Fitting of Analytic Surfaces to Noisy Point Clouds

Oscar Ruiz¹, Santiago Arroyave¹, Diego Acosta²

¹Laboratory of CAD/CAM/CAE, Universidad EAFIT, Universidad EAFIT, Medellin, Colombia
²Design and Development of Processes (DDP) Group, Universidad EAFIT, Medellin, Colombia

Email: oruiz@eafit.edu.co

Received February 5, 2013; revised March 9, 2013; accepted March 25, 2013

ABSTRACT

Fitting $C^2$-continuous or superior surfaces to a set $S$ of points sampled on a 2-manifold is central to reverse engineering, computer aided geometric modeling, entertaining, modeling of art heritage, etc. This article addresses the fitting of analytic (ellipsoid, cones, cylinders) surfaces in general position in $\mathbb{R}^d$. Currently, the state of the art presents limitations in 1) automatically finding an initial guess for the analytic surface $F$ sought, and 2) economically estimating the geometric distance between a point of $S$ and the analytic surface $F$. These issues are central in estimating an analytic surface which minimizes its accumulated distances to the point set. In response to this situation, this article presents and tests novel user-independent strategies for addressing aspects 1) and 2) above, for cylinders, cones and ellipsoids. A conjecture for the calculation of the distance point-ellipsoid is also proposed. Our strategies produce good initial guesses for $F$ and fast fitting error estimation for $F$, leading to an agile and robust optimization algorithm. Ongoing work addresses the fitting of free-form parametric surfaces to $S$.

Keywords: Surface Fitting; Optimization; Analytic Surfaces

1. Introduction

Surface reconstruction is a widely studied field because of its importance in CAD-CAM applications, virtual reality, medical imaging and movie industry. Particularly, the reconstruction of analytic surfaces is important since they are frequently used in mechanical parts [1]. Surface reconstruction process consists in obtaining a surface that minimizes the distance between each point $p_i$ of a point sample $S$ and its corresponding point on surface $F$. It is assumed that $S$ fulfills the Nyquist-Shannon criteria [2,3].

1.1. Optimization Approach

The optimization problem of fitting $F$ to $S$ is described by the objective function $f$ shown in Equation (1).

$$f = \sum_{i=1}^{n} d_i^w$$  \hspace{1cm} (1)

where the residual $d_i$ represents the minimum distance between the $i$-th point of $S$ and its corresponding point on $F$ and $w$ indicates the order of the residual. Then $d_i$ is given by:

$$d_i = \min_k \|F - p_i\|$$  \hspace{1cm} (2)

where $k$ is the norm-degree to calculate the distance.

<table>
<thead>
<tr>
<th>Surface</th>
<th>Decision Variables</th>
<th>Type</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cylinder</td>
<td>Axis vector</td>
<td>Vector</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Axis pivot point</td>
<td>Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Radius</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td>Cone</td>
<td>Apex</td>
<td>Vector</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Axis vector</td>
<td>Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Semi angle</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td>Ellipsoid</td>
<td>Center point</td>
<td>Vector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Angle 1</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Angle 2</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Angle 3</td>
<td>Scalar</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Semi axis 1</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Semi axis 2</td>
<td>Scalar</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Semi axis 3</td>
<td>Scalar</td>
<td></td>
</tr>
</tbody>
</table>

To minimize $f$ and find the best surface fit, some variables are tuned. These variables are specific for each situation. Table 1 shows the decision variables for each surface addressed in this paper. On the other hand, norm $k$ remains constant in the optimization process and is considered as a parameter of the problem.
The number of decision variables \( V \) and the number of equality constraints \( E \) in an optimization problem, allow to know the degrees of freedom with the equation \( G = \nabla - \nabla \cdot \). Table 1 presents the degrees of freedom for each analytic surface addressed in this paper. Notice that \( G \) corresponds to \( V \) because the problem is unconstrained.

1.2. Optimization Method

The Gauss-Newton iterative method for solving non-linear optimization problems uses the Hessian approximation \( H = J^* J^\top \) to calculate the next iteration, as is shown in Equation (3).

\[
x_{k+1} = x_k - (J(x)^\top J(x))^{-1} J(x)^\top r_k
\]

where \( x \) is the decision variables vector and \( r \) is residuals vector.

Notice that in the case in which \( f \) is not strictly convex, \( J \) can be singular at some iteration possibly causing the algorithm to diverge. This problem can be overcome by using the Levenberg-Marquardt (LM) Method (4,5):

\[
x_{k+1} = x_k - (J(x)^\top J(x) + \mu I)^{-1} J(x)^\top r_k
\]

where \( \mu \) is the LM parameter and \( I \) is the identity matrix.

1.3. Function and Region Convexity

The convexity condition of an objective function and its feasible region determines if a local extrema corresponds to a global extrema or not. In order to evaluate this condition in \( f \), it is required to examine the eigenvalues \( e \) of its Hessian matrix \( H_f \) by solving Equation (5).

\[
\det[H_f(x') - eI] = 0
\]

If all eigenvalues of \( H_f \) are positive, then \( f \) is strictly-convex, but if at least one eigenvalue is equal to zero, \( f \) is convex [6]. In the case studies on this research, an exact calculation of \( H_f \) is not possible. Thus, a numerical calculation is required by approximating partial derivatives numerically.

2. Literature Review

2.1. Objective Function and Distance Measurement

Some authors have researched the calculation of the distance between a point and an analytic surface. Sappa and Rouhani [7] present a new technique for the estimation a pseudo geometric distance by calculating the height of a small tetrahedron intersecting the surface. This technique is prone to yield accuracy loss in the distance metric when applied to surfaces with high curvatures. Wang and Yu [1] present a comparison of the fitting processes implementing the algebraic, Euclidean, tangent or squared distance for fitting quadric surfaces. Zhou and Salvado [8] compare the geometric and algebraic distances in fitting ellipsoids. The authors estimate the geometric distance as the difference between the length of the ray connecting the point to ellipsoid center and the radius in the intersection of the ray and the surface. This estimation being a fast solution, only works well in cases of quasi-spherical ellipsoids.

2.2. Optimality Conditions

Just like Zhou and Salvado [8], Jiang and Cheng [9] classify the surface fitting problem as a non-convex one. The authors do not discuss the convexity analysis neither for the objective function nor for the optimization region. Other references reviewed do not report any classification of the surface fitting problem in terms of the convexity.

2.3. Optimization Methods


Jiang and Cheng [9] apply a decomposition technique to reduce the dimensions of the optimization space. This implies that the possibilities of dropping into local minima decrease. However, as the approach is out of the geometrical field, coming up with an initial guess of the parameters is not an easy task.

Li and Griffiths [12] fit ellipsoids by a least squares method using quadrics. This method does not require an initial estimation but, as it is not based on real geometrical distances, the results do not provide the best geometric fitting ellipsoid.

References [8,13] report the use of LM method to fit analytic surfaces without mentioning the selection of the LM parameter and its influence on the optimization process.

2.4. Initial Guess

Numerical optimization strategies are sensitive to the initial guess. The closer to the ideal solution is the initial guess, the less number of iterations in the optimization process. On the other hand, a bad initial guess could make the algorithm to diverge.

Just like Ruiz and Cadavid [14], Kwon et al. [15] use PCA for finding an approximation to axis orientation, center coordinates and radius of point clouds belonging to circular cylindrical surfaces. Because this technique reduces the dimensionality of the data giving the direc-
tion of largest dispersion [16], it is limited to cylinders with aspect ratio lower than 5.0 [14] and to cylindrical caps. Similarly, Zhou and Salvado [8] use the eigenvectors of the covariance matrix of the whole data as the axes of the ellipsoid coordinate system. As noted, this technique gives axes that probably will not correspond to the ellipsoid coordinate system in the case of ellipsoidal caps.

Simari and Singh [17] estimate the ellipsoid’s center as the geometric centroid of the data set. This proposal works in the cases of ellipsoids completely sampled but it loses validity when the cloud is only a subsample of the whole ellipsoid.

References [13,18] report the implementation of an algebraic method based on a least squares solution of the general quadric equation to find an initial guess of the parameters of analytic surfaces.

2.5. Literature Review Conclusions and Contribution of This Paper

As was shown in the taxonomy conducted in the literature review, there are several issues that remain open in optimized analytic surfaces fitting which are studied in this work: 1) Estimation of the real geometric distance between a point and an ellipsoid, 2) Identification of the effect of the parameters such as the norm $k$ in the distance measurement, 3) Analysis of the optimality conditions effect on the convergence of the algorithm.

3. Methodology

3.1. Circular Cylinder Fitting

A circular cylinder is defined by a radius $R$, an axis vector and its pivot point $\hat{v}$ respectively. For purposes of this research, no assumption is made on the orientation or position in space of the cylinder from which the data set belongs.

3.1.1. Initial Guess for Circular Cylinder

The initial guess of the cylinder’s parameters is obtained with a statistical and geometrical procedure as explained below and shown in Figure 1:

1) Random seed points are selected from $S$ and a local neighborhood $Ln$ found for each seed. See Figure 1(a).

2) Crossing each other the line segments defined by a seed point and the normal vector $\vec{n}$ of the best plane fit to $Ln$, a set of points $Cp = \{q_1, q_2, \ldots, q_n\}$ passing near $\hat{v}$ are found. See Figure 1(b).

3) A PCA is executed over $Cp$ for finding an initial approximation of the cylinder axis and its pivot point, $\hat{v}'$ and $p'v'$ respectively. See Figure 1(c).

4) An approximation to the cylinder radius is calculated as the average of the minimum distances between $S$ and $\hat{v}$.

This method allows processing both complete cylindrical surfaces and cylindrical caps.

3.1.2. Estimation of Point-Cylinder Distance

The minimum distance $d_i$ between the point $p_i$ and a circular cylindrical surface is calculated as:

$$d_i = |R - \|p_i - pr_i\||$$

where $R$ is the radius of the cylinder and $pr_i$ is the orthogonal projection of $p_i$ onto $\hat{v}_i$ as is shown in Figure 2.

3.2. Circular Cone Fitting

A circular right cone can be defined by an axis vector $\hat{v}$, an apex $Ap$ and a semi-opening angle $\psi$.

3.2.1. Initial Guess for Circular Cone

The initial approximation of a circular conical surface to a point cloud is obtained by an statistical and geometrical procedure as is depicted in Figure 3 and explained as follows:

1) A set of seed points and local neighborhoods $Ln$ are taken from $S$. See Figure 4(a).

2) The minimum curvature direction $\hat{K}_{min}$ of $Ln$,
Figure 3. Procedure for obtaining the initial guess of conical sample.

being collinear with a generatrix of the cone, is found by fitting a paraboloid \( p(x, y) = ax + by + cy + dxy + ex^2 + fy^2 \) and calculating the eigenvectors of it Hessian matrix

\[
H(p) = \begin{bmatrix} 2e & d \\ d & 2f \end{bmatrix}
\]

See Figure 4(b).

3) \( A_p^\ast \) being the first approximation to \( A_p \), is obtained by averaging the crossing points of all lines defined by a seed point and its corresponding \( \hat{K}_{\min} \). Notice that \( A_p' \) represents an statistical apex. See Figure 4(c).

4) By finding the center of gravity of the center of the circumferences passing trough the points \( p_1 = A_p + \lambda \hat{e}_1 \), \( p_2 = A_p + \lambda \hat{e}_2 \) and \( p_3 = A_p + \lambda \hat{e}_3 \), where \( \lambda \in \mathbb{R} \) and \( \hat{e}_1, \hat{e}_2, \hat{e}_3 \) are unitary vectors in direction of minimum curvatures, the initial guess of the axis vector \( \hat{v} \) is calculated. See Figure 4(d).

5) The initial estimation of \( \psi \) is taken as the average of the angles between the vectors \( \hat{K}_{\min} \) and \( \hat{v} \). See Figure 4(e).

3.2.2. Point-Cone Distance Estimation

The distance \( d_i \) from a point \( p_i \) and a cone is calculated by solving the Equation (7) for \( \alpha, \beta \) and \( \gamma \in \mathbb{R} \).

\[
(p_i - A_p) = \begin{bmatrix} \hat{u} \\ \hat{n}_i \\ \hat{n}_z \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}
\]
where $\hat{u}$ is a rotation of $\hat{v}$ around $\hat{n}_1$ an angle $-\psi$.

$$\hat{n}_1 = \frac{\hat{u} \times \hat{v}}{||\hat{u} \times \hat{v}||} \quad \text{and} \quad \hat{n}_2 = \frac{\hat{u} \times \hat{n}_1}{||\hat{u} \times \hat{n}_1||}$$ as is shown in the Figure 5. Finally, $d_i = \beta$ is the signed distance between $p_i$ and the cone.

### 3.3. Ellipsoid Fitting

As is shown in Table 1, an ellipsoid in general position is defined by the center coordinates $(C_x, C_y, C_z)$, the semi axes $R_x, R_y, R_z$ and the Euler angles $\theta_x, \theta_y, \theta_z$.

#### 3.3.1. Initial Guess for Ellipsoid

The first approximation to the parameters of the ellipsoid can be obtained with the following procedure:

1) A general quadric surface is defined by Equation (8).

$$a_{11} x^2 + a_{22} y^2 + a_{33} z^2 + a_{12} xy + a_{13} xz + a_{23} yz = 0$$

Rearranging Equation (8) appropriately [18], it can be written as:

$$b_1 (x^2 + y^2 - 2z^2) + b_2 (x^2 - 2y^2 + z^2) + 4b_3 xy + 2b_4 yz + b_5 z = 0$$

In compact form, the above equation is:

$$AB = R$$

If the $n$ points of $S$ are taken into account, $A$ is a $(9 \times n)$ matrix where its row $i$-th is:

$$x_i^2 + y_i^2 - 2z_i^2, x_i^2 - 2y_i^2 + z_i^2, 4x_i y_i, 2x_i z_i, 2y_i z_i, x_i, y_i, z_i, 1$$

$R$ is a $(n \times 1)$ vector with row $i$-th: $x_i^2 + y_i^2 + z_i^2$.

$B$ is the coefficients vector and it can be obtained by solving the linear system of Equation (9). As the system is overspecified, a least squares solution is calculated with the pseudo-inverse matrix of $A$, $A^+$:

$$B = A^+ R$$

2) The initial approximation of the center coordinates $C_x', C_y', C_z'$, the semi axes $R_x', R_y', R_z'$ and the Euler angles $\theta_x', \theta_y', \theta_z'$, are obtained from the subdiscriminant $A$ in the matrix notation of a general oriented quadric:

$${X^T}AX + VX + d = 0$$

where

$$A = \begin{bmatrix} a_{11} & \frac{1}{2} a_{12} & \frac{1}{2} a_{13} \\ \frac{1}{2} a_{12} & a_{22} & \frac{1}{2} a_{23} \\ \frac{1}{2} a_{13} & \frac{1}{2} a_{23} & a_{33} \end{bmatrix}$$

$$V = (a_1, a_8, a_9)^T$$

$$d = a_{10}$$

$${X} = (x, y, z)^T.$$  

The eigenvectors of $A$ represent the axes of the ellipsoid, then $\theta_x', \theta_y'$ and $\theta_z'$ can be calculated. The eigenvalues of $A$ are proportional to $\frac{1}{(R_x')^2}, \frac{1}{(R_y')^2}$ and $\frac{1}{(R_z')^2}$ [19], then $R_x', R_y'$ and $R_z'$ can be obtained. The initial guess of the ellipsoid center can be obtained as $[C_x', C_y', C_z'] = -A^+V$.

#### 3.3.2. Point-Ellipsoid Distance Estimation

We present the following conjecture (Figure 6).

**Conjecture.** Let $E$ be an ellipsoid centered in $(C_x, C_y, C_z)$ with Euler angles $\theta_x, \theta_y, \theta_z$ and semi axes $R_x, R_y, R_z$. Let $p_i = (x_i, y_i, z_i)$ be a point in $\mathbb{R}^3$. Then, an ellipsoid $E'$ exists with the same center and orientation to $E$, but with semi axes $R_x' = R_x + d_i, R_y' = R_y + d_i, R_z' = R_z + d_i$, that contain $p_i$.

If $E'$ and $p_i$ are translated to the origin and aligned with the principal axes by a rigid transformation, the following equation can be posed:

$$\frac{x^2}{(R_x + d_i)^2} + \frac{y^2}{(R_y + d_i)^2} + \frac{z^2}{(R_z + d_i)^2} = 1$$

Rearranging (11):

$$k_x d_i^6 + k_y d_i^4 + k_z d_i^2 + k_2 d_i^6 + k_3 d_i^4 + k_4 d_i^2 + k_5 d_i + k_6 = 0$$

where $k_j = f(x_i, y_i, z_i, R_x, R_y, R_z)$ with $j = 0, 1, \cdots, 6$.

![Figure 5. Calculation of point-cone distance.](Image)

![Figure 6. Calculation of point-ellipsoid distance.](Image)
The minimum absolute real root of the polynomial in Equation (12) corresponds to the minimum signed distance from \( p_i \) to \( E \).

4. Results and Discussion

In order to test our fitting routines two study cases were proposed as follows.

4.1. Data Set 1. Frog

In order to prove the algorithm for fitting ellipsoids, a subset of the frog shown in Figure 7(a) was taken. In Figure 7(b) the highlighted points to which the ellipsoids were fitted can be seen. The result of the fitting process is presented in Figures 7(c) and (d).

A good initial guess found by an algebraic approach, let to a fast convergence of the algorithm. In Figure 8 it may be seen that the longest fitting process required of 12 iterations for finding the optimum according to the termination criteria. In Figure 9 the initial estimation ellipsoid and the best geometrical ellipsoid fit are shown. Notice that the initial surface wraps most of the points, giving a good starting point for the LM algorithm. Table 2 presents a comparison between the the initial ellipsoids and the optimized ones.

![Figure 7. Ellipsoid fitting results. Original 3D model obtained from Thingiverse® [20].](image)

![Figure 8. Behavior of objective function in the optimization process.](image)

![Figure 9. Comparison between initial guess and optimization results in ellipsoid fitting.](image)

<table>
<thead>
<tr>
<th>Data</th>
<th>Eye</th>
<th>Pupil</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Points</td>
<td>1215</td>
<td>667</td>
</tr>
<tr>
<td>Number of Iterations</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Final Residual</td>
<td>(1.5 \times 10^{-4})</td>
<td>(2.1 \times 10^{-5})</td>
</tr>
<tr>
<td>Radius 1 Variation</td>
<td>67.7%</td>
<td>86.6%</td>
</tr>
<tr>
<td>Radius 2 Variation</td>
<td>47.1%</td>
<td>177.9%</td>
</tr>
<tr>
<td>Radius 3 Variation</td>
<td>152.4%</td>
<td>245.9%</td>
</tr>
<tr>
<td>Angle 1 Variation</td>
<td>2.9%</td>
<td>25.1%</td>
</tr>
<tr>
<td>Angle 2 Variation</td>
<td>11.1%</td>
<td>15.9%</td>
</tr>
<tr>
<td>Angle 3 Variation</td>
<td>78.2%</td>
<td>7.1%</td>
</tr>
<tr>
<td>Center Variation</td>
<td>2.9278</td>
<td>1.5143</td>
</tr>
</tbody>
</table>
4.2. Data Set 2. Fan

To test the cylinder and cone fitting algorithms some parts of the fan shown in Figure 10(a) were processed with the algorithms. Figure 10 displays the results of the optimization process of two conical surfaces and one cylinder. As in the case of ellipsoids, the algorithm found the optimal surface after a few iterations. The history of the optimized function for the Fan Data Set is shown in Figure 11. The cones (Figure 12) required 4 iterations while the cylinders (Figure 13) required 6 iterations. The good initial estimation of the surfaces allows the convergence of the algorithm and to reduce the number of iterations, therefore saving computing resources. Geometrical statistics for the Fan Data Set appear in Table 3.

Figure 11. Behavior of objective function in the optimization process of the fan parts.

Figure 12. Comparison between initial guess and optimization results in cone fitting.

Figure 13. Comparison between initial guess and optimization results in cylinder fitting.

5. Conclusions and Future Work

This article presents the fitting of analytic surfaces (such as cylinders, cones and ellipsoids) in the sense of mathematical optimization. The objective function for each surface was implemented in terms of the real geometric distance. In the case of cylinder and conical surfaces this metric is formulated and calculated easily. However, in the ellipsoid case the measurement of the distance be-
between a point and the surface is not trivial. In response to this situation this work presented a novel methodology to calculate this distance. The addressed results allow to check that the proposed distance calculation works fine.

The routines for the initial guess of the surfaces were implemented using geometrical and statistical procedures. The study cases allow to prove that the iterative optimization algorithms converge fast with a good initial guess.

Future work includes the extension of the optimization strategies to other analytic and to free form parametric surfaces.

REFERENCES


Nomenclature

- $S$: \{p_0, p_1, \ldots, p_n\}$ Noisy point sample
- $F$: Best fit surface to $S$
- \textit{PCA}: Principal Component Analysis
- \textit{LM}: Lenvenberg-Marquardt
- $k$: Norm degree
- $f$: Objective function
- \(d_i\): Minimum distance between the $i$-th point of $S$ and its corresponding point on $F$