Least Squares Ellipsoid Specific Fitting

Qingde Li and John G.Griffiths
Department of Computer Science, University of Hull, Hull, HU6 7RX, UK
q.li@dcs.hull.ac.uk

Abstract

In this paper, a sufficient condition for a quadric surface to be an ellipsoid has been developed and a closed-form solution for ellipsoid fitting is developed based on this constraint, which is a best fit to the given data amongst those ellipsoids whose short radii are at least half of their major radii, in the sense of algebraic distance. A simple search procedure is proposed to pursuit the 'best' ellipsoid when data cannot be well described by this type of ellipsoid. The proposed fitting algorithm is quick, stable and insensitive to small errors in the data.

1. Introduction

Fitting an ellipsoid specifically to a set of 3D scattered points occurs in the areas of pattern recognition, machine vision, 3D graphics and spatial data analysis[1][3] [4][5][10][12] [13][18]. Ellipsoids, though a bit simple in representing 3D shapes in general, are the only bounded and centric quadrics that can provide information of center and orientation of an object. Fitting algebraic surface with scatter 3D points has been discussed widely and some excellent work has been done in [8] [9][14][15] [16] [17]. However, none of these fitting techniques are not ellipsoid specific. In theory, the conditions that ensure a quadratic surface to be an ellipsoid have been well investigated and explicitly stated in analytic geometry and can be found in most analytic geometry textbooks. In fact, when its leading form is positive definite, the solution of a quadratic equation must be bounded and thus represents an ellipsoid [2]. Ellipsoid fitting can be performed directly in several ways by applying some known bounded surface fitting techniques such as those presented in [8] and [17]. However, these techniques involve a highly nonlinear optimization procedure, which often stops at a local minimum and can not guarantee an optimal solution. In this paper, we aim to develop an effective and efficient ellipsoid fitting algorithm and will not consider those fitting methods based on the geometric distance or the parametric representation of an ellipsoid as these will inevitably lead to a nonlinear optimization procedure, which we want to avoid.

Generally, a quadric surface is defined as the locus of points such that their coordinates satisfy the most general equation of the second degree in three variables, namely

$$\begin{aligned} ax^2 + by^2 + cz^2 + 2fyz + 2gxz + 2hxy & \\ +2px + 2qy + 2rz + d &= 0. \end{aligned} \tag{1}$$

Let

$$I = a + b + c \tag{2}$$

$$J = ab + bc + ac - f^2 - g^2 - h^2 \tag{3}$$

$$K = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} \tag{4}$$

then it is known that they are invariants under rotation and translation and equation (1) represents an ellipsoid if $J > 0$, $I \times K > 0$ [7].

It is shown in this paper that when $4J - I^2 > 0$, equation (1) must represent an ellipsoid. On the other hand, for an ellipsoid, when its short radius is at least half of its major radius, then we must have $4J - I^2 > 0$. Based on this condition, a direct fitting method is developed similarly to the work presented in [6]. However, $4J - I^2 > 0$ is just a sufficient condition to guarantee that an equation of second degree in three variables represents an ellipsoid, but it is not necessary. Therefore the ellipsoids that satisfy the condition $4J - I^2 > 0$ are just a subset of the whole ellipsoid family. Note that, for any ellipsoid, there must exist a real number $\alpha \geq 4$ such that $\alpha J - I^2 > 0$. Although $\alpha J - I^2 > 0$ cannot guarantee a quadric surface to be an ellipsoid when $\alpha > 4$, a simple search procedure can be established based on the above fact for fitting an ellipsoid in general. As can be seen from the discussions in the following sections, the proposed method is quick and stable.

Proceedings of the Geometric Modeling and Processing 2004 (GMP'04)
0-7695-2078-2/04 $ 20.00 © 2004 IEEE

2. The Properties of $\alpha J - I^2$

In this section, we list a few properties of the invariant $\alpha J - I^2$ concerning an ellipsoid. The detailed proofs for these properties are omitted due to paper length limitations.

Proposition 1. *$I^2 \geq 3J$, and $I^2 = 3J$ if and only if equation (1) represents a sphere.*

Proposition 2. *Quadric surface (1)must be an ellipsoid if $4J - I^2 > 0$.*

For any ellipsoid,there must exist a positive number α such that $\alpha J - I^2 > 0$ since $J > 0$. From Property 2, it is clear that $\alpha J - I^2 > 0$ is a sufficient condition for a quadric surface to be an ellipsoid when $3 \leq \alpha \leq 4$. However, it is no longer sufficient when $\alpha > 4$. In fact, we have

Proposition 3. *The maximum value of $\alpha > 0$ for which a quadric surface is guaranteed to be an ellipsoid, when $\alpha J - I^2 > 0$, is $\alpha = 4$.*

It can be shown directly that for any $\alpha > 4$, there always exists a quadric such that $\alpha J - I^2 > 0$ but that the quadric is not an ellipsoid.

At the end of the section, we investigate the geometric meaning of the invariant $4J - I^2$ for an ellipsoid. Let A, B, C be the roots of the characteristic equation

$$\mu^3 - I\mu^2 + J\mu - K = 0.$$

Then it is known that A, B, C will be the reciprocals of the major, the medium, and the short radius of the ellipsoid. Define

$$\rho = \frac{4J - I^2}{\{A^2 + B^2 + C^2\}},$$

Then ρ is an invariant under rotation and translation. It can be shown that $|\rho| \leq 1$ and $\rho = 1$ if and only if $A = B = C$ when the equation (1) defines a sphere. Further, It can be observed that when one of the roots tends to infinity or when two of the roots tend to zero, the value of ρ tends to -1. In this case the corresponding ellipsoid will be flat shaped or bar shaped. Thus, we can see that the value of ρ can be used to measure the roundness of an ellipsoid. The bigger the value of ρ, the more nearly spherical the quadric. Conversely, the smaller the ρ is, the flatter or more bar-shaped the ellipsoid is. The following proposition states precisely when an ellipsoids satisfies $4J - I^2 > 0$.

Proposition 4. *For an ellipsoid*

$$Ax^2 + By^2 + Cz^2 + D = 0, \tag{5}$$

if we assume that $A \geq B \geq C > 0$, then

1. *If $B > A/4, C > A/4$, then $4J - I^2 > 0$;*
2. *If $B \leq A/4, C \leq A/4$, then $4J - I^2 \leq 0$;*
3. *If $B \geq A/4, C < A/4$, we can write $C = (\frac{1}{4} - c)A$, $B = (\frac{1}{4} + \alpha c)A$ for some positive number $0 < c < \frac{1}{4}$ and some number $\alpha \geq 0$. Then $4J - I^2 > 0$ if and only if*

$$f(\alpha) = \frac{2(\alpha - 1)}{(\alpha + 1)^2} > c$$

Above property indicates that the types of ellipsoids that cannot be characterized by $4J - I^2 > 0$ are those that are either long-thin or compressed.

3. ellipsoid specific fitting method

In this section, we first discuss how to fit data with a quadratic equation (1) under constraint $kJ > I^2$, where k is a positive number. When $k = 4$, the fitted shape will be an ellipsoid. The fitting technique is quite similar to the work presented in paper[6]. Let $\{p_i(x_i, y_i, z_i)\}_{i=1}^n$ be the set of points to which an ellipsoid is to be fitted. For each point $p_i(x_i, y_i, z_i)$, let

$$\mathbf{X}_i = (x_i^2, y_i^2, z_i^2, 2y_iz_i, 2x_iz_i, 2x_iy_i, 2x_i, 2y_i, 2z_i, 1)^T,$$

For given equation (1), let

$$\mathbf{v} = (a, b, c, f, g, h, p, q, r, d)^T,$$

the least squares fitting problem based on algebraic distance with constraint $kJ - I^2 > 0$ is:

$$min\|D\mathbf{v}\|^2 \qquad \text{subject to} \qquad kJ - I^2 = 1 \tag{6}$$

where D is the design matrix of size $10 \times n$ defined as $D = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n)$. If we define C_1 as a 6×6 matrix in the following way:

$$C_1 = \begin{pmatrix} -1 & \frac{k}{2} - 1 & \frac{k}{2} - 1 & 0 & 0 & 0 \\ \frac{k}{2} - 1 & -1 & \frac{k}{2} - 1 & 0 & 0 & 0 \\ \frac{k}{2} - 1 & \frac{k}{2} - 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -k & 0 & 0 \\ 0 & 0 & 0 & 0 & -k & 0 \\ 0 & 0 & 0 & 0 & 0 & -k \end{pmatrix} \tag{7}$$

and define

$$C = \begin{pmatrix} C_1 & 0_{6\times4} \\ 0_{4\times6} & 0_{4\times4} \end{pmatrix}, \tag{8}$$

$kJ - I^2 = 1$ can be written as $\mathbf{v}^T C\mathbf{v} = 1$ and the constraint minimization problem (6) becomes that of solving a set of equations using the Lagrange multiplier method:

$$\begin{aligned} DD^T\mathbf{v} &= \lambda C\mathbf{v} & (9) \\ \mathbf{v}^T C\mathbf{v} &= 1 & (10) \end{aligned}$$

Note that matrix C has the eigenvalues $\{k-3, -\frac{k}{2}, -\frac{k}{2}, -k, -k, -k, 0, 0, 0, 0\}$. Using the same inference as given in [6], we state that equation (9) has only one solution when $k > 3$, which is the general eigenvector associated with the unique positive eigenvalue of the general eigenvalue system $DD^T\mathbf{v} = \lambda C\mathbf{v}$.

Note that the elements of matrix C are zeros when the corresponding row and column number are larger than 6. If we write

$$DD^T = \begin{pmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{pmatrix} \qquad \mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix} \tag{11}$$

where matrices S_{11}, S_{12}, S_{22} are of size 6×6, 6×4, 4×4 and vector $\mathbf{v}_1$, $\mathbf{v}_2$ are of size 6 and 4. Then the eigen system (9) becomes

$$(S_{11} - \lambda C_1)\mathbf{v}_1 + S_{12}\mathbf{v}_2 = 0 \tag{12}$$
$$S_{12}^T\mathbf{v}_1 + S_{22}\mathbf{v}_2 = 0 \tag{13}$$

When the given data set is not coplanar, S_{22} will be nonsingular and from (13),

$$\mathbf{v}_2 = -S_{22}^{-1}S_{12}^T\mathbf{v}_1.$$

Substituting this equation for $\mathbf{v}_2$ in equation (12), we obtain the following general eigen system

$$(S_{11} - S_{12}S_{22}^{-1}S_{12}^T)\mathbf{v}_1 = \lambda C_1\mathbf{v}_1 \tag{14}$$

which can be further rewritten as an ordinary eigen system

$$C_1^{-1}(S_{11} - S_{12}S_{22}^{-1}S_{12}^T)\mathbf{v}_1 = \lambda\mathbf{v}_1, \tag{15}$$

Since C_1 is nonsingular.

In most cases, matrix $(S_{11} - S_{12}S_{22}^{-1}S_{12}^T)$ will be positive and eigen system (15) will have and only have one positive eigen value. Let $\mathbf{u}_1$ be the eigenvector associated with the only positive eigenvalue of the general eigen system(14), and let $\mathbf{u}_2 = -S_{22}^{-1}S_{12}^T\mathbf{u}_1$, then $\mathbf{u} = (\mathbf{u}_1^T, \mathbf{u}_2^T)^T$ will be the solution to (6). However, the matrix $(S_{11} - S_{12}S_{22}^{-1}S_{12}^T)$ can be singular in some cases. In this situations, the corresponding $\mathbf{u}_1$ can be replaced with the eigenvector associated with the largest eigenvalue.

Remark

1. Direct spherical fitting. From property 1, a sphere can be characterized by $3J = I^2$. When $k = 3$, C_1 will be singular. It can be shown that the reciprocal eigen system of (14)

$$(S_{11} - S_{12}S_{22}^{-1}S_{12}^T)^{-1}C_1\mathbf{v}_1 = \lambda\mathbf{v}_1 \tag{16}$$

will have only one non-negative eigenvalue 0 when $S_{11} - S_{12}S_{22}^{-1}S_{12}^T$ is positive definite, and the vector $\mathbf{v}_1$ corresponding to the nontrivial solution for problem (6) will be the eigenvector of eigen system (16) sociated with eigenvalue 0.

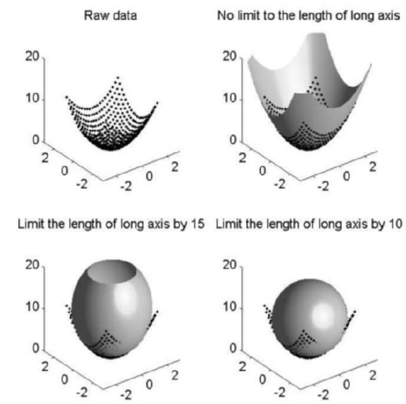

Figure 1. Fitting an ellipsoid with points from an elliptic paraboloid

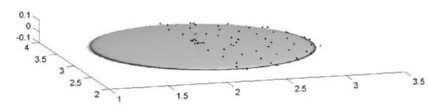

Figure 2. Fitting an ellipsoid with planar points

2. When S_{22} is almost singular, S_{22}^{-1} can be replaced with its generalized inverse $S_{22}^{\dagger}$ and the corresponding solution to $\mathbf{v}_2$ for (13) can be replaced with $-S_{22}^{\dagger}S_{12}^T\mathbf{v}_1$, which has following properties[11]:
   (a) It is the least squares solution of $S_{22}\mathbf{v}_2 = -S_{12}^T\mathbf{v}_1$ when there is no solution.
   (b) It is the unique solution when there is but one solution.
   (c) It is the minimum norm solution when there are an infinite number of solutions.

As already pointed out, $4J - I^2 > 0$ is just a sufficient condition to confirm that a quadric is an ellipsoid, but it is not a necessary condition. Thus, the above ellipsoid fitting by setting $k = 4$ is 'best' just for those ellipsoids satisfying $4J - I^2 > 0$, which is only a subset of the whole ellipsoid family. In practice, when data

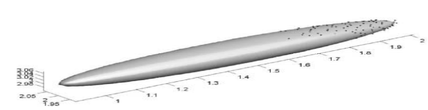

Figure 3. Fitting an ellipsoid to noisy data sampled from a long-thin ellipsoid

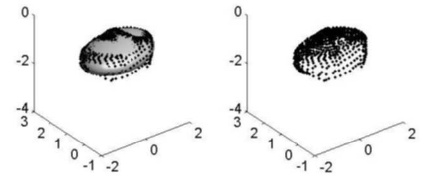

Figure 4. Fitting an ellipsoid with points from the top surface of an actual tibia

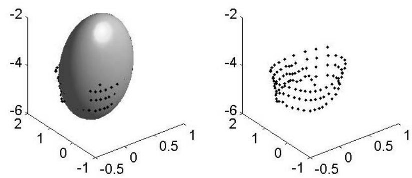

Figure 5. Fitting an ellipsoid with points from the bottom surface of an actual tibia

are sampled from a roughly spherical ellipsoid, the algorithm is just a one step fitting. However, when data are sampled from the surface of a long-thin or a flat object, a large fitting error may be encountered. In this case we need to enlarge the family of ellipsoids by choosing a real number $k > 4$ and use $kJ - I^2 = 1$ to constrain the fitting. The problem is that the fitted equation under the constraint $kJ - I^2 > 0$ may not be an ellipsoid when $k > 4$. Note that when $k' > k$, a quadric satisfying $kJ > I^2$ must also satisfy $k'J > I^2$. Thus, the fitting based on the constraint $k'J - I^2 = 1$ will be preferable to the fitting based on $kJ - I^2 = 1$ if the fitted quadric surface is an ellipsoid, since there are more ellipsoids satisfying condition $k'J > I^2$. Though in theory we do not know exactly the location for the largest number $k_0 (\geq 4)$, such that an ellipsoid is always guaranteed under the constraint $kJ > I^2$ for $3 < k < k_0$,

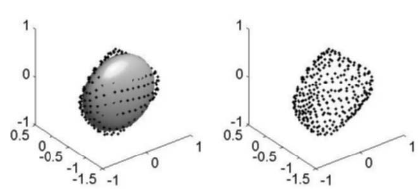

Figure 6. Fitting an ellipsoid with points from the surface of an actual patella

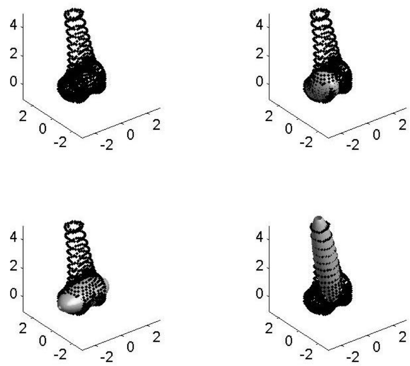

Figure 7. Fitting an ellipsoid to different parts of an actual femur

in practice, a simple search procedure can be easily devised to trace such a number. This can be done first by finding an appropriate range $[a, b]$ for k_0, such that $k_0 \in [a, b]$. Basically, there are two ways to compute it. One way is to begin from $k = 4$ and increase k by a step δ until we find a number b such that the fitting with constraint $bJ > I^2$ will no longer result in an ellipsoid. The other way is to begin with a number b as large as possible and decrease it by a step δ until we get to a number $a (> 3)$ such that the corresponding fitting is an ellipsoids. As in most cases, the fitting based on constraint $kJ > I^2$ results in an ellipsoid even when k is extremely large, it is more efficient in general to start the search from a large number. To quickly locate a upper bound for those plausible k, we can decrease the number k by re-scaling k with a small number, say, let the new k be $k/2$ or $k/10$. Once an upper bound b has been found in this way, we could further compute the least upper bound by a bisection method as described in the following algorithm.

Iterative ellipsoid specific fitting:

1. Set k to a large positive number b.
2. Find the solution for equation:

   $$DD^T\mathbf{v} = \lambda\mathbf{v} \qquad \text{subject to} \qquad \mathbf{v}^T C\mathbf{v} \;=\; 1$$

   where $\mathbf{v}C\mathbf{v} = 1$ corresponding to $kJ - I^2 = 1$.
3. If the fitting is an ellipsoid, STOP; else
   (a) While the fitting is not an ellipsoid and $k \geq 3$, replace k by $k/2$ and fit the data with the constraint $kJ - T^2 = 1$.
   (b) Set $a = max(k, 3), b = 2k$.
   (c) set k=(a+b)/2 and fit the data with the constraint $kJ > I^2$. If the fitted shape is an ellipsoid set a=k; else set $b = k$.

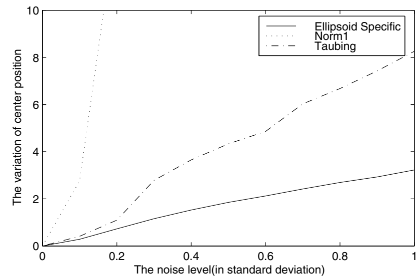

Figure 8. The variation of center position vs. noise level

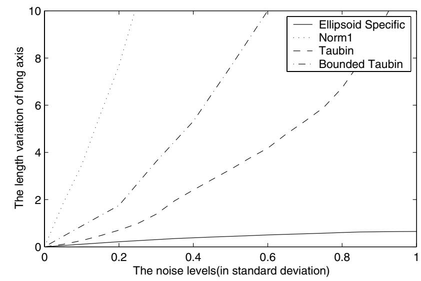

Figure 9. The length error of long axis vs. noise level

(d) If $|a-b|$ is less than a preset tolerance, STOP, otherwise go to 3c.

In this algorithm, we need to check whether a fitted shape is an ellipsoid. Since an ellipsoid can be degenerated into other kinds of elliptic quadrics, such as an elliptic paraboloid, when the lengths of one or two of its principal axes tend to infinity. Therefore a proper constraint must be added. For example, an ellipsoid can be defined as those shapes such that the coefficients A, B, C in its standard form given in (5) are all positive and larger than a preset positive number ϵ.

It can be shown that when the short radius of an ellipsoid is at least $\frac{1}{\sqrt{k}}$ multiple of its major radius, then we will have $kJ - I^2 > 0$. Therefore, the initial value b can be easily estimated from the practical problems.

The number of iterations depends on the data. When the data can be described well by an ellipsoid, it is just a one step fitting. In the worst case for a data set containing about 1000 points, it just takes less than a second to find the optimal solution. For a data set that cannot be properly described by an ellipsoid, we have a trade off between the approximation accuracy and the size of fitted ellipsoid. For example, when the data is precisely from an elliptic paraboloid (see Figure 1), if we want the fitting to be as accurate as possible, the fitted ellipsoid can be very big, which may not what we expect in some cases. To limit the size within an appropriate range, we can discard those ellipsoids whose size is larger than given bounds. Figure 1 shows the fitting results for same data set with different limitations on the length of the long axis. The above algorithm has been tested with data from various kinds of shape such as the hyperboloid, paraboloid, cone, cylinder, tibia, patella, and femur. The fitted results are quite satisfactory.

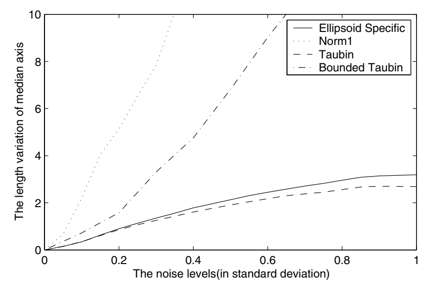

Figure 10. The length error of median axis vs. noise level

4. Experimental results

In this section we show the robustness of our fitting method by comparing it with other least-square fitting approaches. For the sake of simplicity, the fitting algorithm using the constraint that the sum of squares of the estimated coefficients is one is referred to as unit norm method and the method given in [16] is referred as Taubin's method. The bounded fitting method presented in [17] is referred to as the bounded-

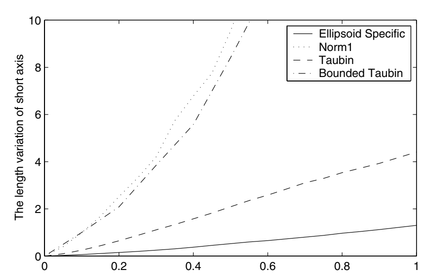

Figure 11. The length error of short axis vs. noise level

Taubin method. In figures 8 to 11, we present the mean error of the center, the lengths of long axis, median axis and short axis. For each level of deviation, these mean errors are computed as the average variation over 500 runs. The data used for obtaining these figures are from ellipsoids centered at the origin with randomly generated lengths for the three principal semi-axis. The sampled data are then corrupted by adding a small perturbation from a Gauss distribution with standard deviation ranging from 0 to 1. These data are then fitted with a quadric surface with four approaches: our ellipsoid specific fitting, unit norm fitting, Taubin's fitting and bounded-Taubin fitting. Since unit norm fitting and Taubin's fitting are not ellipsoid specific, they will not always return an ellipsoid. To compute the average errors for these two methods, we have removed those cases where the fitted shape is not an ellipsoid. As is shown in these figures, our fitting method is very robust compared with the other three fitting techniques. For our new fitting method, the levels of error relating to the levels of noise are much smaller, except for the lengths of median axis, where the error from Taubin's method is slightly smaller. As can be seen from Figures 8 to 11, the fitting based on the bounded-fitting method is quite unsatisfactory when the error level is large. Figures 2 to 7 show that our fitting method always fits an ellipsoid to a set of given data no matter what kind of data are used.

5. Summary

In this paper, we have developed an algorithm for fitting an ellipsoid. We showed that when data are from a surface that is not quite long-thin or compressed, the given method is just a one step direct fitting. When the data are from a very flat and narrow shape, or when the data are from a surface that cannot be well described by an ellipsoid, the data can be fitted with an ellipsoid by a simple searching procedure starting with a large number k. In most cases, this will also be a one-step fitting except for some extreme cases, for example, when the data are precisely from a quadric surface other than ellipsoid. Usually, it just takes less that a second to complete the fitting for data containing about 1000 points. The experiment presented in section 4 shows the method is very stable and robust to noise.

References

[1] P.A. Beardsley. Pose estimation of the human head by modelling with an ellipsoid. *Proc. of IEEE Conf on Automatic Face & Gesture Recognition*, Nara, 1998.

[2] R. Bellman. *Introduction to Matrix Analysis*. SIAM, Philadelphia, 1997.

[3] Douglas DeCarlo and Dimitris Metaxas. Blended deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):443–448, 1996.

[4] Douglas DeCarlo and Dimitris Metaxas. Shape evolution with structural and topological changes using blending. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1186–1205, 1998.

[5] H. Fan. On an earth ellipsoid best-fitted to the earth surface. *Journal of Geodesy*, 72(9):511–515, 1998.

[6] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 21(5):476–480, May 1999.

[7] J. W. Harris and H. Stocker. *Handbook of Mathematics and computational Science*. Springer-Verlag, New York, 1998.

[8] D. Keren, D. Cooper, and J. Subrahmonia. Describing complicated objects by implicit polynomials. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 16(1):38–53, January 1994.

[9] D. Keren and C. Gotsman. Fitting curves and surfaces with constrained implicit polynomials. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 21(1):31–41, January 1999.

[10] A. Langenbucher and B. Seitz et al. Ellipsoidal fitting of corneal topography data after arcuate keratotomies with compression sutures. *Ophthalmic Surgery and Lasers*, 29(9):Clinical Science(abstract), 1998.

[11] W. H. Marlow. *Mathematics for Operation Research*. John wiley & sons, New York, 1978.

[12] M.J.Post. Minimum spanning ellipsoids. *ACM Symposium on Theory of computing*, Washington, D.C.:124–139, 1984.

[13] A. P. Pentland. Perceptual organization and the representation of natural form. *Readings in computer vision, M. Kaufmann publisher, Los Altos, CA*, pages 680–699, Nara, 1998.

[14] V. Pratt. Direct least-squares fitting of algebraic surfaces. *Computer Graphics*, 21(4):145–151, July 1987.

[15] S. Sullivan, L. Sandford, and J. Ponce. Using geometric distance fits for 3-d object modeling and recognition. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 16(12):1183–1195, December 1994.

[16] G. Taubin. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 13(11):1115–1137, November 1991.

[17] G. Taubin, F. Cukierman, S. Sullivan, J. Ponce, and D. J. Kriegman. Parameterized families of polynomials for bounded algebraic curve and surface fitting. *IEEE Transaction on Pattern Analysis and Machine intelligence*, 16(3):287–303, March 1994.

[18] W.C.Karl, G.C.Verghese, and A.S.Willsky. Recontructing ellipsoids from projections. *CVGIP-Graphical models and image processing*, 56(2):124–139, 1994.