), when the function $r(t)$ is constant.

In this article, our objective is to reconstruct 3D canal surfaces from two images, with arbitrary radius variation. This means that we only discuss about curve skeletons. Our first contribution is the description of the relation between the canal surface and its perspective projection (Section 3.2.1); we compute the projection of the skeleton from the projection of a canal surface.

## 3. Skeleton projection

We project a 3D canal surface onto the image plane. In the orthographic setting, the projection of the skeleton of a canal surface is, in general, the skeleton of the projection. In the perspective setting, we show that the perspective projection of a skeleton is related to maximal ellipses in the projected shape, and we propose a method to compute these maximal ellipses.

### 3.1. Orthographic projection

Suppose that we orthogonally project a canal surface on a plane, assuming that there is no self-occlusion of the shape from the considered viewpoint. The 2D skeleton which is computed from this 2D shape, may be the projection of the skeleton of the canal surface. We study here under which assumptions this assertion holds.

The skeleton projection is the skeleton of the projected shape if and only if the boundary of each projection of a maximal sphere is a maximal circle in the projected shape. In a 2D shape, a circle is maximal if and only if it shares at least two points with the boundary of the shape (*cf.* Fig. 2). Thus, if each projected circle shares two points with the boundary of the projected shape, the assertion holds.

First, we characterize the points on a maximal sphere, which projection is on the boundary of the projected shape. Then we describe the necessary conditions to find at least two such points, which implies that the projected circle is maximal.

Let $\mathscr{S}$ be a canal surface, modeled by the functions $C(t)$ for the centers and $r(t)$ for the radii (*cf.* Fig. 3). We define each sphere of this canal surface by a function $\Sigma(t)$. Let $\overrightarrow{D}$ be a unit vector, orthogonal to a plane $\Pi_{\overrightarrow{D}}$. Let $\mathscr{S}_{\overrightarrow{D}}$ denote the orthogonal projection of the shape $\mathscr{S}$ on $\Pi_{\overrightarrow{D}}$. We consider the sphere $\Sigma(t)$. Its orthogonal projection (*cf.* Fig. 4) is a disc $\Sigma_{\overrightarrow{D}}(t)$ on the plane $\Pi_{\overrightarrow{D}}$.

Let us consider a point $P$ on $\Sigma(t)$. We show that its projection is on the boundary of $\mathscr{S}_{\overrightarrow{D}}$ if and only if the two following conditions hold (*cf.* Fig. 4):

(i) $P$ is on the boundary of $\mathscr{S}$.

(ii) The projection of $P$ is on the boundary of $\Sigma_{\overrightarrow{D}}(t)$.

Proof of $\Rightarrow$: Suppose that the projection of $P$ is on the boundary of $\mathscr{S}_{\overrightarrow{D}}$. Properties (i) and (ii) are obvious if $P$ was inside the shape or the sphere, its projection would be inside the projection of the shape or the sphere. ■

Proof of $\Leftarrow$: Suppose that (i) and (ii) are true. Property (ii) implies that the normal to $\Sigma(t)$ in $P$ is orthogonal to $\overrightarrow{D}$. As $P$ is on $\Sigma(t)$, so the normal to the sphere at $P$ is the same as the normal to $\mathscr{S}$ at $P$. We assume that there is no self-occlusion, so this implies that the projection of $P$ is on the boundary of $\mathscr{S}_{\overrightarrow{D}}$. ■

As $\mathscr{S}$ is a canal surface, (i) can be reformulated: $P$ has to be on the characteristic circle of the sphere $\Sigma(t)$, so it lies on the intersection of the plane $\Pi(t)$ (*cf.* Equation (1)) and the sphere $\Sigma(t)$ (*cf.* Fig. 3).

Property (ii) can also be reformulated: the projection of a point $P$ on the sphere is on the boundary of the projected circle if and only if $\overrightarrow{PC(t)}$ is orthogonal to $\overrightarrow{D}$. Thus we define the projection circle as the intersection of the sphere $\Sigma(t)$ with the plane $\Pi_{\overrightarrow{D}}(t)$, parallel to $\Pi_{\overrightarrow{D}}$ and passing through $C(t)$.

Thus, the necessary and sufficient condition for a point $P$ on $\Sigma(t)$ to have its projection on the boundary of the projection circle (*cf.* Fig. 4) is

$$P \in \Sigma(t) \cap \Pi(t) \cap \Pi_{\overrightarrow{D}}(t) \,. \tag{2}$$

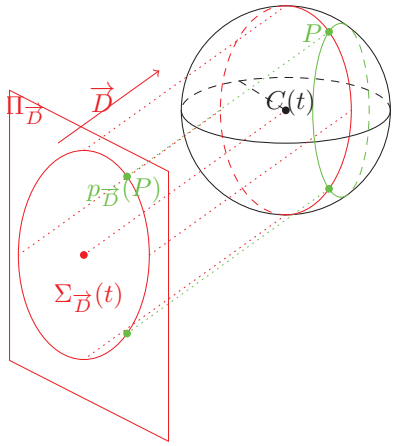


Figure 4. The sphere $\Sigma(t)$ is tangent to the canal surface along the characteristic circle (in green). The projection of $\Sigma(t)$ on $\Pi_{\overrightarrow{D}}$ is a disc. The points on $\Sigma(t)$ projected on its boundary are located on a circle centered on $C(t)$, the projection circle (in red).

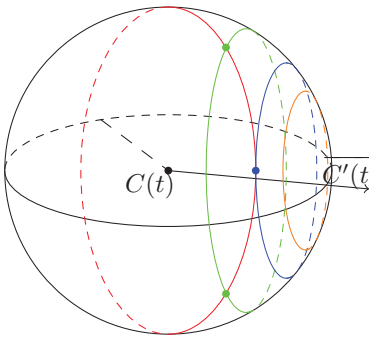


Figure 5. Three possible configurations for the projection circle (in red) and the characteristic circle. The green characteristic circle corresponds to the case shown in Figure 4.The blue one corresponds to a limit case: its intersection with the projection circle is only one point. At most one point of the sphere is on the boundary of the projected shape. The intersection between the orange and the red circles is empty: in this case, the sphere does not contain any point on the boundary of the projection.

We now study in which case there are at least two intersection points. The 3D points which project on the boundary of the projected shape must lie on two circles on $\Sigma(t)$, the characteristic circle ($\Sigma(t) \cap \Pi(t)$) and the projection circle ($\Sigma(t) \cap \Pi_{\overrightarrow{D}}(t)$) (*cf.* Fig. 4). We now determine the conditions for these circles to intersect in two points. Otherwise, two other types of configurations occur (illustrated in blue and orange on Figure 5) for which the projection of $\Sigma(t)$ is not a maximal circle in the projected shape. We now describe these different configurations by a condition on $\overrightarrow{D}$.

If $\overrightarrow{C'(t)}$ and $\overrightarrow{D}$ are not collinear, the intersection of $\Pi(t)$ and $\Pi_{\overrightarrow{D}}(t)$ is a line with direction vector $\overrightarrow{C'(t)} \wedge \overrightarrow{D}$, and passing through the point $N(t)$, where

$$N(t) = C(t) + \lambda_1 \overrightarrow{C'(t)} + \lambda_2 \overrightarrow{D}$$

$$\text{with} \quad \lambda_1 = \frac{r(t)r'(t)}{\|\overrightarrow{C'(t)}\|^2 - (\overrightarrow{C'(t)} \cdot \overrightarrow{D})^2}$$

$$\text{and} \quad \lambda_2 = \frac{-r(t)r'(t)\overrightarrow{C'(t)} \cdot \overrightarrow{D}}{\|\overrightarrow{C'(t)}\|^2 - (\overrightarrow{C'(t)} \cdot \overrightarrow{D})^2} .$$

As $N(t)$ is in the plane defined by $C(t)$, $\overrightarrow{C'(t)}$ and $\overrightarrow{D}$, knowing the relative position of $N(t)$ and $\Sigma(t)$ is sufficient to know the number of intersection points between the line and the sphere: if $N(t)$ is inside/lies on/is outside the sphere, there is respectively two/one/no point(s) in the intersection. By computing $\|\overrightarrow{C(t)N(t)}\|^2$, this condition is equivalent to the following condition on the angle $\alpha(t)$ between $\overrightarrow{C'(t)}$ and $\overrightarrow{D}$

$$\cos^2(\alpha(t)) < 1 - \frac{r'(t)}{\|\overrightarrow{C'(t)}\|^2} \tag{3}$$

implies that there are two intersection points. In case of equality, there is only one intersection point, and if Equation (3) does not hold, there is no intersection point.

Finally, if $\overrightarrow{C'(t)}$ and $\overrightarrow{D}$ are collinear, there are two possible cases. When $r'(t)$ vanishes, the circles do overlap: if $r'(t) = 0$, the plane $\Pi(t)$ contains the center of the sphere (*e.g.* for all $t$ in the case of a cylinder), so it also contains the projection circle. If $r'(t) \neq 0$, there is no common point between the two circles.

Note that the condition "$\Sigma(t) \cap \Pi(t) \cap \Pi_{\overrightarrow{D}}(t)$ contains at least two points" is only local. If a part of the shape is occluded by the shape itself, its projection circle may not be on the boundary – a topological change can even appear in the projected shape. So we need to consider viewpoints avoiding self-occlusions of the shape.

We gave conditions for the projection of the skeleton of a canal surface to be the skeleton of the projection. If these conditions are respected, the projected skeleton is computed using a classical skeletonization algorithm. We use here a Voronoï skeletonization, which provides points on the medial axis but also the skeleton topology.

### 3.2. Perspective projection

For the perspective projection setting, the shape of the projected sphere is not a disc, as a perspective projection is

not an affine map. Indeed, the boundary of the perspective projection of a sphere is an ellipse if the center of projection is outside the sphere (which is always true). In Section 3.2.2, we show how to compute sets of three points to define maximal ellipses. But first, in Section 3.2.1, we show how to define a cone from the three points on the ellipse, then we expose how to recover the center of the sphere from its radius and the cone.

### 3.2.1 Perspective projection of a sphere

Let $\Sigma(t)$ be a sphere, of center $C(t)$ and radius $r(t)$. The central projection of the sphere $\Sigma(t)$, with center $O_3$, origin of the frame, on the plane $z = 1$, is a particular ellipse, intersection between the projection plane and a cone tangent to the sphere. We will construct here a point $C_c(t)$ (*cf.* Fig. 6), such as for each point $Q$ on the ellipse

$$\overrightarrow{O_3Q} \cdot \overrightarrow{O_3C_c(t)} = \|\overrightarrow{O_3Q}\| \,. \tag{4}$$

Let us show such a point exists. We know that the cone axis is $(O_3C(t))$, and its opening angle corresponds to the angle between $\overrightarrow{O_3C(t)}$ and $\overrightarrow{O_3P}$, with $P$ a tangent point between the sphere and the cone. Using the fact that $\|\overrightarrow{C(t)P}\| = r(t)$, a point $A$ lies on the cone if the following equation holds

$$\cos(\overrightarrow{O_3A}, \overrightarrow{O_3C(t)}) = \cos(\overrightarrow{O_3P}, \overrightarrow{O_3C(t)}) \,.$$

Developing this expression gives

$$\frac{\overrightarrow{O_3A}}{\|\overrightarrow{O_3A}\|} \cdot \frac{\overrightarrow{O_3C(t)}}{\|\overrightarrow{O_3C(t)}\|} = \frac{\sqrt{\|\overrightarrow{O_3C(t)}\|^2 - r(t)^2}}{\|\overrightarrow{O_3C(t)}\|} \,.$$

We define $C_c(t)$ such as

$$\overrightarrow{O_3C_c(t)} = \frac{\overrightarrow{O_3C(t)}}{\sqrt{\|\overrightarrow{O_3C(t)}\|^2 - r(t)^2}} \,.$$

In consequence, $A$ is on the cone if and only if

$$\overrightarrow{O_3A} \cdot \overrightarrow{O_3C_c(t)} = \|\overrightarrow{O_3A}\| \,.$$

We can deduce that for each point $Q$ on the ellipse, as $Q$ is also on the cone, so Equation (4) holds. Now let us prove its uniqueness: using three points $\{Q_i, i = 1, 2, 3\}$ on the ellipse, we have a linear system on $\overrightarrow{O_3C_c(t)}$, which is well-posed, as the $\overrightarrow{O_3Q_i}$ vectors are not collinear.

Now, we show that given the radius $r(t)$, the sphere center $C(t)$ can be computed. Let $R$ be the central projection of $Q$ on the unit sphere (*cf.* Fig. 6). According to Thales theorem, we have

$$\frac{\|\overrightarrow{O_3C(t)}\|}{\|\overrightarrow{O_3C_c(t)}\|} = \frac{r(t)}{\|\overrightarrow{RC_c(t)}\|} \,.$$

But

$$\|\overrightarrow{RC_c(t)}\| = \sqrt{\|\overrightarrow{O_3C_c(t)}\|^2 - \|\overrightarrow{O_3R}\|^2}$$

So, $O_3$, $C(t)$ and $C_c(t)$ are collinear, $\|\overrightarrow{O_3R}\| = 1$, and we obtain

$$\overrightarrow{O_3C(t)} = r(t) \frac{\overrightarrow{O_3C_c(t)}}{\sqrt{\|\overrightarrow{O_3C_c(t)}\|^2 - 1}} \,. \tag{5}$$

From the ellipse corresponding to the boundary of the sphere projection, it is possible to recover the center $C(t)$ and the radius $r(t)$ up to a scale factor. Thus, estimating the perspective projection of a skeleton is equivalent to searching three points on maximal ellipses in the 2D shape. We now propose an algorithm to compute these sets of three points corresponding to maximal ellipses.

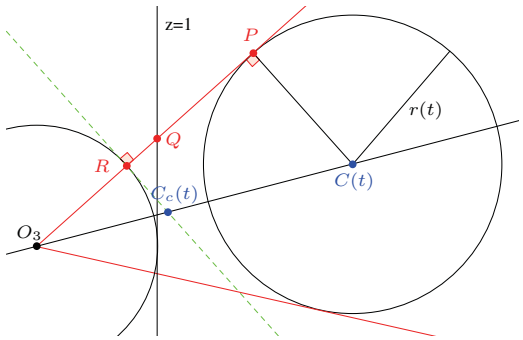


Figure 6. The sphere projection on the plane $z = 1$ (seen from the plane $(yO_3z)$). The boundary of the projection of the sphere $\Sigma(t)$ on the plane is an ellipse. This ellipse corresponds to the intersection between the cone of vertex $O_3$ tangent to the sphere and the plane $z = 1$. That cone is defined by the vector $\overrightarrow{O_3C_c(t)}$, which can be computed by knowing ellipse points.

### 3.2.2 Perspective skeleton estimation

As the perspective projection of a sphere is not a circle, the perspective projection of the skeleton is not the skeleton of the perspective projection. So we seek maximal ellipses (defined above) in the shape rather than maximal circles. This set of ellipses defines what we call here the **perspective skeleton** associated with a perspective projection.

To compute this perspective skeleton, as we are not looking for maximal circles, the same algorithm as for computing the orthographic skeleton is not appropriate. Instead,

we propose Algorithm 1, based on Delaunay tetrahedralization, to estimate sets of three points lying on each maximal ellipse.

**Algorithm 1:** Perspective skeleton estimation

**Data**: 2D discrete shape boundary
$\mathcal{Q} = \left\{Q_i \in \mathbb{R}^3, z_{Q_i} = 1\right\}_{i=1,\cdots,n}$
**Result**: $\mathcal{S}_p$, a set of cones
1 Build the projection of $\mathcal{Q}$ to the unit sphere:
$\mathcal{R} = \left\{ R_i, \exists Q_i \in \mathcal{Q} \text{ s.t. } \overrightarrow{O_3R_i} = \dfrac{\overrightarrow{O_3Q_i}}{\|\overrightarrow{O_3Q_i}\|} \right\}$
2 Compute the 3D Delaunay triangulation of $\mathcal{R} \cup \{O_3\}$
3 Compute a cone for each tetrahedron

The following properties hold for each tetrahedron given by Algorithm 1:

(i) The origin $O_3$ is a vertex of the tetrahedron.

(ii) The vector $\overrightarrow{O_3C_c} = 2\ \overrightarrow{O_3D}$, where $D$ is the center of the sphere, defines a cone with vertex $O_3$ passing through the three remaining tetrahedron vertices.

(iii) The intersection of the cone with the plane $z = 1$ corresponds to a maximal ellipse.

The 3D-Delaunay triangulation generates maximal cones, corresponding to maximal ellipses of the projected shape, from which the perspective skeleton is defined. Moreover, this algorithm derives a topology, as neighbor tetrahedrons lead to adjacent ellipses, whose centers are adjacent points of the skeleton.

To show this, suppose that we apply Algorithm 1 to a set of points $\mathcal{Q}$ on the image plane $z = 1$. The set $\mathcal{R}$ is the projection of the points in $\mathcal{Q}$ on the unit sphere. The second step is Delaunay tetrahedalization of the set $\mathcal{R} \cup \{O_3\}$, that returns a set of tetrahedrons.

Proof of property (i): Let $\mathcal{T}$ be a tetrahedron and suppose that $O_3$ is not a vertex of $\mathcal{T}$. By definition of the 3D-Delaunay triangulation, the sphere $\Sigma$ passing through the vertices of the tetrahedron does not contain any other points in $\mathcal{R}$. Furthermore, all points except $O_3$ are on the unit sphere. But the unit sphere is not a maximal sphere of $\mathcal{R} \cup \{O_3\}$, since the associated ball contains $O_3$. The four vertices are on the intersection of the unit sphere and $\Sigma$, which is a circle. But a circle does not define a unique sphere, which leads to a contradiction. So $O_3$ is one of the vertices of the tetrahedron $\mathcal{T}$. ■

Proof of property (ii): Let us define $C_c$ as $\overrightarrow{O_3C_c} = 2\ \overrightarrow{O_3D}$, thus $D$ is the center of $[O_3C_c]$, so $C_c$ is on the sphere $\Sigma$ of center $D$ and radius $\|\overrightarrow{O_3D}\|$.

Let $R_1$, $R_2$ and $R_3$ be the remaining vertices of the tetrahedron (on the unit sphere), corresponding respectively to $Q_1$, $Q_2$ and $Q_3$ in $\mathcal{Q}$. The triangle $O_3R_iC_c$ is right angle at $R_i$ ($i = 1, 2, 3$). So for each $i$, we have

$$\begin{aligned}\overrightarrow{O_3R_i} \cdot \overrightarrow{O_3C_c} &= \overrightarrow{O_3R_i} \cdot \overrightarrow{O_3R_i} + \overrightarrow{O_3R_i} \cdot \overrightarrow{R_iC_c} \\ &= \|\overrightarrow{O_3R_i}\|^2 = 1\ .\end{aligned}$$

Furthermore, as the sphere passing through $R_1$, $R_2$, $R_3$ and $O_3$ is unique, these four points are not coplanar, so vectors $\overrightarrow{O_3R_i}$ are linearly independent. So the point $C_c$ is the unique point such that $\overrightarrow{O_3R_i} \cdot \overrightarrow{O_3C_c} = \|\overrightarrow{O_3R_i}\|$. Furthermore we have

$$\begin{aligned}\overrightarrow{O_3Q_i} \cdot \overrightarrow{O_3C_c} &= (\|\overrightarrow{O_3Q_i}\|\overrightarrow{O_3R_i}) \cdot \overrightarrow{O_3C_c} \\ &= \|\overrightarrow{O_3Q_i}\|\ .\end{aligned}$$

According to Equation (4), $\overrightarrow{O_3C_c}$ defines a cone of vertex $O_3$ that passes through all the vertices of the tetrahedron, and an ellipse that passes through $Q_1$, $Q_2$ and $Q_3$. ■

Proof of property (iii): Suppose that there exists a point $Q_4$ in $\mathcal{Q}$ in the interior of the cone, and for which the projection on unit sphere is $R_4$, we have

$$\begin{aligned}\overrightarrow{O_3Q_4} \cdot \overrightarrow{O_3C_c} &\geq \|\overrightarrow{O_3Q_4}\| \\ \overrightarrow{O_3R_4} \cdot \overrightarrow{O_3C_c} &\geq 1 \\ \overrightarrow{O_3R_4} \cdot \overrightarrow{O_3D} &> \frac{1}{2}\ .\end{aligned}$$

So, the point $R_4$ is in the ball associated to $\Sigma$ since

$$\begin{aligned}\|\overrightarrow{R_4D}\|^2 &= \|\overrightarrow{R_4O_3}\|^2 + \|\overrightarrow{O_3D}\|^2 - 2\overrightarrow{O_3R_4} \cdot \overrightarrow{O_3D} \\ \|\overrightarrow{R_4D}\|^2 &< 1 + \|\overrightarrow{O_3D}\|^2 - 2\frac{1}{2} \\ \|\overrightarrow{R_4D}\| &< \|\overrightarrow{O_3D}\|\ .\end{aligned}$$

This leads to a contradiction, since the sphere $\Sigma$ does not contain any other point $R_i$, $i \neq 1, 2, 3$. So the proposed algorithm defines a set of maximal ellipses. Similarly, each maximal ellipse corresponds to a tetrahedron. ■

To conclude, the algorithm that we propose gives us a set of maximal ellipses associated to the discretisation of a boundary. This set of maximal ellipses represents here the perspective skeleton that we search. Furthermore, the use of Delaunay tetrahedralization gives us the links between the different cones of the skeleton, which assures that the skeleton topology is associated to the shape topology [19].

## 4. Triangulate the 3D skeleton

We have extracted in 2D the perspective skeleton on the image of a canal surface. Now, we see how to compute

the inverse projection of the skeletons extracted from each image. First, we show that the perspective skeleton on an image corresponds to a two-dimensional space in $\mathbb{R}^4$. We then show how to estimate the intersection between these two spaces in $\mathbb{R}^4$ and so, compute the 3D skeleton defining the shape.

### 4.1. Representation of a perspective skeleton in $\mathbb{R}^4$

Under the assumptions in Section 3.1 (no self-occlusions from the considered viewpoint), we have a discrete skeleton of the shape projection. After a pruning step, we can assume that this skeleton corresponds to a unique branch. We model this branch by a one parameter function, which associates an ellipse to each value of $t$.

Each ellipse from a perspective skeleton is the image of a set of spheres. This set of spheres is a line in $\mathbb{R}^4$, where each sphere of center $(x, y, z)$ and radius $r$ is modeled by a point $(x, y, z, r)$. For perspective projection, the branch skeleton is modeled by the function: $t \mapsto \overrightarrow{O_3C_c(t)}$, representing the axis of a projection cone in the camera frame, for each value of $t$. The camera frame is given by a rotation matrix $\mathcal{R}$ and a translation $\overrightarrow{T} = \overrightarrow{O_3\Omega}$ (where $\Omega$ is the camera position). The set of spheres corresponding to a given skeleton point can be represented by

$$\{S \in \mathbb{R}^4, \exists(t, \lambda), \overrightarrow{O_3S} = \overrightarrow{O_3S_o(t)} + \lambda\overrightarrow{D_S(t)}\} \,. \qquad (6)$$

where $S_o(t)$ is the point of $\mathbb{R}^4$ representing the coordinates of the sphere of center $O_3$ and radius 0 (which corresponds to the sphere at the origin of the cone), and $\overrightarrow{D_S(t)}$ is the vector representing the center and radius variation of the spheres along the cone. Using this notation, each $\lambda$ represents a sphere tangent to the cone at time $t$.

### 4.2. Reconstruction of the 3D skeleton

We model a perspective skeleton by a surface in $\mathbb{R}^4$. Each image generates a perspective skeleton, so we have two $\mathbb{R}^4$ surfaces, associated to the bivariate functions $\overrightarrow{O_3S_1}(t, \lambda) = \overrightarrow{O_3S_{o1}(t)} + \lambda\overrightarrow{D_{S1}(t)}$ and $\overrightarrow{O_3S_2}(t, \lambda) = \overrightarrow{O_3S_{o2}(t)} + \lambda\overrightarrow{D_{S2}(t)}$. The resulting 3D skeleton is, in general, the intersection of these two surfaces.

The 3D skeleton, modeled by a function $\Sigma(t)$, associates a sphere to each parameter value $t$, if there exists two monotonous functions $t_1$ and $t_2$ such that $\forall t, \exists(\lambda_1, \lambda_2)$

$$\overrightarrow{O_3\Sigma(t)} = \overrightarrow{O_3S_1}(t_1(t), \lambda_1) = \overrightarrow{O_3S_2}(t_2(t), \lambda_2) \,.$$

So each sphere of the skeleton is on the intersection of two 3D lines, containing $S_{o1}(t)$ and $S_{o2}(t)$, and with direction vector $\overrightarrow{D_{S1}(t)}$ and $\overrightarrow{D_{S2}(t)}$. These two vectors correspond to two different views of the same sphere, so we can assume that they are not collinear. We estimate two functions $t_1$ and $t_2$ such that the lines $\mathcal{D}(S_{o1}(t_1(t)), \overrightarrow{D_{S1}(t_1(t))})$ and $\mathcal{D}(S_{o2}(t_2(t)), \overrightarrow{D_{S2}(t_2(t))})$ intersect for each value of $t$. So, the distance between the two lines, $d_{12}(t) = \|\overrightarrow{H_1(t_1(t))H_2(t_2(t))}\|$ should cancel, where $H_1(t_1(t))$ is the closest point of $\mathcal{D}(S_{o2}(t_2(t)), \overrightarrow{D_{S2}(t_2(t))})$ on the line $\mathcal{D}(S_{o1}(t_1(t)), \overrightarrow{D_{S1}(t_1(t))})$ and reciprocally. Thus we can show that $\overrightarrow{O_3H_i}(t_i(t)) = \overrightarrow{O_3S_{oi}}(t_i(t)) + \lambda_i\overrightarrow{D_{Si}(t_i(t))}$ and

$$(\lambda_1 \; \lambda_2)^\top = \mathcal{A}^{-1}b$$

where $\mathcal{A}$ is

$$\begin{pmatrix} \|\overrightarrow{D_{S1}(t_1(t))}\|^2 & -\overrightarrow{D_{S1}(t_1(t))} \cdot \overrightarrow{D_{S2}(t_2(t))} \\ -\overrightarrow{D_{S1}(t_1(t))} \cdot \overrightarrow{D_{S2}(t_2(t))} & \|\overrightarrow{D_{S2}(t_2(t))}\|^2 \end{pmatrix}$$

and $b$ is

$$\begin{pmatrix} \overrightarrow{S_{o1}(t_1(t))S_{o2}(t_2(t))} \cdot \overrightarrow{D_{S1}(t_1(t))} \\ \overrightarrow{S_{o2}(t_2(t))S_{o1}(t_1(t))} \cdot \overrightarrow{D_{S2}(t_2(t))} \end{pmatrix} \,.$$

So, the functions $t_1$ and $t_2$ minimize the following criterion

$$\int \|\overrightarrow{H_1(t_1(t))H_2(t_2(t))}\|^2 \, dt \,.$$

The solution of this problem is not always unique. Indeed, in the case of a 3D skeleton with constant radius which is on a plane, containing the projection vectors, each couple of points on the projected skeletons respects the previous criterion. So a criterion for limiting the length of the skeleton is added and we minimize

$$\int \|\overrightarrow{H_1(t_1(t))H_2(t_2(t))}\|^2 + \lambda(t_1'(t)^2 + t_2'(t)^2) \, dt$$

where the second part of the function represents the length of the curve, weighted by a parameter $\lambda$. To minimize it, we discretise the 2D curve skeletons, then use dynamic programming to find the best match between the curves.

## 5. Results

### 5.1. Evaluation of the skeleton based reconstruction on synthetic images

The ground truth is a canal surface representing a single branch curve skeleton. Here, we suppose that we fulfill the optimal acquisition constraints, namely no self-intersection and no self-occlusion of the skeleton from the two considered viewpoints.

Once we have a canal surface and two viewpoints, we generate the images of the canal surface from the two viewpoints. We then extract its boundary using a marching square algorithm (2D version of the marching cubes algorithm [17]). The projective skeleton is then extracted as described in Section 3.2.2, and the 3D skeleton is triangulated. Results are presented in table 1. The quality of the results are quite similar when $\lambda$ belongs to $[0.2, 0.8]$ and here, we present the results for $\lambda = 0.5$.

| $\lambda$ | 0.2 | 0.4 | 0.6 | 0.8 |
|---|---|---|---|---|
| Shape 1 | 2.10% | 2.13% | 2.40% | 2.44% |
| Shape 2 | 2.28% | 2.21% | 2.10% | 2.04% |
| Shape 3 | 0.85% | 1.04% | 0.93% | 1.11% |

Table 1. Quantitative evaluation of the proposed skeleton based reconstruction, on synthetic images. We evaluate the Hausdorff distance between the reconstructed shape and the reference shape. The distance is normalized by the bounding box diagonal, in order to estimate a relative error. Several values of $\lambda$ are used for the evaluation. We see here that relative errors are of the order of some percents, and that the value of $\lambda$ does not affect too much the reconstruction. See the reconstructed shapes at http://durix.perso.enseeiht.fr.

### 5.2. Skeleton reconstruction using two images

Each object is reconstructed from two images taken from two different viewpoints. The objects are segmented from the background with the semi-automatic grabcut algorithm [22], and the perspective skeleton is computed on each image as described in Section 3.2.2. Then the skeleton is simplified, using the scale axis transform [12] (which only needs one constant parameter, representing the pruning level). Branches of each skeleton are paired using a semi-automatic algorithm (we click the end points of the skeleton, and propagate the matches along branches). Finally, we triangulate each branch of the 3D skeleton. A discrete graph skeleton is obtained and each branch is fitted with a B-Spline. This gives a set of canal surfaces, for which we compute an associated mesh (*cf.* Fig. 1).

A qualitative evaluation of our reconstruction is presented in Fig. 7. First, we can see that our reconstruction generates a mesh which gives a complete approximation of the shape. Second, we see that as expected, tubular objects are well reconstructed, since they can be easily approximated by a curve skeleton. On the contrary, flat objects are approximated by rounded shapes (like the blue cuddly cat in the last row of Fig. 7), because we are using a curve skeleton for the body, where the real skeleton is a surface.

### 5.3. Comparison with multiview stereo

As we already mentioned, 3D reconstruction from multiple images is classically done by multiview stereo (MVS). So, for comparison, a reconstruction with the multiview stereo method is computed. For that, we had to take textured objects (objects on Fig. 1 or Fig. 7 lead to very poor results for MVS). Based on a video acquisition of the object to be reconstructed, the external calibration parameters of the camera are estimated by structure-from-motion. We use here the openMVG library [18], given the camera internal parameters. Then the reconstruction is refined using the PMVS library [11]. The result is a 3D point cloud, that we can visually compare with our result. As input to our approach, we chose only two images in the video, and compute the 3D skeleton with the pipeline described above. Figure 8 shows a qualitative comparison between the two methods, because a quantitative evaluation is not possible since we do not have dataset with ground truth.

This figure highlights that the multiview stereo gives more details, in particular on the belly of the teddy, whereas our method is able to provide a full 3D surface model.

Figure 7. Reconstruction of different objects using the proposed method. Left column: one of the two input images. Middle, right columns: front and side views of the reconstruction. Tubular parts like arms are well reconstructed, whereas flat parts like the body of the third object are not, giving a rounded shape to the objects.

## 6. Conclusion

In this article, we presented a novel method to reconstruct objects from images, using skeletons. For this, we use two calibrated views of the object, extract the perspective skeleton in each image, and triangulate the 3D skeleton. This method gives interesting results: a complete 3D model with topology, well adapted for animation and generated from only a couple of images. Unlike multiview stereo, this method is also able to handle untextured scenes.

Figure 8. Comparison between reconstruction by multiview stereo (on the right), and the skeleton reconstruction (on the middle), with one of the origin images (on the left). We consider here textured objects, since multiview stereo can not reconstruct untextured objects. We can see here that both methods are complementary: multiview stereo gives details of the object, and skeleton based reconstruction gives a complete model with topology.

There are some limitations. First, we have to avoid viewpoints of the object leading to self-occlusion. Second, we do not have a global skeleton of the object, but a set of its different parts, which leads to one connected mesh for each skeleton branch. Finally, we can not reconstruct flat surfaces, since we only deal with curve skeletons.

In consequence, the first perspective is to solve the self-occlusion problem, by adding views of the object to retrieve correct topological information, and readjusting the 3D skeleton after its first estimation. Secondly, as the skeleton is represented by a graph, we can topologically associate the graphs estimated from the images to retrieve the 3D graph of the skeleton. Thirdly, to deal with flat surfaces, we can also readjust the skeleton, to match with the projected masks, then compute the surface associated to the obtained surface skeleton [25]. Finally, our method could be combined with multiview stereo, to refine the geometry of the model.

## References


[1] N. Amenta, S. Choi, and R. K. Kolluri. The power crust. In *ACM Symposium on SMA*, 2001. 2

[2] D. Attali and A. Montanvert. Computing and simplifying 2D and 3D continuous skeletons. *CVIU*, 1997. 2

[3] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *IJCV*, 2007. 1

[4] H. Blum. A transformation for extracting new descriptors of shape. *Models for the Perception of Speech and Visual Form*, 1967. 2

[5] V. Caglioti and A. Giusti. Reconstruction of canal surfaces from single images under exact perspective. In *ECCV*. 2006. 3

[6] V. Caglioti and A. Giusti. Recovering ball motion from a single motion-blurred image. *CVIU*, 2009. 3

[7] T. Chen, Z. Zhu, A. Shamir, S.-M. Hu, and D. Cohen-Or. 3-sweep: Extracting editable objects from a single photo. *ACM TOG*, 2013. 1

[8] T. Delamé, C. Roudet, and D. Faudot. From a medial surface to a mesh. In *Symposium on Geometry Processing*, 2012. 2

[9] H. Edelsbrunner. The union of balls and its dual shape. *DCG*, 1995. 2

[10] H. Edelsbrunner. Deformable smooth surface design. *DCG*, 1999. 2

[11] Y. Furukawa and J. Ponce. Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on PAMI*, 2010. 1, 8

[12] J. Giesen, B. Miklos, M. Pauly, and C. Wormser. The scale axis transform. In *ACM SOCG*, 2009. 8

[13] V. Hiep, R. Keriven, P. Labatut, and J.-P. Pons. Towards high-resolution large-scale multi-view stereo. In *CVPR*, 2009. 1

[14] T. Igarashi, S. Matsuoka, and H. Tanaka. Teddy: a sketching interface for 3D freeform design. In *ACM SIGGRAPH 2007 courses*. 1

[15] N. Kruithof and G. Vegter. Meshing skin surfaces with certified topology. In *International Conference on CAD/Graphics*, 2005. 2

[16] L. Lam, S.-W. Lee, and C. Suen. Thinning methodologies - a comprehensive survey. *IEEE Transactions on PAMI*, 1992. 2

[17] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *SIGGRAPH Computer Graphics*, 1987. 8

[18] P. Moulon, P. Monasse, and R. Marlet. OpenMVG. 8

[19] R. Ogniewicz and M. Ilg. Voronoi skeletons : Theory and applications. In *IEEE conference on CVPR*, 1992. 2, 6

[20] M. Peternell and H. Pottmann. Computing rational parametrizations of canal surfaces. *Journal of Symbolic Computation*, 1997. 2

[21] D. Rigaudière, G. Gesquière, and D. Faudot. Shape modelling with skeleton based implicit primitives. In *Graphicon*, 2000. 2

[22] C. Rother, V. Kolmogorov, and A. Blake. "GrabCut": Interactive foreground extraction using iterated graph cuts. *ACM TOG*, 2004. 8

[23] R. Smith. Computer processing of line images : A survey. *Pattern Recognition*, 1987. 2

[24] Z. Xu, R. Feng, and J.-G. Sun. Analytic and algebraic properties of canal surfaces. *Journal of Computational and Applied Mathematics*, 2006. 2

[25] P. Yushkevich, P. T. Fletcher, S. Joshi, A. Thall, and S. M. Pizer. Continuous medial representations for geometric object modeling in 2D and 3D. *Image and Vision Computing*, 2002. 9

[26] C. Zanni, E. Hubert, and M.-P. Cani. Warp-based helical implicit primitives. *Computers & Graphics*, 2011. 2