A Hierarchical Shape Representation for Vision-Guided Robotics

Begoña Martínez-Salvador **Angel P. del Pobil**

Computer Science Department, Jaume-I University
Campus Penyeta Roja, E12071 Castellón (Spain)
{bmartine, pobil}@inf.uji.es

Abstract

Using an adequate representation is often the key to solve complex problems in Artificial Intelligence. Hierarchical shape representations are very convenient in domains -such as vision-based robotics- that require a trade-off between efficiency and accuracy. In this paper, we present new techniques and results concerning a hierarchical shape representation based only on spheres. We focus on the global aspects of the model that are relevant to AI applications and present a summary of its main features and the employed procedures. Since the underlying object model is the generalized cylinder, the representation lends easily itself to its utilization in conjunction with a vision system.

Introduction

A suitable spatial representation is the key for solving a great number of problems in some domains, such as Artificial Intelligence and Robotics. A good approximation model should be simple enough to simplify the solution of the problem and at the same time, accurate in order not to loose information. There exist two ways of dealing with complex objects (Chazelle 1987): simplify them by means of computing an approximation or rewrite the objects as a combination of simpler parts.

Hierarchical representations are very useful in Robotics for path planning applications. They are a good spatial model in those applications where a trade between accuracy and simplicity is needed. Different approximations of the object with different levels of accuracy are generated and used when needed, so the application can work in real time.

Robot Motion planning has been studied for nearly two decades and many important contributions to the problem have been made. However, it has made few inroads into practical applications in the real world (Gupta and del Pobil 1998). A fact that partly accounts for this situation is that motion planning algorithms are often tested in simulations using simplistic geometric models. The examples reported in the literature usually involve simple geometries such as line segments. When current approaches are applied to real-world problems with complex geometric models (Chang 1995), relatively simple applications took 23 hours, and even in some cases almost 100 hours.

Another important point for practical applications of motion planning in the future (Gupta and del Pobil 1998) is that the input to the system cannot be assumed to be available as a CAD model –which is always the case in the literature– but rather should be based on computer vision.

The challenge is then to obtain very fast algorithms resulting in acceptable computation times for large complex domains with multiple moving objects. The efficiency of the algorithms is critically dependent on the representation that is used to model the robots and the environment. Bounding volumes have been used, typically axis-aligned boxes (Hayward 1986), (Lin et al. 1996) and, more recently, hierarchical spatial representations (del Pobil, Serna and Llovet 1992), (Lin et al. 1996), (Quinlan 1994).

Current approaches usually constrain objects to be described as the union of convex polytopes. If curved objects are modeled as polyhedra in a realistic way as is usual, the great number of involved features will make these methods inefficient, this is particularly important in the case of curved concavities, where no partition algorithm into convex parts can be used.

We have developed a hierarchical shape representation, where the only primitive used is the sphere. Starting with a single sphere that covers the whole object, the system can successively refine the representation by replacing spheres and always considering the shape of the object to find the best fitting set of spheres. The system can work in two different ways: the global mode and the local mode. The global mode is intended as an approximation of the global shape of the object. The local mode takes advantage of the system's capacity to automatically build new representations to improve locally the approximation in some parts of the object. In this paper, we are going to focus on the global shape representation as a general tool for AI applications; the local mode is more relevant to collision detection and is described in detail elsewhere (Gupta and del Pobil 1998) and (Martínez-Salvador, del Pobil and Pérez- Francisco 1998).

Our system is suitable for large and complex scenarios. The model aims at 3D objects described as a generalized cylinder (GC) or as a set of them. GCs are well-known in Computer Vision. They can describe a vast quantity of every-day objects. They are used for part-based recognition and, due to the invariant and quasi-invariant properties of their contours, they are often used for recovering 3D volumes with curved sides from 2D contours.

Recovering and representing the shape of complex objects is a main task in Computer Vision. A good model is useful not only for object recognition but also for manipulation, navigation and even learning (Zerroug and Nevatia 1996). Articulated objects can be naturally represented as a set of simpler parts and their relationships. Each part can be described by a volumetric primitive such a GC.

Related Work

O'Rourke and Badler (1979) described a method for obtaining a unique representation of an object as a set of overlapping spheres. More recently, Quinlan (1994) and Hubbard (1996) build hierarchies of spheres for objects. Both approaches build the hierarchy in a bottom-up fashion starting from a predefined lower level of spheres that cannot be improved. Comparative results of our approach in terms of quality and time have been provided with respect to the most closely related approach (Martínez-Salvador, del Pobil, Pérez-Francisco 1998).

Spherical representations have also been used for recognition of complex curved surfaces (Hebert et al. 1995) and for achieving a stable volumetric representation based on spheres (Ranjan and Fournier 1994).

The spatial representation we present is based on a initial representation developed by del Pobil et al. (1992). That model was limited to straight prisms having convex polygonal cross-sections. The system has been extended to approximate the shape of a vast quantity of objects described as GCs. It can approximate non-convex objects and curved surfaces with the desired accuracy. Our approach is intended as a shape representation.

The model has successfully been applied to efficiently solve the collision detection problem in motion planning (Gupta and del Pobil 1998) and (Pérez-Francisco, del Pobil and Martínez-Salvador 1998).

The Spherical Representation

We aim at obtaining a set of spherical approximations of an object described as a generalized cylinder. The set of approximations represent the shape of the object with different level of accuracy and tends to a zero-error model.

The class of solids whose shape we can approximate are those that can be described as a generalized cylinder or a set of them. A GC is a swept volume described by a *cross-section*, a *sweeping axis* (or *spine*), a *sweeping rule* and a *sweeping angle*. Our system can approximate GCs whose cross-section is any planar shape (*generalized polygon*) that can change its size along the spine according to the sweeping rule. The sweeping axis can be any curve in the plane or space. The angle between the cross-section plane and the tangent to the sweeping axis in a point is referred as the *sweeping angle*. The sweeping angle can be different of 90º but it remains constant for all the points in the spine.

Given a solid, we want to obtain a set of exterior spheres that completely covers its boundary surface. The problem we are concerned with is NP-hard (Meggido and Supowit 1984) even in the planar case, so the solution of the problem must have a great heuristic component.

In order to obtain a complete covering of the solid, a covering of the cross-section by a set of circles is obtained. The cross-section is swept along the spine and each circle gives rise to a circular generalized cylinder (CGC). Then, each CGC is covered by spheres.

To compute the 2D covering, we follow a top-down approach. First the whole cross-section is covered by a unique circle. Then, this circle is replaced by two new ones that completely cover the planar shape. In the successive steps, one circle is selected and replaced by two new ones.

Curved sides can be represented in a realistic way by using a great number of vertices without impairing the performance of the system thanks to the *efficient edge* heuristics.

The *efficient edge* heuristics proposes a qualitative vision of the shape that does not rely on the number of vertices used to represent the shape. It can be said that there is not a direct relation between the shape and the number of points. Two neighbor edges in a polygon are considered to belong to the same efficient edge if the angle between then is greater than a certain threshold angle (135º in our case). See (del Pobil and Serna 1995) for more details about the efficient edge heuristics in the general case.

The system internally groups the edges of the generalized polygon into efficient edges. The error set is made of those points in the representation that do not belong to the real object. The error set is divided into the efficient edges (del Pobil and Serna 1995). The *worst efficient edge* is that with greater error area associated. The set of efficient edges covered by the same circle make a *list*. Thus, each *list* is covered by a single circle.

The algorithm for covering the boundary of a planar shape can be outlined as follows:

- Group the boundary into efficient edges (number of efficient edges must be greater than 2).
- Compute the error surface corresponding to each efficient edge.
- Initially, there is only one *list*.
- Compute the smallest enclosing circle for the list.
- Repeat
  - Select the list that contains the worst efficient edge.
  - Divide the selected list into two sublists.
  - Compute the smallest enclosing circle for each sublist.
  - Compute the error areas for the edges of both sublists.

Few approaches deal with non-convex objects. In our system non-convex generalized polygons are treated as a whole, without using partitioning algorithms. Thus, this approach is useful to deal with curved concavities where no partitioning algorithm can be applied.

As it has been stated, the two-dimensional covering is swept and each circle gives rise to a CGC. These CGCs are even more general that the given definition of GCs since the *sweeping angle* might change along the spine. The procedure to cover a CGC by a set of spheres consists in dividing the axis of the CGC into *lists* and compute a sphere that covers the piece of CGC corresponding to each list. The algorithm is described in (Martínez-Salvador and del Pobil 1998).

Stability is the property of a representation such that changes in the data induce commensurate and predictable changes in the representation (Ranjan and Fournier 1994). In our case, the number of spheres in an instance of the representation is basically independent of the number of vertices, edges and faces in the underlying polyhedral model, as long as the shape is not significantly changed. This property allows curves and curved surfaces to be represented with as many details as needed without impairing the performance of the algorithm. Results that demonstrate the stability of the approach can be found in (Martínez-Salvador and del Pobil 1998).

The Expert Spherizer

To automatically build a hierarchy of representations we use a top-down approach which allows a better control over the resulting approximations. Starting with a first representation consisting of a single sphere, the system improves the actual representation by replacing the sphere by two new ones that better fit the shape of the object. The covering is *conservative*, that is, the new spheres will cover the same part of the object the old sphere did. The hierarchy of representations is *improvable* since given a representation it is always possible to obtain a new representation that better fits the shape of the object.

The model fulfills the *coverage criterion* which means that the set of spheres in a instance of the representation completely covers the object. Moreover, we try to find the best representation for a given number of spheres and to keep the number of spheres low.

The resulting hierarchy is *spatially balanced*, whereas most current approaches result only in a structurally balanced tree, as discussed by Xavier (1996). Spatial balance is hardly to implement efficiently, but it greatly improves the quality of the involved approximations and optimizes the ratio between the number of spheres and the resulting local precision.

A solid represented as a GC is composed of three different surfaces: the cross-sections at the initial and final positions of the sweeping axis and the side surface that arises when sweeping the cross-section. In order to fulfill the above mentioned properties, the complete covering of a solid consists of five different coverings: the covering of the cross-sections at the initial and final positions — *top* and *bottom* —, the covering of the side surface —*side*— and, in some cases, two portions of the side surface next to the top and the bottom — *upper* and *lower tip*, respectively.

As it was mentioned, all the points in a representation that belong to any of the spheres but not to the real object, belong to the error set. For the two-dimensional covering the error set is measured by ε, defined as the ratio of the error surface to the covered boundary of the cross section. In 3D, the error set has been divided into two subsets: the error set of the CGC respect to the real object and the error set of the spheres respect to the CGC. In both cases, the quality is measured as the ratio of the volume of the error set to the covered surface and the coefficients are denoted as ξ and δ, respectively. This concept can be extended to any of the partial coverings in a representation.

Given a certain representation characterized by a set of quality coefficients we want to obtain a new representation by modifying its predecessor by means of a certain action on it. The resulting representation must fulfill the desired properties: to be balanced, to keep low the number of spheres and to cover the whole object.

An heuristic system —called the *Expert Spherizer*— decides among all the possible actions which is the best action to obtain the new representation. This system has been implemented with the structure of a rule-based system. A complete description of the rules can be found in (Martínez-Salvador and del Pobil 1996).

Usually, the action consists in replacing a set of spheres by a new set that refines the representation. The *Expert Spherizer* compares the quality coefficient between the different partial coverings and the decides which set of spheres must be modified. Considering how the spherical representation is built, the set of spheres in a covering can be modified in two different ways: modifying the 2D covering —by replacing a circle by two new ones— or adding more spheres to cover the CGCs. The *Expert Spherizer* chooses between both actions by comparing the ξ and δ coefficients.

Moreover, the rules implement other actions to ensure a complete covering of the object and to keep the number of spheres low. When it cannot be ensured that the set of spheres of the side covers the cross-sections at the initial and final positions of the axis, the coverings for the top and bottom are defined. If the error of the spheres of the side covering over the top and/or bottom surfaces is worst than the error of the top and/or bottom coverings, then the chosen rule is to define a different covering for the upper tip and/or lower tip regions. These regions are covered in the same way as the side surface but using more circles in the 2D covering. Thus, the number of spheres is only increased in these parts and not for the whole object (see figures in next section).

Experimental Results

Figure 1 shows the spherical representation for a torus. This is a typical example of a non-convex object that cannot be partitioned into convex subparts. Our system does not apply any partitioning algorithm and the spherical representation is independent of the polygonal approximation of the torus.

The system can represent a vast class of GCs with curved surfaces. Figure 2 depicts the spherical representation of a GC having a non-planar curved axis. Figure 2(b) shows how the cross-section is swept along the axis. Figures 2(c) and 2(d) show different levels in the hierarchical representation. First, only one circle is used to cover the cross-section; consequently, there is only one CGC (figure 2(c)). When the representation is refined, two circles cover the cross-section; therefore, two CGCs are covered with spheres (figure 2(d)).

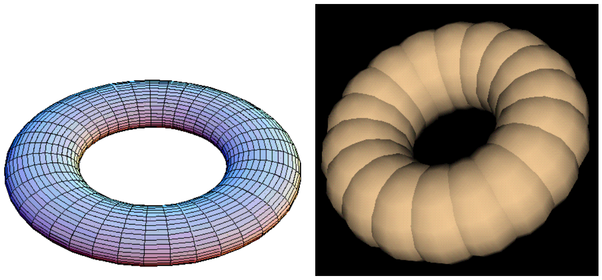
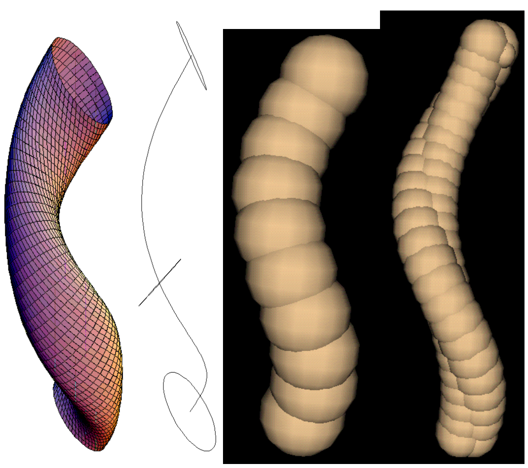


Figure 1 Spherical representation of a torus.

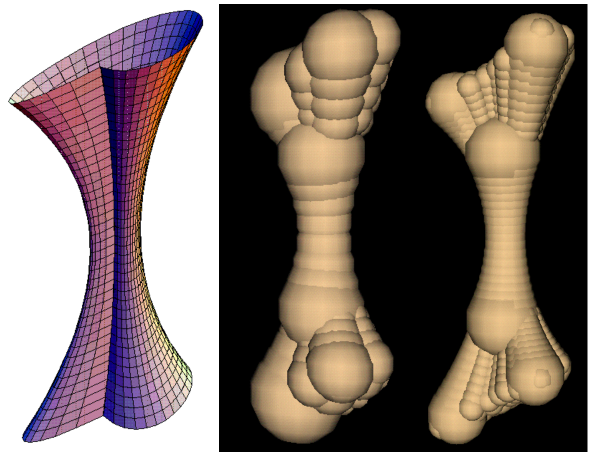
Figure 3 depicts a straight GC with curved non-convex cross-section, the *sweeping angle* is different from 90º (non-right GC) and the cross-section changes its size (decreasing and increasing). The spherical representation for this solid consists of five different coverings: the top and the bottom, the side and the upper and lower tips. The lower and upper tip coverings can clearly be observed in the figures. While for the side covering, the system is using just one CGC, for both tips the number of CGCs is increased to better fit the shape of the object in these parts.



(a) (b) (c) (d)

Figure 2 Representation of a general GC.

The hierarchy of representations tends to a zero-error model. Figure 4 plots the evolution of the δ coefficient —that measures the quality of each approximation— with the number of spheres for the object in Fig. 3. The quality of the approximations is improved — δ decreases very fast—in the first levels of the hierarchy by adding a few spheres. The fluctuations in the plot are due to some actions like making a new covering for the top or bottom or for the tip regions. These decisions are taken by the *Expert Spherizer* to ensure the complete covering of the solid and to keep the number of spheres low.



(a) (b) (c)

Figure 3 Representation of a GC with non-convex cross-section.

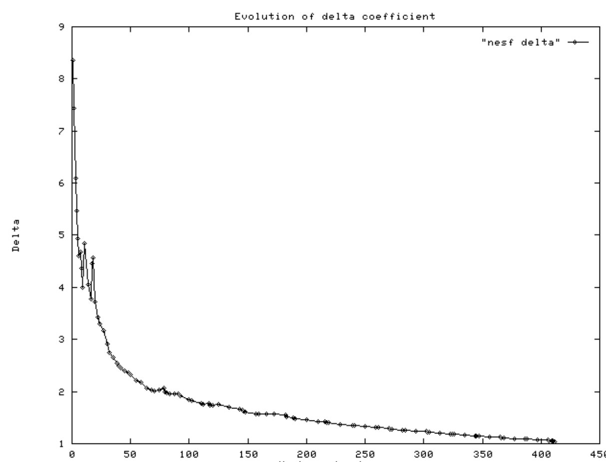


Figure 4 Quality of the hierarchy of approximations.

Contributions and Future Work

We have presented a system for representing general and complex objects described as generalized cylinders. The approach results in a conservative, quasi-optimal, stable and structurally balanced hierarchy of representations that tends to a zero-error approximation by using only spheres.

The model is suitable for dealing with real objects in complex scenarios. First, the use of the sphere as an unique primitive simplifies the solution of many problems. Second, the performance of the system is not impaired by the number of geometric features of the objects, therefore curved surfaces can be represented with as many details as desired and objects can be non-convex. Finally, the hierarchy of approximations allows a trade-off between accuracy and simplicity and it is always possible to obtain a representation as accurate as needed.

The system always considers the shape of the real object. It follows a top-down scheme which allows a better control on the quality of the approximations. Quality coefficients are defined for measuring the accuracy of each approximation with respect to the real object. A rule-based system called the *Expert Spherizer* decides which is the best action to improve the representation.

Non-convex objects can be represented without using partitioning algorithms. The system can handle a vast class of generalized cylinders. In fact, the approach presented is independent of the features of the generalized cylinder that represents the real object and can cover with spheres a more general class than GCs.

The model combines two volumetric primitives. It has the advantages of the simplicity of the sphere and the ability of representing a vast range of volumetric forms as GCs.

Since GCs are very well-suited for representing the parts of the human body, we are also working on real-time cooperation between robots and human models. We are considering environments where there is an interaction between persons and robots for which safety is a critical concern.

Acknowledgments

Support for this work is provided in part by the CICYT under project TAP95-0710, by the Generalitat Valenciana under project GV-2214/94, by Fundació Caixa-Castelló under P1A94-22 and by a scholarship of the FPU Program of the Spanish Department of Education and Science

References

Chang, H., 1995, "Motion Planning in Virtual prototyping: Practical considerations", *Proc. IEEE International Symposium on Assembly and Task Planning*, pp. 427-428.

Chazelle, B., 1987, "Approximation and decomposition of shapes". In Schwartz J. T. and Yap C.K. (eds) *Algorithmic and Geometric Aspects of Robotics*, chapter 4, pages 145-185. Lawrence Erlbaum.

del Pobil, A.P., Serna, M.A., Llovet, J., 1992, "A New Representation for Collision Avoidance and Detection", *IEEE Intl. Conf. on Robotics and Automation*, Nice, France, pp.246-251

del Pobil, A.P., Serna, M.A., 1995, *Spatial Representation and Motion Planning*, No. 1014 Lecture Notes in Computer Science, Springer-Verlag, Berlin.

Gupta, K., del Pobil A.P. (eds), 1998, *Practical Motion Planning in Robotics*, John Wiley & Sons.

Hayward, V., 1986, "Fast Collision Detection Scheme by Recursive Decomposition of A Manipulator Workspace", in IEEE International Conference on Robotics and Automation, pp.1044-1049.

Hebert, M., Ikeuchi, K. and Delingette, H., 1995, "A Spherical Representation for Recognition of free-form surfaces", IEEE Trans. on Pattern Analysis and Machine Intelligence 7(7), pages: 681-689.

Hubbard P.M., 1996. "Approximating polyhedral with spheres for time-critical collision detection". *ACM Transactions on Graphics* 15(3).

Lin, M.C., Manosha, D., Cohen, J. and Gottschalk, S., 1996. "Collision detection: algorithms and applications" in 2nd Workshop on Algorithmic Foundations of Robotics, Toulouse, France.

Martínez-Salvador, B., del Pobil, A.P., 1996, "A Hierarchy of Detail for Representing Non-Convex Curved Objects", Proc. Of the 9th. Intl. Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, Fukuoka, Japan, 1996.

Martínez-Salvador, B., del Pobil, A.P. and Pérez-Francisco, M., 1998, "Very Fast Collision Detection for Practical Motion Planning. Part I: The Spatial Representation", in Proc. *IEEE Intl. Conference on Robotics and Automation*, Leuven, Belgium.

Meggido N., Supowit L.J., 1984. "On the Complexity of some Geometric Location Problems", S*IAM J. Computing* 13(1), pp. 182-196.

O'Rourke, J. and Badler, N., 1979. "Decomposition of 3d Objects into Spheres", *IEEE Trans. on Pattern Analysis and Machine Intelligence*.

Pérez-Francisco, M., del Pobil, A.P., Martínez-Salvador, B., 1998, "Very Fast Collision Detection for Practical Motion Planning. Part II: The Parallel Algorithm", in Proc. *IEEE Intl. Conference on Robotics and Automation*, Leuven, Belgium.

Quinlan, S., 1994, "Efficient distance computation between non convex objects", *Proc. IEEE International Conference on Robotics and Automation*, pp. 3324-3329.

Ranjan, V., Fournier, A., 1994, "Volume Models for Volumetric Data", *IEEE Computer*, Vol. 27, No. 7, pp. 28-36.

Xavier, P.G., "A Generic Algorithm for Constructing Hierarchical Representations of Geometric Objects", *Proc. 1996 IEEE Intl. Conf. On Robotics and Automation*, Minneapolis, Minnesota.

Zerroug, M. and Nevatia, R., 1996. "From an Intensity Image to 3D Segmented Descriptions", in J. Ponce, A. Zisserman and M. Hebert eds., *Object Representation in Computer Vision II*, *Lecture Notes in Computer Science*, no. 1144, Springer-Verlag.