

# Automated B-Spline Curve Representation Incorporating MDL and Error-Minimizing Control Point Insertion Strategies

Tat-Jen Cham and Roberto Cipolla

**Abstract**—The main issues of developing an automatic and reliable scheme for spline-fitting are discussed and addressed in this paper, which are not fully covered in previous papers or algorithms. The proposed method incorporates B-spline active contours, the minimum description length (MDL) principle, and a novel control point insertion strategy based on maximizing the Potential for Energy-Reduction Maximization (PERM). A comparison of test results shows that it outperforms one of the better existing methods.

**Index Terms**—B-spline fitting, curve fitting, minimum description length, collapse mechanism, active contour

## 1 INTRODUCTION

REPRESENTING curves by analytic functions instead of sets of data points has many uses in computer vision, but despite the long existence of B-spline curve-fitting techniques, it is difficult to find a *fully automated spline-fitting method* which consistently performs as well as methods based on human assistance.

A number of schemes have been proposed for fitting analytic functions to image curves. Duda and Hart [1] suggested a polygonal representation scheme based on the subdivision of joined line segments, while other methods [2], [3], [4], [5], [6], [7], [8] have been suggested for fitting B-splines. These methods were found to be unsatisfactory for curved-based applications described in [9], [10] primarily because of poor control-point distribution. In some cases, the control points must be initialized manually. Furthermore, most of these methods terminate the spline fitting based on a predefined error threshold.

The main contributions in our paper are:

- 1) isolating and discussing the main issues to spline-fitting, describing where previous methods fail—past papers have not done this;
- 2) introducing a novel control-point insertion strategy which attempts to maximize error-reduction through the iterative insertion of control-points—this addresses the *correct* issue as opposed to the case in previous heuristic schemes;
- 3) combining different techniques into a highly functional and *automated* spline-fitting algorithm.

## 2 MAIN ISSUES IN FITTING B-SPLINES

### 2.1 Parameterization

When handling image curves, the B-spline parameterization along the fitted curve does not need to be unique and the B-spline model may be simplified by having *control points uniformly spaced at unit intervals along the spline parameter* (it does *not* follow that the control

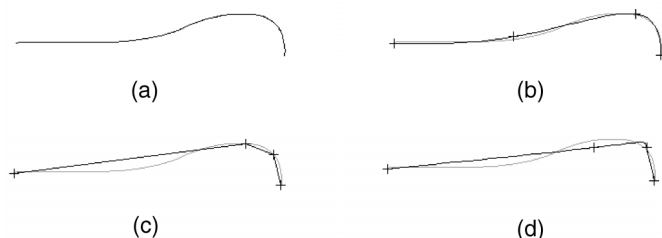


Fig. 1. Nonoptimality of clustering control points around high curvature regions. (a) The target curve. (b) A spline with two control points which is fitted *manually*; the “+”s mark the boundaries between piecewise polynomials. (c) The best polygonal representation used to initialize the control points in (d).

points are uniformly spaced in the image domain). A measure typically used for the quality of a fit is the sum-of-squared-errors (the “error energy”) given by  $E = \sum_i \|\mathbf{x}_i - \mathbf{S}(t_i)\|^2$  where  $\mathbf{x}_i$  are the data samples with associated spline parameter values  $t_i$ , and  $\mathbf{S}(t)$  is the spline with parameter  $t$ .<sup>1</sup> Since  $E$  may be minimized through reparameterization of  $\mathbf{S}$ , such that the  $(\mathbf{x}_i - \mathbf{S}(t_i))$ s are perpendicular to the tangents  $\mathbf{S}'(t_i)$ s of the spline, *the problem may be stated as that of assigning the optimal spline parameter value  $t_i$  to each of the sampled points  $\mathbf{x}_i$* . Bartels et al. [2] proposed the use of normalized arc-length distances between data points for the parameterization, while Guézic and Ayache [4] suggested using arc-length distances along a polygonal representative curve. In both cases, the assigned parameter values are only heuristic estimates. If iterative fitting strategies [4], [5] are further used, the parameter values may be refined numerically but can produce erroneous results if the initial estimates are weak. There are also spline-fitting algorithms which do not take this into account (e.g., [3]), and are likely to assign consecutive data points to a uniformly increasing sequence of parameter values, which gives poor results.

### 2.2 Number of Control Points

The number of control points in a spline model determines the number of degrees of freedom available in the fitting process, and an optimal choice for this is needed to maintain both smoothness and closeness of fit. A number of statistical tests have been proposed including tests for the constancy of an unbiased error variance estimate [6] and Powell’s test for trends [11]. However, in most of the other spline-fitting methods, an empirical error threshold is used.

### 2.3 Distribution of Control Points

Regions which are poorly modeled by polynomials of the given spline order require a greater concentration of control points to reflect the greater information density.<sup>2</sup> Hence, it is *wrong to assume that the optimal distributions of control points are similar between splines of different order, or that control points cluster around regions of higher curvature*. Although the initial positions of the control points may be numerically improved, it is likely that only a local minimum for  $E$  will be found (Jupp’s “lethargy” theorem [12]), and, hence, good initial estimates for the control points have to be obtained. In [4], initial control points estimates are based on Duda and Hart’s polygonal representation. Fig. 1c and Fig. 1d show how the polygonal representation may be used to produce a control-point distribution which will give rise to control points clustering

1. For clarity, we call the points  $\mathbf{S}(t)$  on the spline the *sampling points*.

2. This is analogous to the *Shannon-Nyquist sampling theorem*, applied locally, except that the units are not spatial frequencies but information density.

• T.-J. Cham is with Cambridge Research Laboratory, Compaq Computer Corporation, One Kendall Square, Bldg. 700, Cambridge, MA 02139. E-mail: tjc@crl.dec.com.

• R. Cipolla is with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK. E-mail: cipolla@eng.cam.ac.uk.

Manuscript received 29 Apr. 1997; revised 19 Oct. 1998. Recommended for acceptance by D.J. Kriegman.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 108079.

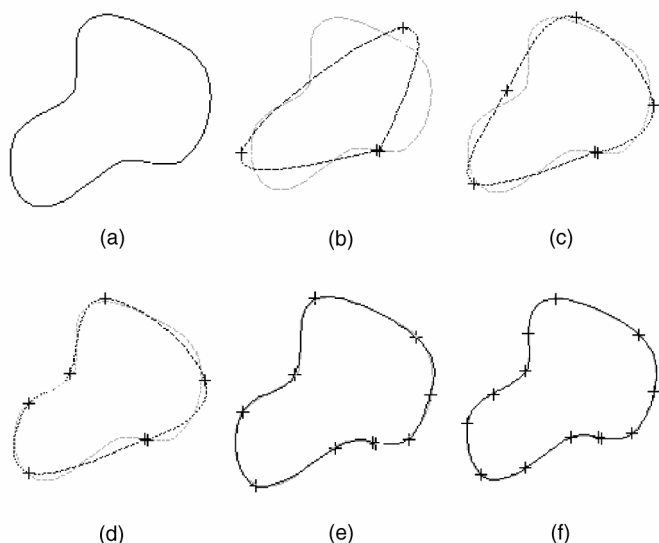


Fig. 2. (a) The target shape, currently treated as an *open* contour. (b)-(f) The spline with increasing number of control points. Of particular interest is that as (c) evolves to (d), a collapse mechanism is set up.

around higher curvature regions. This is clearly weaker than what is possible as shown in Fig. 1b. Others [5] assume that some initial distribution of control points is available or may be obtained heuristically.

### 2.3 Data Sampling

Regions of curves which require a higher proportion of control points to model should have the same higher proportion of sampled points in the vicinity in order to achieve the same rate of oversampling. Many algorithms however treat all available data points equally or weigh them according to their accuracies [4], [6]. Doing this results in simple regions of the curve (e.g., long straight sections) being modeled with too much precision, *at the cost of* poorly representing the complex (high information) regions of the curve. Lu and Milios [5] used curvature-dependent weighing of data points which does not fully address the problem.

## 3 THEORETICAL APPROACH TO B-SPLINE FITTING

The top-level paradigm for fitting splines proposed here is based on the continual *evolution* of a spline such that the spline deforms to approximate the curve to be represented (see Fig. 2).

The development of spline-fitting algorithms has converged on a similar approach; the differences between them lie in the formulation of the various mechanisms, and how the different issues previously listed in Section 2 are handled. The features of our paradigm are:

- 1) **B-Spline Active Contours.** The optimal positions of a fixed set of control points are recovered by treating the B-spline as an active contour.
- 2) **Optimal Control Point Insertion.** A control point insertion strategy is derived which maximizes energy-reduction potential.
- 3) **Minimum Description Length (MDL) Framework.** The minimum description length theory [13] caters for a parameter-free estimation of the optimal number of control points.

### 3.1 Least-Squares Regression

In a least-squares formulation for spline-fitting, it is often impossible to analytically assign the optimal parameter values to the data points. A numerical solution may be achieved using *B-spline active contours*.

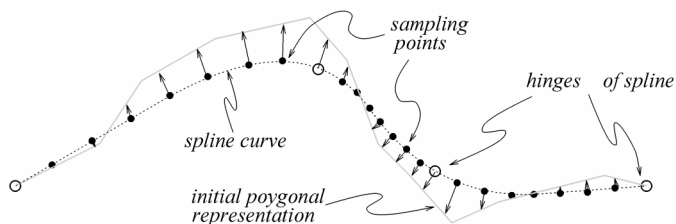


Fig. 3. The number of sampling points per piecewise polynomial, i.e., the oversampling ratio, is constant. Also, all data points associated with the sampling points is found by perpendicular projection.

#### 3.1.1 B-Spline Active Contours

Instead of using point-based models for active contours which involves minimization over a high-dimensional state space, a B-spline active contour has sparse control points but a large number of sampling points. It addresses two fitting issues:

**Data Sampling.** The data is resampled according to some preferred distribution of sampling points on the active contour. This distribution of sampling points may be changed for each iteration. In the proposed model, the number of sampling points per piecewise polynomial is fixed, and is given by  $R = N/(M - 1)$  where  $R$  is the *oversampling ratio*,  $N$  is the number of sampling points, and  $M$  is the number of control points. Hence if two control points are drawn closer, the sampling also becomes proportionally denser in that region. This ratio may be achieved by sampling at fixed intervals of the spline parameter, as shown in Fig. 3.

**Parameterization.** Given a distribution of sampling points, data samples are obtained by searching *perpendicularly* to the spline tangent at each sampling point. These data samples therefore are associated with the spline parameter values at the corresponding sampling points.

#### 3.2 Control Point Insertion

Iterative fitting schemes which involve gradually increasing the number of control points during the fitting process [6], [4], [5], usually require a control point insertion strategy. In this paper, the insertion of control points is considered in its dual form of inserting *hinges*, which are the boundaries between polynomial pieces.

The shortcomings of two current strategies are discussed here:

- 1) **Interval Midpoint Strategy.** Dierckx [6] proposed that a control point be inserted in the middle of the piecewise-polynomial which has large displacements from the data points. However, if these displacements are skewed to one side of the piecewise-polynomial, the optimization routine may be trapped in a weak local minimum.
- 2) **Largest Displacement Strategy.** Lu and Milios [5] suggested that the control point be inserted such that a hinge is formed at the point on the spline which has the maximum displacement from a data point. However, it is possible that this maximum displacement may correspond to a position on or close to a hinge. Introducing a hinge at the position gives very unstable results.

Observation of the failure of the largest displacement strategy reveals that control points must be introduced such that *compatible collapse mechanisms* [14] must be formed. While the original concept of compatible collapse mechanisms was formulated for the plastic deformation of structural beams, in this context, they refer to the family of control point configurations where the error energy can be reduced dramatically (hence, "collapse"). More details may be found in [15].

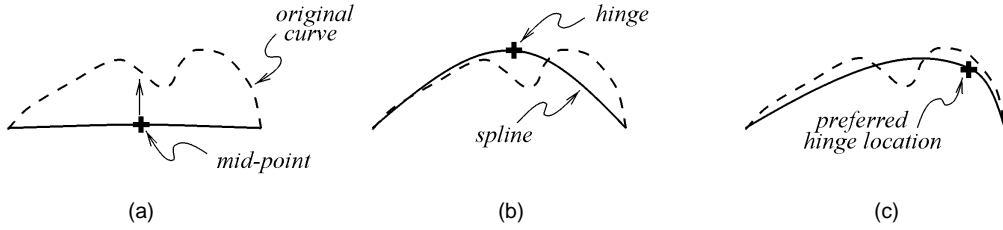


Fig. 4. *Interval Midpoint Strategy*. (a) A hinge is inserted in the midpoint of a piecewise-polynomial interval. (b) Active contour optimization arrives. (c) A preferred solution is shown.

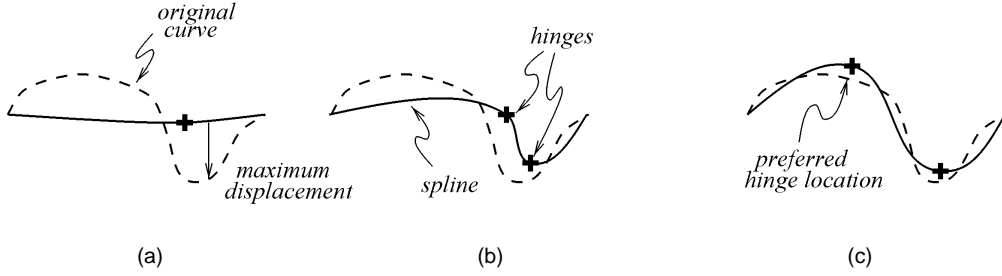


Fig. 5. *Largest Displacement Strategy*. (a) The initial hinge has little effect since a *collapse mechanism* has not been set up. If another hinge is inserted at the maximum displacement, (b) is obtained after optimization. (c) A preferred solution with a collapse mechanism is shown.

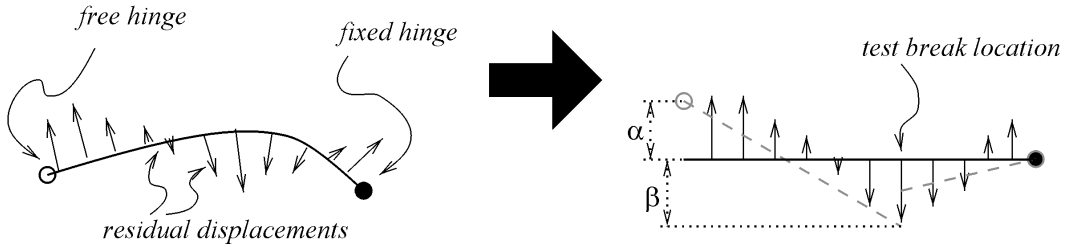


Fig. 6. Figure illustrating the PERM strategy. See text for details.

### 3.2.1 The Potential for Energy-Reduction Maximization (PERM) Strategy

In this strategy, we attempt to estimate the position on the spline to introduce a hinge such that the *potential for reducing the error energy E is maximized*.

Consider a polynomial piece of the spline. By mapping the polynomial piece onto a horizontal line segment, the residual displacements between the sampling points and the data points are also mapped such that they become perpendicular to the line segment. The new vertical displacements  $d_i$  may be written as

$$d_i = \left| \mathbf{x}_i - \mathbf{S}(t_i) \frac{\mathbf{S}'(t_i)}{\|\mathbf{S}'(t_i)\|} \right|. \quad (1)$$

See Fig. 6. The hinges on either side of the polynomial piece are also noted as either fixed (for end-points) or free.

By allowing the line segment to be completely broken at some point, we proceed to calculate the maximum error energy reduction possible on both the smaller line segments which can be achieved by moving the break point and free hinges vertically, as in Fig. 6. For a given break point, the error-energies are quadratic. If the error energy to the left of the break point at  $t_i$  is  $E_l$ , then

$$E_{li}(\alpha, \beta) = \sum_{j=m}^i \left( \alpha + j \frac{\beta - \alpha}{i - m} - d_j \right)^2, \quad (2)$$

where  $\alpha$  and  $\beta$  are the (vertical) displacements of the left hinge and break point, respectively. The maximum energy reduction possible may be easily computed by

$$\Delta_m E_{li} = \Delta E_{li}(\hat{\alpha}, \hat{\beta}) = \frac{1}{2} (\nabla E_{li})^T \mathbf{H}_i^{-1} \nabla E_{li} \Big|_{\alpha=\beta=0}, \quad (3)$$

where  $(\nabla E)$  is the gradient of  $E$ , and  $\mathbf{H}_i$  is the Hessian which is constant with respect to  $\alpha$  and  $\beta$ .

The *Energy-Reduction Potential* for a particular point is defined as the sum of the left and right maximum energy reductions if a break is introduced at that point

$$\mathcal{P}_i = \Delta_m E_{li} + \Delta_m E_{ri}. \quad (4)$$

Intuitively, *this potential measures the energy-reduction potential for a hinge at the point to be part of a compatible collapse mechanism likely to be formed after an additional number of control point (hinge) insertions*.

The PERM strategy is therefore to insert a hinge at the point on the spline corresponding to the maximum energy-reduction potential  $\mathcal{P}$ . As the computational costs of computing (4) is minimal compared to the active contour optimization, the maximum of (4) is found by exhaustive search along the entire B-spline contour, after the active contour optimization has converged. Results based on the PERM strategy have shown that the control points are introduced in near-optimal regions, and compatible collapse mechanisms are formed.

### 3.3 Minimum Description Length (MDL)

The minimum description length (MDL) criterion by Rissanen [13] provides a generic method for comparing the optimality of different models fitted to a particular data set.

Model-fitting may be expressed in a Bayesian framework as *maximizing* the a posteriori probability  $P(\theta|\mathbf{X})$ , where  $\mathbf{X}$  is the

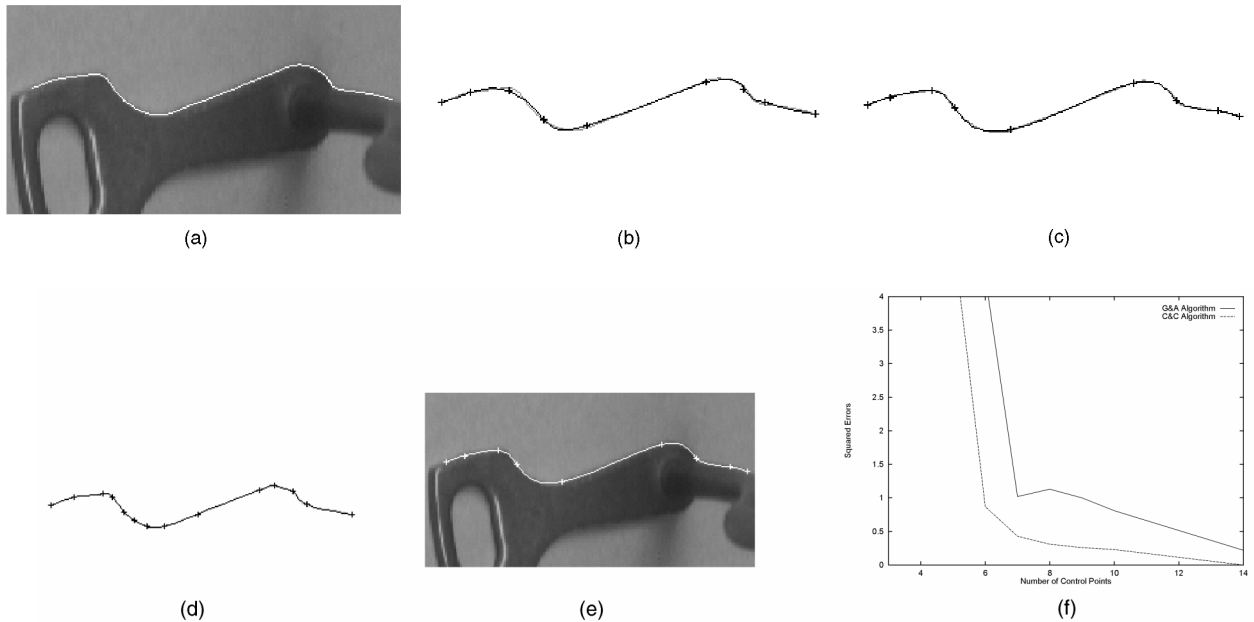


Fig. 7. (a) The target contour. (b) Spline with nine control points using Guéziec and Ayache's algorithm (G&A). (c) Spline with nine control points using our algorithm (C&C). (d) Spline from G&A which has similar averaged errors to the spline in (c). (e) Spline from (c) superimposed on the original image. (f) A graph comparing squared errors when using different numbers of control points with the two algorithms.

matrix of data points given by  $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_N]$ , and  $\boldsymbol{\theta}$  is a variable-length vector containing all model parameters. Then, with the application of Bayes' rule and the assumption that  $P(\mathbf{X}|\boldsymbol{\theta})$  is normal, we obtain an expression to be minimized:

$$L(\mathbf{X}, \boldsymbol{\theta}) = \frac{\log e}{2} (\mathbf{y} - \bar{\mathbf{y}}_{\boldsymbol{\theta}})^T \mathbf{K}_{\boldsymbol{\theta}}^{-1} (\mathbf{y} - \bar{\mathbf{y}}_{\boldsymbol{\theta}}) + M \log N, \quad (5)$$

where  $L(\mathbf{X}, \boldsymbol{\theta})$  is the ideal code length to describe both  $\mathbf{X}$  and  $\boldsymbol{\theta}$ ,  $\mathbf{y} = [\mathbf{x}_1^T \ \mathbf{x}_2^T \ \dots \ \mathbf{x}_N^T]$  is a vector containing the elements of  $\mathbf{X}$ , and  $\mathbf{K}$  is the covariance matrix of  $\mathbf{y}$ .

In circumstances when the covariance matrix  $\mathbf{K}$  is unknown, it is possible to estimate this from the residual errors. If  $\mathbf{K} = \sigma^2 \mathbf{I}$  may be assumed, an unbiased estimate for the uniform variance  $\sigma^2$  is  $\hat{\sigma}^2 = \mathbf{y}^T \mathbf{y} / (N - M - 1)$ . As new control points are inserted, each new estimate  $L_{j+1}$  is calculated based on  $\hat{\sigma}_j^2$  for the current estimate. If  $L_{j+1} < L_j$ , previous estimates  $L_j$  may be revised by back-propagating the new  $\hat{\sigma}_{j+1}^2$ . This scheme is similar to that used in [16]. The minimum  $L$  is usually considered found once  $L$  remains higher than the lowest estimate over a number of iterations (three in our implementation).

#### 4 IMPLEMENTATION AND EXPERIMENTAL RESULTS

The specific problem considered here is that of fitting B-splines to chains of edgels obtained from any edge-detection and linking algorithm. The implemented algorithm will fit cubic B-splines to these chains of edgels.<sup>3</sup>

For the purpose of comparing the performance of our PERM-based algorithm to the latest spline-fitting methods, we implemented Guéziec and Ayache's algorithm [4] which, in our opinion, addresses the main spline-fitting issues most thoroughly among the other published methods. The algorithms were tested on the

contours shown in Fig. 7a, Fig. 8a, and Fig. 8e. The results are compared in terms of errors for the same number of control points, as shown in Fig. 7f, Fig. 8d, and Fig. 8h. For the first curve, the fit obtained by Guéziec and Ayache's algorithm when errors are similar to our fit for nine control points is also shown in Fig. 7d.

The results indicate that the PERM control-point insertion strategy is superior compared to using Duda and Hart's polygonal representation scheme, since our algorithm performs better for any number of control points. However as would be expected, this advantage reduces when linear splines are used. Additionally, Guéziec and Ayache's algorithm requires less computation time. The evolution of our active contours take considerably longer since at each time-step the data needs to be resampled, and convergence is slow in some situations for the gradient-descent scheme used in B-spline active contour optimization.

#### 5 CONCLUSIONS

We have shown that by explicitly describing the main issues in B-spline fitting, a model designed to resolve these issues may be developed. B-spline active contours and the minimum description length principle have been used in various ways but have rarely been applied to spline-fitting, while the PERM control-point insertion strategy is inspired by structural mechanics analysis. However, when these features are combined in the paradigm, it is possible to produce a good quality, *fully automatic* spline-fitting algorithm, which is capable of outperforming existing methods used in computer vision. While this method may be slower, future improvements can be made by using more complex methods (e.g., conjugate-gradient methods) for active contour optimization.

#### REFERENCES

- [1] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. John Wiley, 1973.
- [2] R. Bartels, J. Beatty, and B. Barsky, *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*. Morgan Kaufmann, 1987.

<sup>3</sup> This software is publicly available.  
See <http://svr-www.eng.cam.ac.uk/research/vision/>.

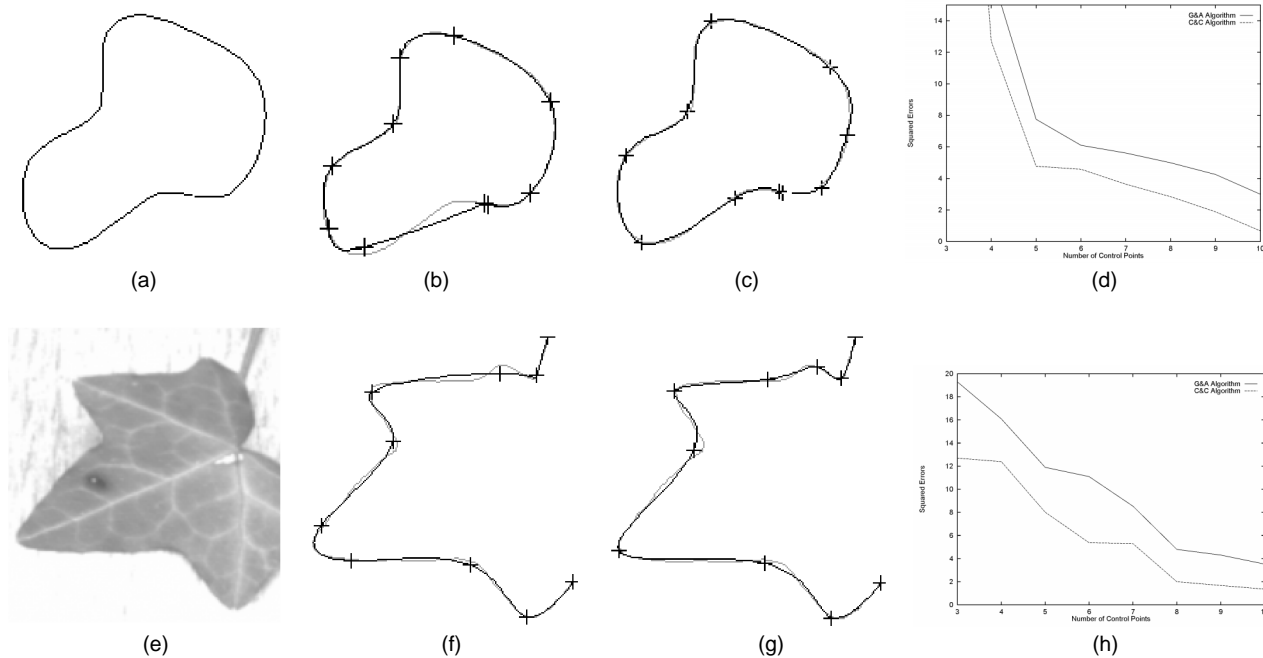


Fig. 8. (a), (e) Target curve.(b), (f), Splines fitted with G&A at 10 control points. (c), (g) Splines fitted with C&C at 10 control points. (d), (h) Comparison results.

[3] P. Saint-Marc, H. Rom, and G. Medioni, "B-Spline Contour Representation and Symmetry Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1,191-1,197, Nov. 1993.

[4] A. Guézic and N. Ayache, "Smoothing and Matching of 3-D Space Curves," *Int'l J. Computer Vision*, vol. 12, no. 1, pp. 79-104, 1994.

[5] F. Lu and E. Milios, "Optimal Spline Fitting to Planar Shape," *Signal Processing*, vol. 37, pp. 129-140, 1994.

[6] P. Dierckx, *Curve and Surface Fitting with Splines*. Clarendon Press, 1993.

[7] T. Cataldi and T. Rotunno, "Kalman Filter Optimized Cubic Spline Functions for Digital Smoothing," *Analytical Methods and Instrumentation*, vol. 2, no. 1, pp. 27-34, 1995.

[8] I. Qamar, "Method to Determine Optimum Number of Knots for Cubic Splines," *Comm. Numerical Methods in Engineering*, vol. 9, pp. 483-488, 1993.

[9] T.-J. Cham and R. Cipolla, "Geometric Saliency of Curve Correspondences and Grouping of Symmetric Contours," *Proc. Fourth Euro. Conf. Computer Vision*, B. Buxton and R. Cipolla, eds., vol. 1,064, Cambridge, England. *Lecture Notes in Computer Science*, pp. 385-398, Springer-Verlag, 1996.

[10] J. Sato and R. Cipolla, "Quasi-Invariant Parameterisations and Matching of Curves in Images," *Int'l J. Computer Vision*, vol. 28, no. 2, pp. 117-136, 1998.

[11] M. Powell, "Curve Fitting by Splines in One Variable," *Numerical Approximation to Functions and Data*, pp. 65-83. Athlone Press, 1970.

[12] D. Jupp, "The 'Lethargy' Theorem—A Property Of Approximation By  $\gamma$ -Polynomials," *J. Approximation Theory*, vol. 14, pp. 204-217, 1975.

[13] J. Rissanen, "Minimum-Description-Length Principle," *Encyclopedia of Statistical Sciences*, vol. 5, pp. 523-527, 1987.

[14] S. Crandall, N. Dahl, and T. Lardner, *An Introduction to the Mechanics of Solids*, Eng. Mechanics Series, Second ed. McGraw-Hill, 1978.

[15] T.-J. Cham, *Geometric Representation and Grouping of Image Curves*. PhD thesis, Dept. of Eng., Univ. of Cambridge, Aug. 1996.

[16] Y. Leclerc, "Constructing Simple Stable Descriptions for Image Partitioning," *Int'l J. Computer Vision*, vol. 3, no. 1, pp. 73-102, 1989.