

Automatic, Accurate Surface Model Inference for Dental CAD/CAM

Chi-Keung Tang¹, Gérard Medioni¹, and François Duret²

¹ Inst. Rob. Intell. Sys., University of Southern California, Los Angeles, CA 90089-0273.
{chitang, medioni}@iris.usc.edu

² School of Dentistry, University of Southern California, Los Angeles, CA 90089-0641.
duret@hsc.usc.edu

Abstract. Dental CAD/CAM offers the prospects of drastically reducing the time to provide service to patients, with no compromise on quality. Given the state-of-the-art in sensing, design, and machining, an attractive approach is to have a technician generate a restorative design in wax, which can then be milled by a machine in porcelain or titanium. The difficulty stems from the inherent outlier noise in the measurement phase. Traditional techniques remove noise at the cost of smoothing, degrading discontinuities such as anatomical lines which require accuracy up to 5 to 10 μm to avoid artifacts. This paper presents an efficient method for the automatic and accurate data validation and 3-D shape inference from noisy digital dental measurements. The input consists of 3-D points with spurious samples, as obtained from a variety of sources such as a laser scanner or a stylus probe. The system produces faithful smooth surface approximations while preserving critical curve features such as grooves and preparation lines. To this end, we introduce the *Tensor Voting* technique, which efficiently ignores noise, infers smooth structures, and preserves underlying discontinuities. This method is non-iterative, does not require initial guess, and degrades gracefully with spurious noise, missing and erroneous data. We show results on real and complex data.

1 Introduction and overall approach

Dental CAD/CAM has been revolutionizing dentistry for the past ten years, and is ubiquitous in every major dental group and laboratory today. Its main components are data acquisition, modeling, and milling systems. The only system widely used commercially is the CEREC system, produced by Siemens Inc. It is a self-contained unit with an imaging camera, a monitor, and an electrically controlled machine to mill inlay and onlay restorations from ceramic blocks. The accuracy is not good, and significant manual processing is required. Another system of interest, developed by Duret *et al.*, was able to produce crowns with an average gap of 35 μm . The system is no longer commercially available, and suffered from lack of speed and a cumbersome interface.

An alternative approach, followed here, is to perform the restoration design manually, in a laboratory, by using a conventional medium such as wax; then to transfer this physical design into a digital model. This information can be directly used to control a CAM machine to mill the restoration from a ceramic (or other) block. Typically, digital measurements are sampled from a wax model, using palpation or optical sensing [4].

Figure 1 shows the set up of one such commercial, multiple-view registration system, from which our data was obtained. As we shall see, though mostly accurate, the point sets obtained also contain many erroneous outlier readings. Given such noisy measurements, the challenge is to derive an accurate shape description automatically. Otherwise, even slight errors may result in artifacts which need to be corrected manually, or worse, may make the restoration unusable.

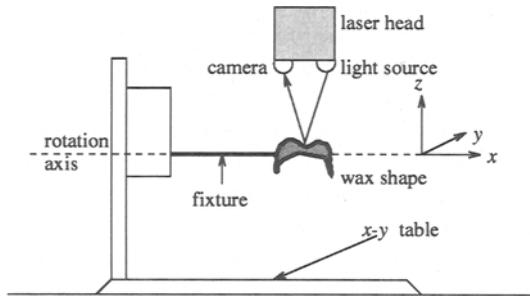


Fig. 1. Data acquisition by a laser digitizer (courtesy of GC Corporation, Tokyo, Japan).

While filter-based techniques such as discrete Fourier transform [10] are effective in suppressing spurious samples, they often “oversmooth”, degrading sharp discontinuities and distinct features that correspond to important anatomical (preparation) lines and features. At present, unless intensive human intervention is used, it is impossible to construct an accurate representation that respects both medical and dental criteria requiring accuracy of up to 5 to 10 μm [5]. Recently, much progress has been made in computer vision for robust surface inference from clouds of points. Their application should result in better 3-D dental model descriptions. This paper presents an integrated approach for inferring dental models in the form of surfaces (for capturing smoothness) and 3-D curves (for preserving shape discontinuities) from noisy dental data.

Much work has been done in surface fitting to clouds of points. The majority of the approaches use the deformable model approach (first proposed by Kass *et al.* in [9] and then in [19]) which attempts to deform an initial shape for fitting the input, using energy minimization. The work by Boulton and Kender [2], Poggio and Girosi [13], Blake and Zisserman [3], Fua and Sander [6], Szeliski *et al.* [14], and many others belong to this category. Physics-based approaches proposed by Terzopoulos *et al.* [17, 18] model the problem as a multi-particle system governed by physics laws and equations. The initial surface (or model), originally in equilibrium, is subject to external forces exerted at each data point. Such forces make the system converge to another equilibrium. Hoppe *et al.* [8] and Boissonnat [1] use computational geometry techniques, treating the data as graph vertices and constructing edges using local properties.

Most methods above are computationally expensive as an iterative process is taking place. Also, they have limited potential in inferring faithful 3-D models from dental

data: most do not work in the presence of spurious noise; and surface discontinuities are usually smoothed out. Our method attacks these problems by applying [7, 15, 16], in which a non-linear voting process is used to achieve efficient feature segmentation and discontinuity detection. Our method is robust to outlier noise since its effect is reduced by accumulating a large number of tensor votes.

Our overall approach is depicted in Figure 2. Each input point is first quantized in a 3-D voxel array. A preprocessing step is then applied to estimate the normal to the surface. This step, as well as the surface and curve inference processes, is realized by *tensor voting*, which is outlined in section 2. Consequently, two independent 3-D *saliency maps*, one for surface (SMap) and one for curve (CMap), are produced. More details can be found in [7, 11]. Local extrema in these maps are extracted, resulting in a triangulation mesh (surfaces) and a set of directed and connected 3-D poly-line segments (curves). These features are then refined (if possible) to localize detected discontinuities and remove inconsistent surface patches (such as the spurious patches labeled in Figure 2). Finally, the surfaces and curves inferred in the 3-D model are coherently integrated. Feature extraction and integration are summarized in section 3 (also detailed in [15, 16]). Results on real dental data are shown in section 4. Owing to space limitation, we omit many details in the present coverage, and refer readers to [7, 11, 15, 16] for more technical and mathematical details.

2 Tensor voting and saliency maps

The noisy dental data is in fact a scalar field, for which a preprocessing step for normal estimation at each point, or “vectorization”, is required. Having produced a vector field, then, a “densification” step is needed for surface and curve extraction. Both processes are realized by *tensor voting* [11], in which data are represented by *tensors*, and data communication is achieved by *voting*.

Tensor representation. A point in the 3-D space can assume either one of the three roles: surface patch, discontinuity (curve or point junctions), or outlier. Consider the two extremes, in which a point on a smooth surface is very certain about its surface (normal) orientation, whereas a point on a (curve or point) junction has absolute orientation uncertainty. This whole continuum is thus abstracted as a general, second-order symmetric 3-D *tensor*, which can be visualized geometrically as a 3-D **ellipsoid** (Figure 3). Such an ellipsoid can be fully described by the corresponding eigensystem with its three eigenvectors \hat{V}_{max} , \hat{V}_{mid} , and \hat{V}_{min} and the three corresponding eigenvalues $\lambda_{max} \geq \lambda_{mid} \geq \lambda_{min}$. Rearranging the eigensystem, the 3-D ellipsoid is given by: $(\lambda_{max} - \lambda_{mid})\mathbf{S} + (\lambda_{mid} - \lambda_{min})\mathbf{P} + \lambda_{min}\mathbf{B}$, where $\mathbf{S} = \hat{V}_{max}\hat{V}_{max}^T$ defines a *stick tensor*, $\mathbf{P} = \hat{V}_{max}\hat{V}_{max}^T + \hat{V}_{mid}\hat{V}_{mid}^T$ defines a *plate tensor*, and $\mathbf{B} = \hat{V}_{max}\hat{V}_{max}^T + \hat{V}_{mid}\hat{V}_{mid}^T + \hat{V}_{min}\hat{V}_{min}^T$ gives a *ball tensor*. These tensors define the three *basis tensors* for any 3-D ellipsoid.

Geometric Interpretation. The *eigenvectors* encode *orientation (un)certainties*: surface orientation (normal) is described by the *stick tensor*, which indicates *certainty* in a single direction. *Uncertainties* are abstracted by two other tensors: *curve junction* results from two intersecting surfaces, where the uncertainty in orientation only spans a single *plane* perpendicular to the tangent of the junction curve, and thus described by a *plate tensor*. At *point junctions* where more than two intersecting surfaces are present, a *ball tensor*

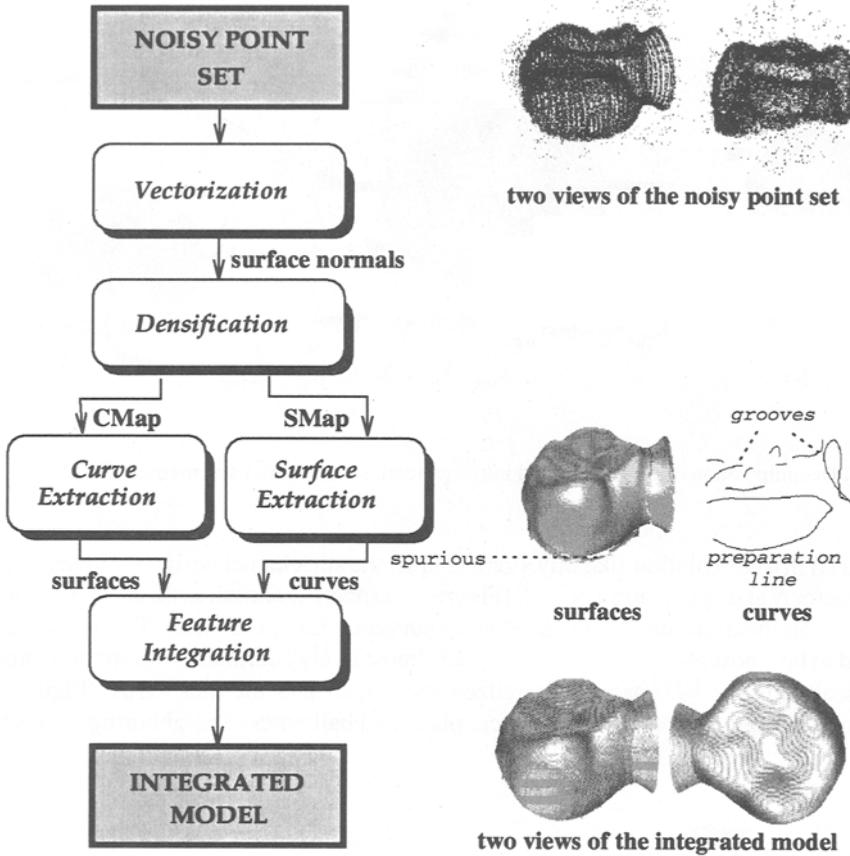


Fig. 2. Overall strategy of our method, using *Crown-24* as illustration.

is used since all orientations are equally probable. The *eigenvalues* encode the *magnitudes* of orientation (un)certainties, since they indicate the size of the corresponding 3-D ellipsoid. We define three *dense vector maps*, in which each voxel of these maps has a 2-tuple (s, \bar{v}) , where s is a scalar indicating *saliency* and \bar{v} is a unit vector:

- *Surface map (SMap)*: $s = \lambda_{max} - \lambda_{mid}$, and $\bar{v} = \hat{V}_{max}$ indicates the normal direction.
- *Curve map (CMap)*: $s = \lambda_{mid} - \lambda_{min}$, and $\bar{v} = \hat{V}_{min}$ indicates the tangent direction.
- *Junction map (JMap)*: $s = \lambda_{min}$, \bar{v} is arbitrary.

Tensor voting. Having outlined the tensor formalism, we now describe the *voting* algorithm for obtaining the tensor representation at *each* voxel and the above dense vector maps, thus achieving field densification. Suppose that we already have a vector field for densification (vectorization will be described shortly). First, each input vector is encoded as a general tensor, which is actually a very thin and elongated ellipsoid. Then, these input tensors are made to align with predefined, discrete versions of the three basis tensors

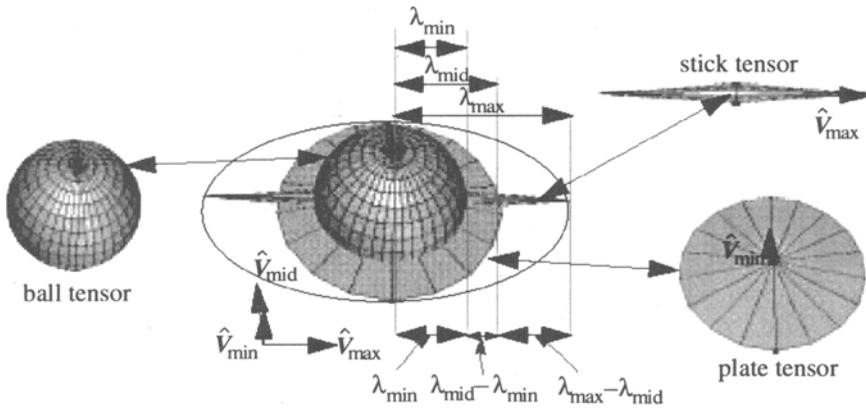


Fig. 3. Decomposition of a 3-D ellipsoid into a general, second-order symmetric 3-D tensor.

(*kernels*) in a convolution-like way. For example, the stick kernel design is: Given a normal vector N at origin O and a point P (Figure 4), a sphere is chosen as the most likely and natural continuation between O and P , since curvature is kept constant. The normal to P is said to be “most likely”. The set of all such “most likely” normals (with strength properly scaled) in the 3-D space is discretized and collected as the stick kernel (Figure 4). When each input tensor has cast its stick, plate, and ball votes to neighboring voxels by

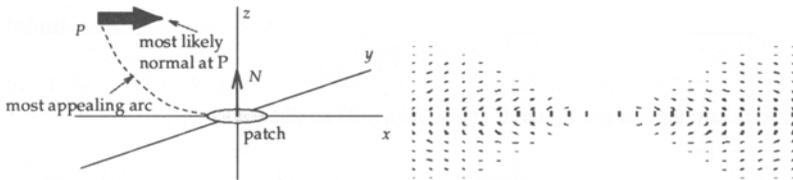


Fig. 4. The design, and one projection view of Stick kernel.

aligning with the respective *dense* basis kernels, each voxel in the volume receives a set of directed votes. These directed votes are collected, using tensor addition, as a 3×3 covariance matrix. Then, we diagonalize this matrix into the corresponding eigensystem and the above tensor formalism is applied.

Vectorizing scalar fields. Vectorizing a scalar field into a dense vector field is implemented *exactly* as the densification described above, except that each scalar input site is encoded as *ball* tensor, since no orientation information is given. After the voting step, the original input “balls” are replaced by true ellipsoids, which encode surface normal and curve tangent information.

3 Feature extraction and integration

Local extrema in CMap (resp. SMap) are extracted by our extremal curve (resp. surface) algorithms [16]. The extracted features are integrated to produce a coherent surface and curve description, whenever possible. These processes are outlined in the following:

Extremal curve extraction. Each voxel in the CMap holds a 2-tuple (s, \bar{t}) , where s is curve saliency and \bar{t} indicates tangent direction. Suppose the CMap is continuous in which (s, \bar{t}) is defined for every point p in 3-D space. A point p with (s, \bar{t}) is on an extremal curve if any displacement from p on the plane normal to \bar{t} will result in a lower s value, i.e. $\frac{\partial s}{\partial u} = \frac{\partial s}{\partial v} = 0$, where u and v define the plane normal to \bar{t} at the voxel center. This definition therefore involves the detection of zero crossing in the u - v plane normal to \bar{t} . To do this, we compute the saliency gradient \bar{g} as, $\bar{g} = \nabla s = \left[\frac{\partial s}{\partial x} \quad \frac{\partial s}{\partial y} \quad \frac{\partial s}{\partial z} \right]^T$. Define $\bar{q} = \mathbf{R}(\bar{t} \times \bar{g})$ where \mathbf{R} defines a frame aligned with the u - v plane. By construction, \bar{q} is the projection of \bar{g} onto the plane normal to \bar{t} . Therefore, an extremal curve is the locus of points for which $\bar{q} = \bar{0}$. The corresponding discrete \bar{q} can be similarly defined, from which a tracing algorithm can be readily defined. The output is a set of connected and oriented poly-line segments representing the extremal curves.

Extremal surface extraction. Each voxel in the SMap holds a 2-tuple (s, \bar{n}) where s indicates surface saliency and \bar{n} denotes normal direction. As before, suppose the SMap is continuous in which (s, \bar{n}) is defined for every point p in 3-D space. A point is on an extremal surface if its saliency s is locally extremal along the direction of the normal, i.e., $\frac{ds}{dn} = 0$. This definition involves the detection of zero crossing on the line aligned with \bar{n} , which is computed by defining a scalar $q = \bar{n} \cdot \bar{g}$, where \bar{g} was defined earlier. Therefore, an extremal surface is the locus of points for which $q = 0$. We define the corresponding discrete q , which can be processed directly by the Marching Cubes algorithm [12]: A polygonal mesh is thus produced.

Feature integration. The final phase integrates surfaces, curves and junctions to produce an unified 3-D model description. While tensor voting produces good results on smooth structures, it only detects discontinuities but does not properly localize them. This is because the SMap and CMap are independently interpreted. To integrate them, a curve detected in CMap is first treated as *surface inhibitor* in SMap so that when a smooth surface is traced, it is forbidden to enter region close to any detected discontinuities. With such a “trimmed” surface, the same curve is then treated as *surface exciter* for computing precise and natural surface junction. A set of “extended” surfaces are produced, which will undergo subsequent refinement.

4 Results

We tested our system on a variety of crowns and inlays. An inlay is a cast filling that is used to replace part of a tooth, while a crown is a larger restoration. The data is acquired using the set-up shown in Figure 1. The wax shape, obtained from the dental laboratory, is rotated about the x -axis in 15° (or 90°) increment so that 24 (or 4) successive views are visible to the sensor. The following tabulates the running times (in min) on a Sun Ultra 2 (Model 1300 with 256 MB RAM).

Data	No. of points	Voting	Feature extraction	Feature integration	Total
<i>Mod-4</i>	5217	4	2	-	6
<i>Inlay-4</i>	2447	3	1	-	4
<i>Crown-4</i>	8844	6	2	-	8
<i>Crown-24</i>	60095	35	5	5	45

Table 1. Processing times (in min) on Sun Ultra 2 (Model 1300 with 256 MB RAM).

Mod-4. We first demonstrate in this and the next two examples the graceful degradation of tensor voting in the presence of spurious and missing samples. Because of this, feature integration is skipped (as outlined in section 3). A set of only 4 views of a *Mod* are digitized and quantized in 100x100x100 array, which contains 5217 points. This data set is difficult because it has a complicated shape, and has many missing and misleading data resulting from fewer registered views and self occlusion. Using the inferred surface and curve model, we can perform **data validation** (Figure 5(a)): after voting, we validate each input point by simple thresholding of its associated surface or curve saliency. This shows a simple application of spurious noise elimination with the availability of a faithful model. The extracted extremal surfaces and curves are depicted in Figure 6(a). Note that surfaces are only correct from the discontinuity curves, for which they have low surface saliencies. However, these discontinuities are detected and marked since they are characterized by high curve saliencies. These detected junction curves allow further refinement or intervention to take place.

Inlay-4. A set of 4 views of an *Inlay* are digitized, having 2447 data points quantized in a 70x50x70 array. This is a very complicated surface, and the data set is noisy, and contains many missing and erroneous readings. Result of data validation is shown in Figure 5(b). The extremal surfaces and curves extracted are shown in Figure 6(b).

Crown-4. A sparse set of only 4 views of a *Crown* are digitized and quantized, using a 100x100x100 voxel array. This data set contains 8844 points. We can recover the underlying surface model, which is automatically marked with curves representing grooves and preparation lines. With a faithfully inferred model, we can perform data validation (Figure 5(c)). Figure 7(a) shows the resultant extremal surfaces and curves.

Crown-24. A set of 24 views of a *Crown* are registered. The data set contains 60095 points, quantized in a 100x100x100 array. We can detect the upper and lower surfaces of the *Crown*. The detected preparation line and the grooves are in turn used to produce a coherently integrated surface and curve description. Result of data validation is shown in Figure 5(d). Figure 7(b) shows the extracted extremal surfaces and curves.

5 Conclusion

We have presented an efficient and robust method, *tensor voting*, to automatically generate faithful dental models, consisting of surfaces and 3-D curves, from very noisy measurements. As shown by our results, while further improvement can still be made (or simply provide us with more views of data), our working prototype indeed demonstrates

a very promising potential on significantly improving current dental CAD/CAM technology, as our system produces faithful results despite that some sample data are noisy, missing and confusing. We not only interpolate smooth structures, but also respect important anatomical lines, and filter out spurious outlier points as well, and hence offer the prospects for reducing time in providing necessary service. We expect to perform more quantitative error analysis, investigate the right level of details (e.g. data quantization), and compare our results on same complex data with outputs from other sites.

References

1. J.D. Boissonnat, "Representation of Objects by Triangulating Points in 3-D space", in *Proc. Intl. Conf. Patt. Recogn.*, pp.830–832, 1982.
2. T.E. Boulton and J.R. Kender, "Visual Surface Reconstruction Using Sparse Depth Data", in *Proc. IEEE Comput. Vision Patt. Recogn.* (Miami Beach, FL), pp.68–76, 1986.
3. A. Blake and A. Zisserman, "Invariant Surface Reconstruction Using Weak Continuity Constraints", in *Proc. IEEE Comput. Vision Patt. Recogn.* (Miami Beach, FL), pp.62–67, 1986.
4. F. Duret, The Dental Robotic: "State of the Science Dental CAD/CAM", Clark's Clinic. Den. Ency, Philadelphia, Lincholl ed.
5. F. Duret, "Functionality and accuracy of Sopha CAD/CAM system today", in *Proc. Intl. Conf. Comput.*, vol. 1, p.500.
6. P. Fua and P. Sander, "Segmenting Unstructured 3D Points into Surfaces", in *Proc. Euro. Conf. Comput. Vision* (Santa Margherita Ligure, Italy), pp.676–680, 1992.
7. G. Guy and G. Medioni, "Inference of Surfaces, 3-D Curves, and Junctions from Sparse, Noisy 3-D Data", *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol.19, no.11, pp.1265–1277, 1997.
8. H.Hoppe, T.DeRose, T.Duchamp, J.McDonald, W.Stuetzle, "Surface Reconstruction from Unorganized Points", *Computer Graphics*, 26, pp.71–78, 1992.
9. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models", *Int. J. Comput. Vision*, pp.321–331, 1988.
10. H.Kimura, H.Sohmura, et al., "Three Dimensional Shape Measurement of Teeth (part 3)", *J. Osaka Uni. Dent Sch.*
11. M.-S. Lee and G. Medioni, "Inferring Segmented Surface Description from Stereo Data", in *Proc. IEEE Comput. Vision Patt. Recogn.*, (Santa Barbara, CA), pp.346–352, 1998.
12. W.E. Lorensen and H.E. Cline, "Marching Cubes: A High Resolution 3D Surface Reconstruction Algorithm", *Computer Graphics*, 21(4), 1987.
13. T. Poggio and F. Girosi, "A theory of networks for learning", *Science*, pp.978–982, 1990.
14. R. Szeliski, D. Tonnesen, and D. Terzopoulos, "Modeling Surfaces of Arbitrary Topology with Dynamic Particles", in *Proc. IEEE Comput. Vision Patt. Recogn.*, pp.82–85, Jun 1993.
15. C.-K. Tang and G. Medioni, "Integrated Surface, Curve and Junction Inference from Sparse 3-D Data Sets", in *Proc. IEEE Intl. Conf. Comput. Vision*, pp.818–824, Jan 1998.
16. C.-K. Tang and G. Medioni, "Extremal Feature Extraction from Noisy 3-D Vector and Scalar Fields", to appear in *Proc. IEEE Visualization Conf.*, Oct 1998.
17. D. Terzopoulos and D. Metaxas, "Dynamic 3D models with local and global deformations: deformable superquadratics", *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol. 13, no. 7, pp.91–123, 1991.
18. D. Terzopoulos and M. Vasilescu, "Sampling and reconstruction with adaptive meshes", in *Proc. IEEE Comput. Vision Patt. Recogn.*, (Lahaina, Maui, HI), pp.70–75, 1991.
19. M. Vasilescu and D. Terzopoulos, "Adaptive Meshes and Shells: Irregular Triangulation, Discontinuities, and Heretical Subdivision", in *Proc. IEEE Comput. Vision Patt. Recogn.*, pp.829–832, 1992.

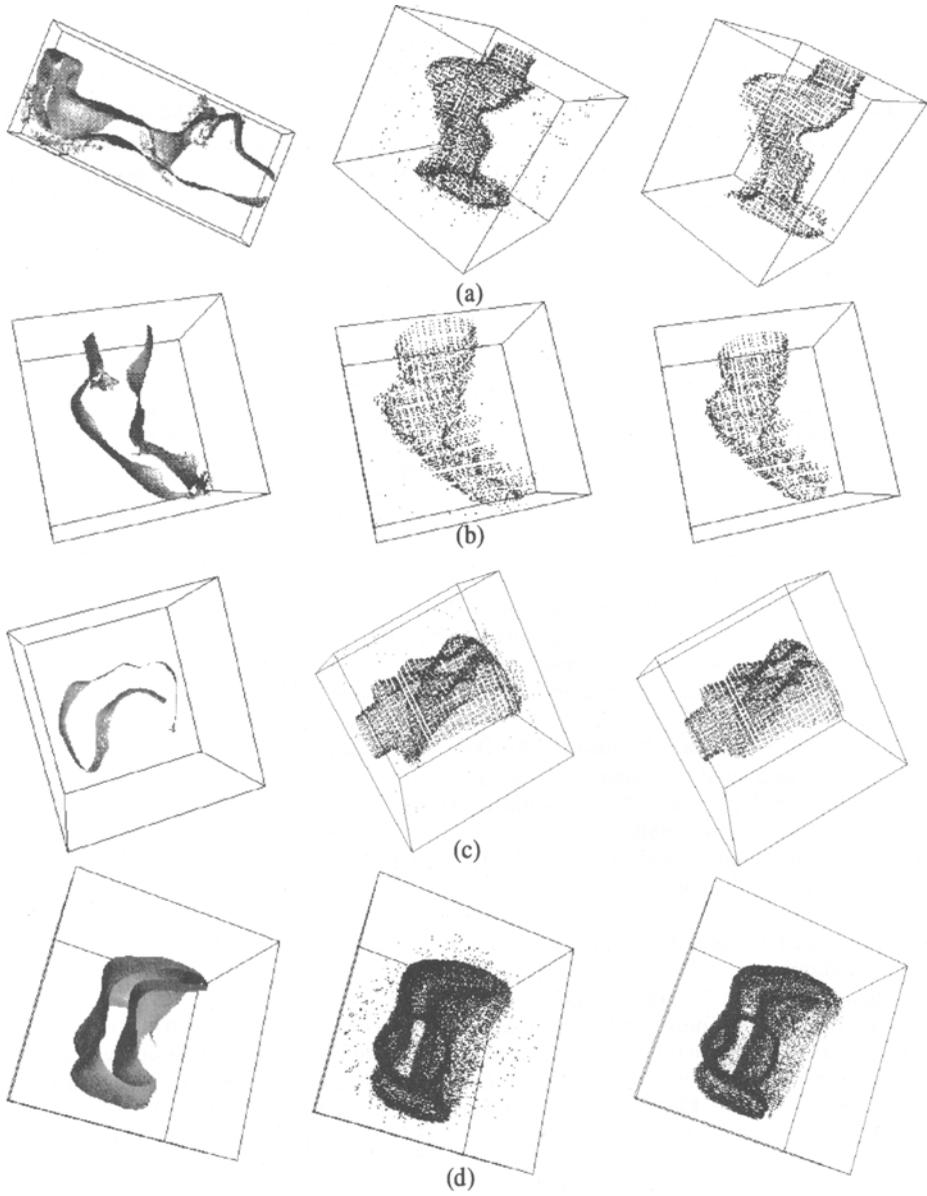


Fig. 5. A middle slice of extremal surfaces, the original noisy data, and the validated data set, for (a) *Mod-4*, (b) *Inlay-4*, (c) *Crown-4*, and (d) *Crown-24*

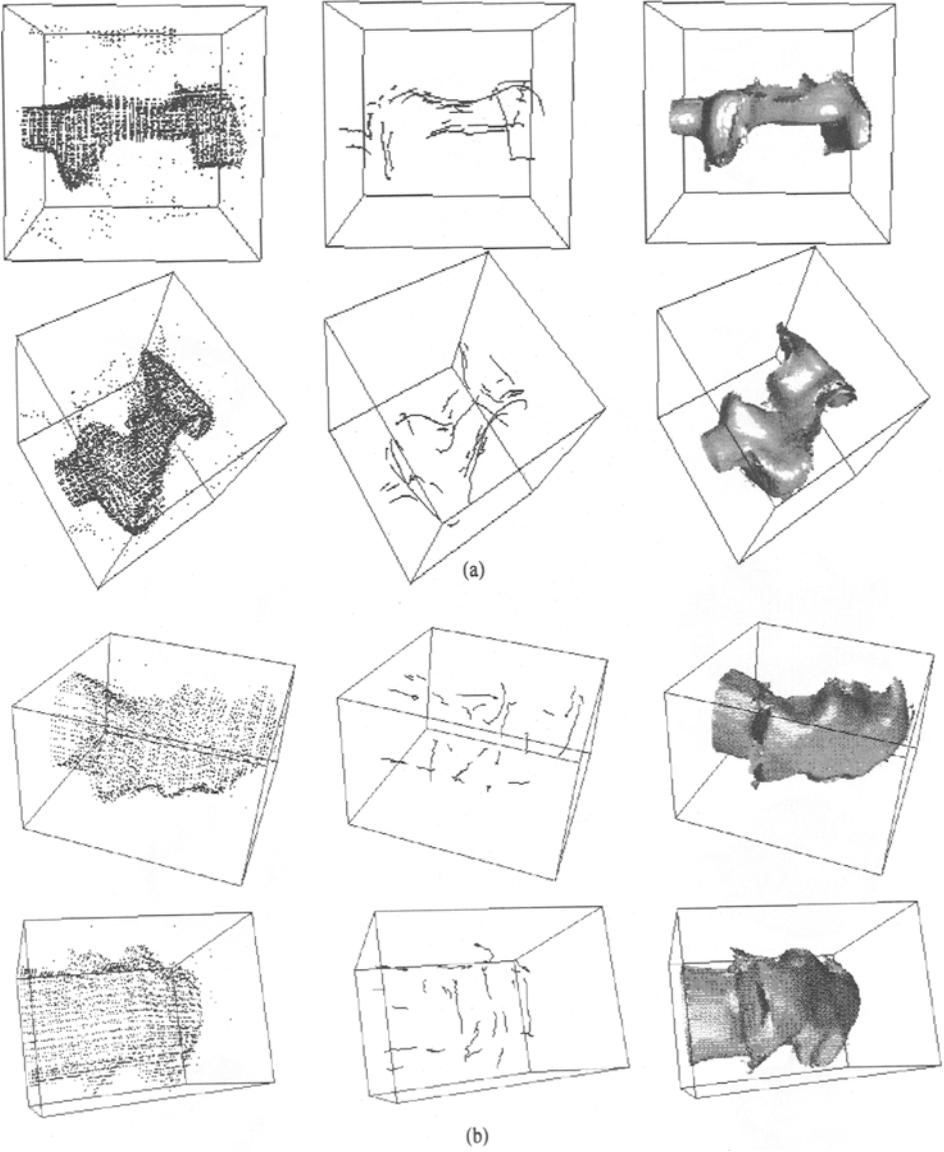


Fig. 6. Two views of the original noisy data, the extremal discontinuity curves and surfaces inferred for (a) *Mod-4* and (b) *Inlay-4*.

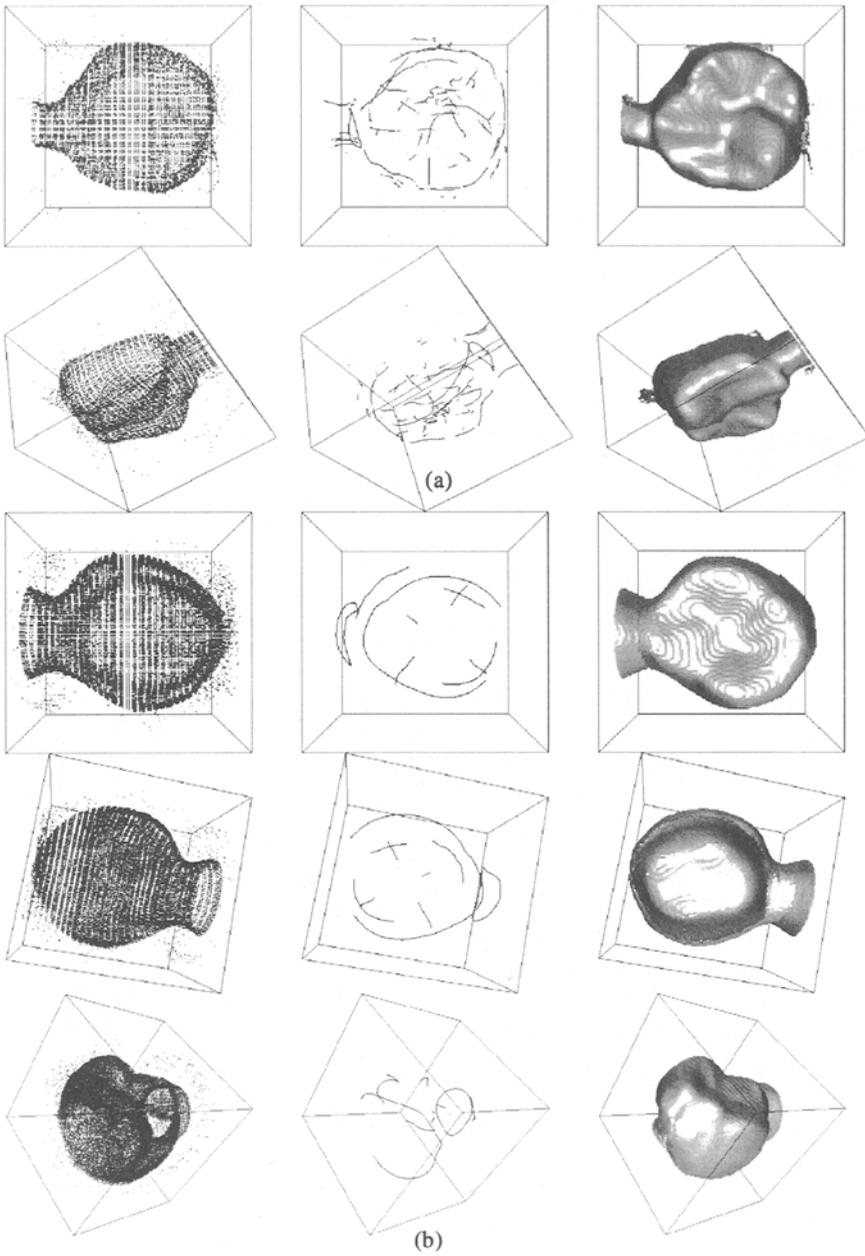


Fig. 7. Different views of the original noisy data, the extremal discontinuity curves and surfaces inferred for (a) *Crown-4* and (b) *Crown-24*.