



Contour-based shape representation and description techniques for 3D Objects

Rajesh Chhikara

Research Scholar, Singhania University, Rajasthan, India

Sunil Kumar

Maharaja Agrasen Institute of Technology, Delhi, India

Gopinath S

Robert Bosch Engineering and Business Solutions Ltd, Bangalore, India

Rita Chhikara

ITM University, Gurgaon, Haryana, India

Abstract: Recent developments in techniques for modeling, digitizing and visualizing 3D shapes has led to an explosion in the number of available 3D models on the Internet and in domain-specific databases. This has led to the development of 3D shape retrieval systems that, given a query object, retrieve similar 3D objects. For visualization, 3D shapes are often represented as a surface, in particular polygonal meshes, for example in VRML format. Often these models contain holes, intersecting polygons, are not manifold, and do not enclose a volume unambiguously. On the contrary, 3D volume models, such as solid models produced by CAD systems, or pixels models, enclose a volume properly. This paper surveys the literature on methods for content based 3D retrieval, taking into account the applicability to surface models as well as to volume models.

Keywords: 3D shape retrieval, Contour-based shape, global descriptors, Correspondence-based shape matching, Shape signature.

I. Introduction

The methods are evaluated with respect to several requirements of content based 3D shape retrieval, such as shape representation requirements, properties of dissimilarity measures, efficiency, discrimination abilities, ability to perform partial matching, robustness, and necessity of pose normalization. Finally, the advantages and limits of the several approaches in content based 3D shape retrieval are discussed. Shape representation and description techniques can be generally classified into two methods contour-based methods and region-based methods. The classification is based on whether shape features are extracted from the contour only or are extracted from the whole shape region. Under each class, the different methods are further divided into structural approaches and global approaches. This sub-class is based on whether the shape is represented as a whole or represented by segments/sections (primitives). These approaches can be further distinguished in to space domain and transform domain, based on whether the shape features are derived from the spatial domain or from the transformed domain.

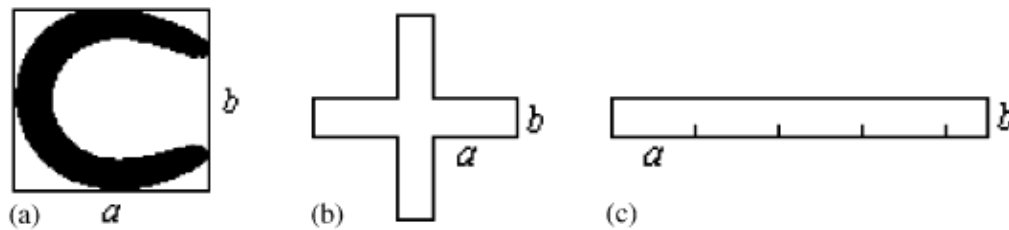
Laser scanning has been applied to obtain archives recording cultural heritage like the David of Michelangelo [1], and the Minerva of Arezzo [2]. Furthermore, archives containing domain-septic shape models are now accessible on the Internet. Examples are the National Design Repository, an online repository of CAD models [3, 4], Shape Sifter, a retrieval system for CAD databases [5, 6] and the Protein Data Bank, an online archive of structural data of biological macromolecules [7, 8]. Very recently, the first benchmarks for mesh models have been made available by Princeton University [9, 12], the University of Konstanz [10], Utrecht University [11, 13], and Purdue University [14, 15]. The Princeton Shape Benchmark database contains 1,814 models downloaded from the web, subdivided into a training set and a test set, containing 907 models each, classified into 90 and 92 classes respectively. The database at the University of Konstanz provides a test set containing 473 models classified into 55 classes, and 1,366 unclassified models. The database at Utrecht University consists of 512 models classified into six categories. The Purdue Engineering Shape Benchmark contains 1,391 models classified into solids of revolution, prisms, and flat-thin wall models

II. Contour-based shape representation and description techniques

Contour shape techniques only exploit shape boundary information. There are generally two types of very different approaches for contour shape modeling: continuous approach (global) and discrete approach (structural). Continuous approaches do not divide shape into sub-parts; usually a feature vector derived from the integral boundary is used to describe the shape. The measure of shape similarity is usually a metric distance between the

acquired feature vectors. Discrete approaches break the shape boundary into segments, called primitives using a particular criterion.

Fig. 1 Shape eccentricity and circularity



The final representation is usually a string or a graph (or tree), the similarity measure is done by string matching or graph matching. In the following we discuss these two types of approaches.

Global contour shape representation techniques usually compute a multi-dimensional numeric feature vector from the shape boundary information. The matching between shapes is a straight forward process, which is usually conducted by using a metric distance, such as Euclidean distance or city block distance. Point (or point feature) based matching is also used in particular applications.

III. Simple shape descriptors

Common simple global descriptors are area, circularity (perimeter²=area), eccentricity (length of major axis/length of minor axis), major axis orientation, and bending energy. These simple global descriptors usually can only discriminate shapes with large differences; therefore, they are usually used as filters to eliminate false hits or combined with other shape descriptors to discriminate shapes. They are not suitable to be standalone shape descriptors. For example, the eccentricity of the shape in Fig. 1(a) is close to 1 ($a = b$), it does not correctly describe the shape, because perceptually it is an elongated shape. In this case, circularity is a better descriptor. The two shapes in Fig. 1(b) and (c) have the same circularity ($a = 2b$), however, they are very different shapes. In this case, eccentricity is a better descriptor.

Other simple global contour shape descriptors have been proposed by Peura and Iivarinen. These descriptors include convexity, ratio of principle axis, circular variance and elliptic variance.

IV. Correspondence-based shape matching

Correspondence-based shape matching works in the space domain. In contrast to feature-based shape representation techniques, correspondence-based shape matching measures similarity between shapes using point-to-point matching. In other words, every point on the shape is treated as a feature point. The matching is conducted on 2-D space.

Hausdorff distance is a classical correspondence-based shape matching method, it has often been used to locate objects in an image and measure similarity between shapes [4 -9]. Given two shapes represented by two set of points:

$A = \{a_1; a_2; : : : a_p\}$ and $B = \{b_1; b_2; : : : b_q\}$, the Hausdorff distance between A and B is defined as:

$$H(A; B) = \max(h(A; B); h(B; A)); \quad (2.1)$$

where

$$h(A; B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (2.2)$$

$\in A \in B$

and $\| \cdot \|$ is the underlying norm on the points of A and B, usually Euclidean distance. However, this distance measure is too sensitive to noise or outlier. A single point in A that is far from anything in B will cause $h(A; B)$ to be large.

Therefore, a modified Hausdorff distance is introduced by Rucklidge:

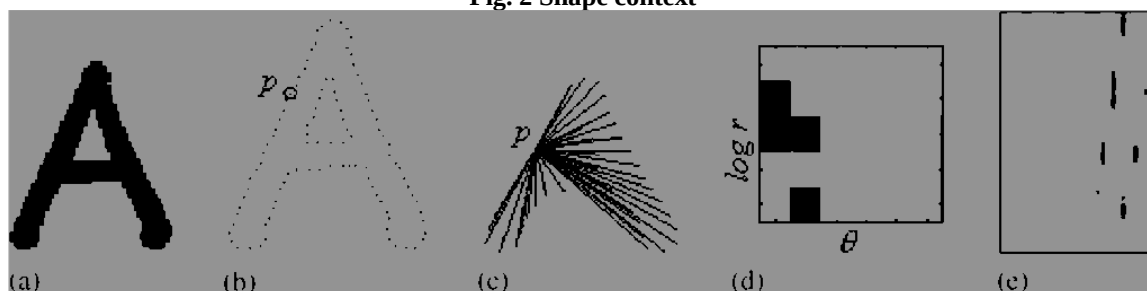
$$hf(A, B) = fth_{a \in A} \min_{b \in B} \|a - b\|,$$

where $fth_{a \in A} g(x)$ denotes the f th quantile value of $g(x)$ over set X, for some value of f between 0 and 1. For example, the 1th quantile value is the maximum and the 1/2th quantile value is the median. In practice, f is usually set to be 1/2. The advantage of shape matching using Hausdorff distance is that shape can be matched partially. However, the Hausdorff distance is not translation, scale and rotation invariant. In order to match a model shape with a shape in the image, the model shape has to be overlapped on the image in different positions, different orientations and different scales. As the result, the matching is prohibitively expensive.

Chetverikov and Kenokh propose an efficient matching by chamfer distance transform. This reduces a translation invariant matching from $O(N^2)$ of computations to $O(N)$, where N is the number of image points. Traditional Hausdorff shape matching only allows similarity matching, that is, shape matching allowing translation, rotation and scaling changes. Rucklidge extends Hausdorff distance matching into affine invariant matching. For this purpose, a set of affine models are generated from the model shape.

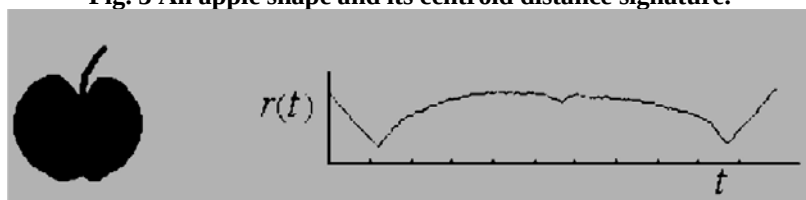
Since the space of affine transformations from the model shape is large, an efficient matching scheme is introduced by only examining a small part of the space of the affine transformations. Despite the efficiency effort, the matching load is still unacceptably high. The matching of even a small model shape with a normal image can take half an hour on an eight-processor Sun SPARC Server 1000.

Fig. 2 Shape context



Shape matching using Hausdorff distance is sensitive to noise and slight variations. Recently, Belongie et al. propose a correspondence-based shape matching method using shape contexts. Shape matching using shape contexts is an improvement to traditional Hausdorff distance based methods. It extracts a global feature, called shape context, for each corresponding point. The matching between corresponding points is then the matching between the context features. To extract the shape context at a point p , the vectors of p to all the other boundary points are found (Fig. 2(c)). The length r and orientation θ of the vectors are quantized to create a histogram map which is used to represent the point p (Fig. 2(d)). The histogram of each point is flattened and concatenated to form the context of the shape (Fig. 2(e)). To make the histogram more sensitive to positions of nearby points than to those of points farther away, these vectors are put into log-polar space. For example, in Fig. 2, (a) is a character shape, (b) is the edge image of (a), (c) is the set of vectors at point p , (d) is the histogram or context extracted from the vectors, (e) is the context of the shape.

Fig. 3 An apple shape and its centroid distance signature.



The matching of two shapes is done by matching two context maps of the shapes, which is a matrix-based matching. It minimizes the total cost of matching between one context matrix and all the permutations of another context matrix. In order to reduce the matching overhead, it samples the boundary at a number of points and uses the shortest augmenting path algorithm for the matrix matching.

V. Shape signature

A Shape signature represents a shape by a one dimensional function derived from shape boundary points. Many shape signatures exist, they include centroidal profile, complex coordinates, centroid distance (Fig. 3), tangent angle, cumulative angle, curvature, area and chord-length. Shape signatures are usually normalized into being translation and scale invariant. In order to compensate for orientation changes, shift matching is needed to find the best matching between two shapes. Most of the signature matching is normalized to shift matching in 1-D space, however, some signature matching requires shift matching in 2-D space, such as the matching of centroidal profiles. In either case, the matching cost is too high for online retrieval.

In addition to the high matching cost, shape signatures are sensitive to noise, and slight changes in the boundary can cause large errors in matching. Therefore, it is undesirable to directly describe shape using a shape signature. Further processing is necessary to increase its robustness and reduce the matching load. For example, a shape signature can be simplified by quantizing the signature into a signature histogram, which is rotationally invariant.

VI. Boundary moments

Boundary moments can be used to reduce the dimensions of the boundary representation. Assuming the shape boundary has been represented as a shape signature $z(i)$, the r th moment m_r and central moment μ_r can be estimated as

$$m_r = \frac{1}{N} \sum_{i=1}^N [z(i)]^r, \quad \mu_r = \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^r, \quad (2.4)$$

where N is the number of boundary points. The normalized moments and $N m_r = m_r / (\mu_2)^{r/2}$ and $\mu_r = \mu / (\mu_2)^{r/2}$ are invariant to shape translation, rotation and scaling. Less noise-sensitive shape descriptors can be obtained from $F1 = (\mu_2)^{1/2} / m_1$:

$$F2 = \mu_3 / (\mu_2)^{3/2} \quad F3 = \mu_4 / (\mu_2)^2.$$

The method in treats the amplitude of the shape signature function $z(i)$ as a random variable v and creates a histogram $p(v_i)$ from $z(i)$. Then, the r th moment is obtained by

$$\mu_r = \sum_{i=1}^k (v_i - m)^r p(v_i) \quad \text{and} \quad m = \sum_{i=1}^k (v_i p(v_i)). \quad (2.5)$$

The advantage of boundary moment descriptors is that it is easy to implement. However, it is difficult to associate higher order moments with physical interpretation.

VII. Elastic matching

Bimbo and Pala have proposed the use of elastic matching for shape based image retrieval. According to this approach, a deformed template is generated as the sum of the original template $\tau(s)$ and a warping deformation $\theta(s)$

$$\varphi(s) = \tau(s) + \theta(s), \quad (2.6)$$

where $\tau = (\tau_x, \tau_y)$ is a second order spline and $\theta = (\theta_x, \theta_y)$ is the deformation. The similarity between the original shape of the template and the shape of the object in the image is measured by minimizing a compound function:

$$\begin{aligned} F &= S+B+M \\ &= \alpha \int_{\theta} [(d\theta_x)^2 + (d\theta_y)^2] ds \\ &+ \beta \alpha \int_{\theta} [(d^2\theta_x)^2 + (d^2\theta_y)^2] ds \\ &+ \alpha \int_{\theta} IE(\varphi(s)) ds \end{aligned} \quad (2.7)$$

where IE is the object image, S and B are called strain energy and bend energy respectively, while M measures the degree of overlapping between the deformed template and the object in the image. The three quantitative measures are not sufficient to measure the similarity between shapes, therefore shape complexity N (measured as the number of 0's of the curvature function associated with the templates contour) and correlation C (between the curvature function associated with the template and that associated with the deformed one) are also taken into account in the similarity measure. Finally, the 1ve parameters (S ; B ; M ; N ; C) are classified by a back-propagation neural network.

This approach is not practical for online image retrieval, mainly because of the computation and matching complexity.

VIII. Stochastic method

Time-series models and especially autoregressive (AR) modeling has been used for calculating shape descriptors. Methods in this class are based on the stochastic modeling of a 1-D function f obtained from the shape as described in Section 2.1.3. A linear autoregressive model expresses a value of a function as the linear combination of a certain number of preceding values. Specifically, each function value in the sequence has some correlation with previous function values and can therefore be predicted through a number of, say, M observations of previous function values.

The autoregressive model is a simple predictor of the current radius by a linear combination of M previous radii plus a constant term and an error term:

$$f_i = \alpha + \sum_{j=1}^m \theta_j f_{i-j} + \sqrt{\beta} \omega_t, \quad (2.8)$$

where θ_j are the AR-model coefficients, m is the model order, i.e., tells how many preceding function values the model uses.

$\sqrt{\beta} \omega_t$ is the current error term or residual, reflecting the accuracy of the prediction. α is proportional to the mean of function values. The parameters $\{\alpha, \theta_1, \dots, \theta_m, \beta\}$ are estimated by using the least square (LS) criterion. The estimated θ_j are translation, rotation and scale invariant. Parameters α and β are not scale invariant, but the quotient $\alpha/\sqrt{\beta}$, which reflects signal-to-noise ratio of the boundary, is regarded as a invariant. Therefore, the feature vector $[\theta_1, \dots, \theta_m, \alpha/\sqrt{\beta}]$ is used as the shape descriptor.

The disadvantage of the AR method is that in the case of complex boundaries, a small number of AR parameters is not sufficient for an adequate description. The choice of m is a complicated problem and is usually decided empirically. Besides, the physical meaning associated with each θ_j is not clear.

IX. Scale space method

The problem of noise sensitivity and boundary variations in most spatial domain shape methods inspires the use of scale space analysis. The scale space representation of a shape is created by tracking the position of infection

points in a shape boundary filtered by low-pass Gaussian filters of variable widths. As the width (σ) of Gaussian filter increases, insignificant infections are eliminated from the boundary and the shape becomes smoother. The infection points that remain present in the representation are expected to be 'significant' object characteristics. The result of this smoothing process is an interval tree, called 'fingerprint', consisting of infection points. The difficulty with this type of approach is the interpretation of the final result.

Asada and Brady first attempted to interpret the interval tree acquired from scale space. Their interval trees are acquired from both Gaussian filter and second derivatives of Gaussian filter. The interpretation of the interval tree is based on detecting the peaks of the tree branches from higher scales to lower scales (Fig. 4(c)). Since the shapes under analysis are from hardware application, it is possible to interpret higher level primitive events from the detected peaks. In this application, the primitive events are defined as corner, smooth joint, end, crank, bump/dent.

Mokhtarian and Mackworth adopted Asada and Brady's interpretation method and extend it to shoreline registration. They call the acquired scale space signature as a curvature scale space (CSS) contour image. The peaks of individual branches in the CSS are detected. However, instead of interpreting these peaks as higher level primitive events, they are used for matching two curves under analysis. The matching proves to be very complex and expensive. It tries to find the best match between the contour branches in the two interval trees. For each individual matching, the template contour branch has to be scaled and shifted to accommodate scale and translation invariance.

During each individual matching, the template curve is also affinely transformed to match with the model curve. The method is later extended for shape retrieval. Since the old matching algorithm is too complex for retrieval, the new method developed a matching algorithm based on the two highest peaks in each of the two contour images. All the images under analysis are also scaled into the same size (same number of boundary points) prior to applying scale space. However, the four essential empirical parameters involved in the feature extraction and matching processes make this algorithm unstable.

Daoudi and Matusiak interpret the interval tree obtained from scale space as a geodesic topology. The matching between two shapes now turns into matching between two scale space images using the geodesic distance measure suggested by Eberly. The matching is actually a point to point 2-D matching between the two scale space images, this can be impractical if shapes in the database are complex, resulting in very high interval tree.

X. Spectral transform

Spectral descriptors overcome the problem of noise sensitivity and boundary variations by analyzing shape in spectral domain. Spectral descriptors include Fourier descriptor (FD) and wavelet descriptor (WD), they are derived from spectral transforms on 1-D shape signatures described in Section 2.1.3.

One of the most widely used shape description methods is FD. Conventional FD methods only deal with closed curve, however, Lin et al. and Mitchell et al. used FD to describe partial shapes. Arnbø et al. introduced the affine-invariant FD to take into consideration of affine shape description. Granlund introduced the Fourier invariants which describe the rotational symmetry of shapes. Rauber proposed a UNL FD (named after Universidade Nova de Lisboa, Portugal) which is able to describe disjointed or articulated contour shape. The UNL FD is acquired by applying 2-D Fourier transform on the UNL transformed shape image. Even though a feature selection process is followed, the dimension of the feature vector acquired this way is very high. Richard and Hemami introduced a complex distance measurement, called the true distance measurement, for measuring the similarity between two sets of FDs. Since the true distance measurement requires two Fourier transforms for each matching, it involves 15 times more computation than a normal distance measurement. Rui et al. proposed a distance measurement to classify similarity transformed characters using Fourier transformed coefficients. This distance measurement is the weighting sum of the variance of magnitude ratios and the variance of phase difference between two sets of Fourier coefficients. The Fourier coefficients are derived from Fourier reconstructed shape boundary rather than from original boundary. This is not different from FD derived from a smoothed boundary. Eichmann et al. proposed the use of a short-time Fourier descriptor (SFD) for shape description, however, Zhang and Lu have found that SFD is outperformed by conventional FD methods in shape retrieval. This is because SFD cannot capture global shape features although it can capture local shape features more accurately.

Recently, several researchers have proposed the use of WD for shape description. Although WD has the advantage over FD in that it is of multi-resolution in both spatial space and spectral space, the increase of spatial resolution will certainly sacrifice frequency resolution. For example, in, only wavelet coefficients of the few low frequencies are used to represent shape. Most importantly, the complicated matching scheme of wavelet representation makes it impractical for online shape retrieval. In, the similarity measurement algorithm needs $2L \times N$ all-level shift matching, where L is the number of levels of resolution of the wavelet transform and N is the number of normalized boundary points. In, the number of matchings for similarity measurement is not only large but also dependent on the complexity of the shape, since the similarity measurement is the all level shift matching of all the zero-crossing points of the wavelet approximation of the shape. Apart from the matching complexity, the dyadic wavelets used can rarely associate the feature segments on the shape boundary. Therefore, WD suffers the same drawback in primitive determination as that in the structural approach.

FD is backed by the well-developed and well-understood Fourier theory. The advantages of FD over many other shape descriptors are (i) simple to compute; (ii) each descriptor has specific physical meaning; (iii) simple to do normalization, making shape matching a simple task; (iv) captures both global and local features. With sufficient features for selection, FD overcomes the weak discrimination ability of those simple global descriptors. FD also overcomes the noise sensitivity and difficult normalization in the shape signature representations.

Most FD based works are dedicated to character recognition and object classification. The complex coordinates and the cumulative angle function are dominantly used in these works to derive FD. However, Zhang and Lu have found that for general shapes, the centroid distance function is the most desirable shape signature to derive FD. They have also found that 10 FD features are sufficient to represent shape, this is a significant reduction in dimensions of FD compared with 60 FD features usually used in shape representation. Their results show that FD outperforms CSS method in terms of retrieval performance and robustness.

XI. RESULTS

Figure lists the results obtained with the proposed method in a CAT image of the media stinum, in which a liver tumor can be observed. First the original image is presented, then image 5b) after applying the Alternating Sequential Filters with structuring elements of growing size. This allows to homogenize the areas to be detected without subjecting them to deformations. Image 5c) depicts different connected components marked in different colors. Finally, Figure 5 d) shows the correct tumor segmentation. The exact boundaries of the segmented tumor as well as its localization can be calculated in relation to it.

Figure depicts several CAT mediastinum images and their corresponding segmented images with the proposed method. Original images can be observed in the first column; while the second column shows the images resulting from the segmentation of the different tumors (lung, kidney and liver, respectively).

The third column depicts the contour extraction of said tumors. As seen in the figure, the proposed segmentation is genuinely satisfactory in all cases, having been tested in a group of more than 50 testing images.

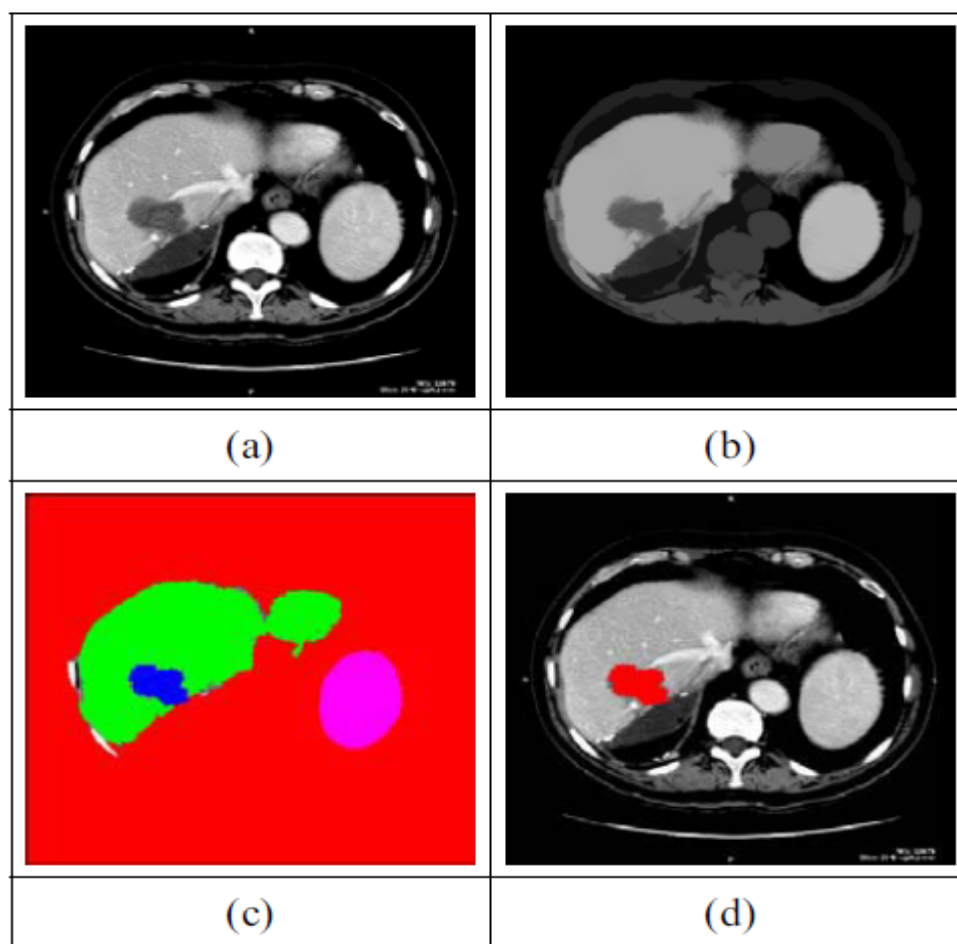


Fig.5 CATmedia stinum image. a) Original image b) Image obtained after applying sequential filters, c) Image with different connected components; and d) Image resulting from liver tumor segmentation.

Topology-based methods are of increasing importance in the analysis and visualization of dataset from a wide variety of scientific domains such as biology, physics, engineering, and medicine. Current challenges of topology-based techniques include the management of time-dependent data, the representation large and complex datasets,

the characterization of noise and uncertainty, the effective integration of numerical methods with robust combinatorial algorithms, etc.

While we see an increasing number of high-quality publications in this field, many fundamental questions remain unsolved. New focused efforts are needed in a variety of techniques ranging from the theoretical foundations of topological models, algorithmic issues related to the representation power of computer-based implementations as well as their computational efficiency, user interfaces for presentation of quantitative topological information, and the development of new techniques for systematic mapping of science problems in topological constructs that can be solved computationally.

XVIII. References

- [1] Levoy M, Pulli K, Curless B, Rusinkiewicz S, Koller D, Pereira L, Ginzton M, Anderson S, Davis J, Ginsberg J, Shade J, Fulk D (2000) "The digital Michelangelo project: 3D scanning of large statues." Proc. SIGGRAPH 2000, pp 131–144
- [2] Fontana R, Greco M, Materazzi M, Pampaloni E, Pezzati L, Rocchini C, Scopigno R (2002) "Three-dimensional modeling of statue" the Minerva of Arezzo. Cultural Heritage Journal 3(4):325–331
- [3] National design repository. <http://www.designrepository.org/>
- [4] Regli WC, Cicirello VA (2000) "Managing digital libraries for computer-aided design" J Computer Aided Des 32(2):119–132
- [5] ShapeSifter. <http://www.shapesearch.net/index.html>
- [6] Sung R, Rea HJ, Corney JR, Clark DER, Pritchard J, Breaks ML, Macleod RA (2002) ShapeSifter "a retrieval system for databases of 3D engineering data" New Rev Inf Netw 8:33–54
- [7] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242
- [8] Protein data bank. <http://www.rcsb.org/pdb/>
- [9] Shilane P, Kazhdan M, Min P, Funkhouser T (2004) "The princeton shape benchmark" Proc. shape modeling international 2004, pp 157–166
- [10] 3D model search engine. <http://shape.cs.princeton.edu>
- [11] 3D model similarity search engine. <http://merkur01.inf.uni-konstanz.de/CCCC>
- [12] 3D shape retrieval engine. <http://www.cs.uu.nl/centers/give/multimedia/3Drecog/3Dmatching.html>
- [13] Tangelder JWH, Veltkamp RC (2003) "Polyhedral model retrieval using weighted point sets" Int J Image Graphics 3(1):209–229
- [14] Engineering shape benchmark. <http://purdue.edu/shapelab>amani K (2005) "An engineering shape benchmark for 3D models" Proceedings of the ASME international design engineering technical conference & computers and information in engineering conference, pp 24–28