Topological Active Volumes

N. Barreira

Grupo de Visión Artificial y Reconocimiento de Patrones (VARPA), LFCIA, Departamento de Computación, Universidad de la Coruña, 15071 A Coruña, Spain
Email: noelia@dc.fi.udc.es

M. G. Penedo

Grupo de Visión Artificial y Reconocimiento de Patrones (VARPA), LFCIA, Departamento de Computación, Universidad de la Coruña, 15071 A Coruña, Spain
Email: cipenedo@dc.fi.udc.es

Received 29 December 2003; Revised 4 February 2005

The topological active volumes (TAVs) model is a general model for 3D image segmentation. It is based on deformable models and integrates features of region-based and boundary-based segmentation techniques. Besides segmentation, it can also be used for surface reconstruction and topological analysis of the inside of detected objects. The TAV structure is flexible and allows topological changes in order to improve the adjustment to object's local characteristics, find several objects in the scene, and identify and delimit holes in detected structures. This paper describes the main features of the TAV model and shows its ability to segment volumes in an automated manner.

Keywords and phrases: image segmentation, 3D reconstruction, active nets, active volumes.

1. INTRODUCTION

The scene in a 3D image is usually made of an object or several objects. The segmentation task consists of isolating the points that belong to each object from the image and integrating them into a coherent and consistent model of the detected structures. Unfortunately, the detection of objects inside 3D data is difficult due to the complex topology of the objects and the cost of the operations in a 3D space. Furthermore, the available means of acquisition introduce noise in data, so direct segmentation is often inadequate and the use of an advanced segmentation technique is needed.

There are many approaches for segmentation proposed in the literature. Nowadays, level set methods and deformable models have become two of the most promising techniques for segmentation, mapping, tracking, and modelling tasks. Level set methods were introduced by Osher and Sethian [1] and have found several applications in image processing, such as segmentation and reconstruction of complex shapes [2, 3] due to their ability to perform topological transformations. However, level sets’ formulation is more complex than deformable models’. Deformable models were introduced by Kass et al. [4] in 2D as explicit deformable contours and generalised to the 3D case by Terzopoulos et al. [5]. In recent years, many models have been developed for the treatment of 3D scenes [6, 7, 8]. They have tried to solve some of the limitations of classical deformable models, such as the sensitivity to the initialisation or the parametric definition of the model, which restricts their topology to segment simple objects. These models have also adapted their operation to concrete domains by means of the definition of different energy terms or the use of several topologies. In particular, some models have increased their degree of flexibility, providing a way to perform topological transformations on their structures [9, 10]. However, all these models are mainly interested in surface extraction and shape recovery, not in the treatment of the whole detected solid.

In this paper a general 3D deformable model is proposed, the topological active volumes, as an extension of the topological active nets [11], focused on segmentation tasks by means of a volumetric distribution of nodes. It is fully automatic, so no a priori knowledge is needed to initialise the model. It also integrates information of edges and regions in the adjustment process and allows to obtain topological information inside the objects found. The model has a dynamic behaviour by means of topological changes in its structure which enables accurate adjustments and the detection of several objects in the scene.

This paper is organised as follows. Section 2 describes the model and the mechanisms that govern its behaviour. Section 3 explains the methodology used in the segmentation process. Section 4 presents several examples of image segmentation and reconstruction, and finally the conclusions are exposed in Section 5.
2. TOPOLOGICAL ACTIVE VOLUMES

The model presented in this paper is an extension of topological active nets [11] to the 3D world. Its operation is focused on extraction, modelization, and reconstruction of volumetric objects present in the scene.

A topological active volume (TAV) is a 3D structure composed by interrelated nodes, where the basic repeated structure is a cube (Figure 1). Parametrically, a TAV is defined as \( v(r,s,t) = (x(r,s,t), y(r,s,t), z(r,s,t)) \), where \( (r,s,t) \in ([0,1] \times [0,1] \times [0,1]) \). The state of the model is governed by an energy function defined as follows:

\[
E(v) = \int_0^1 \int_0^1 \int_0^1 E_{\text{int}}(v(r,s,t)) + E_{\text{ext}}(v(r,s,t)) \, dr \, ds \, dt,
\]

where \( E_{\text{int}} \) and \( E_{\text{ext}} \) are the internal and the external energy of the TAV, respectively. The former controls the shape and the structure of the net. Its calculus depends on first- and second-order derivatives which control contraction and bending, respectively. The internal energy term is defined by

\[
E_{\text{int}}(v(r,s,t)) = \alpha\left( |v_r(r,s,t)|^2 + |v_s(r,s,t)|^2 + |v_t(r,s,t)|^2 \right) + \beta\left( |v_{rr}(r,s,t)|^2 + |v_{ss}(r,s,t)|^2 + |v_{tt}(r,s,t)|^2 \right) + 2\gamma\left( |v_{rs}(r,s,t)|^2 + |v_{rt}(r,s,t)|^2 + |v_{ts}(r,s,t)|^2 \right),
\]

where subscripts represent partial derivatives and \( \alpha, \beta, \) and \( \gamma \) are coefficients that control the smoothness of the net. In order to compute the energy, the parameter domain \([0,1] \times [0,1] \times [0,1]\) is discretized as a regular grid defined by the internode spacing \((k,l,m)\) and the first and second derivatives are estimated using the finite differences technique in 3D.

\( E_{\text{ext}} \) represents the characteristics of the scene that guide the adjustment process. As can be seen in Figure 1, the model has two types of nodes: internal and external. Each type of node is used to represent different features of the object: the external nodes fit the surface of the object and the internal nodes model the internal topology of the object. So the external energy has to be different for both types of nodes. This fact allows the integration of information based on discontinuities and information based on regions. The former is associated to external nodes and the latter, to internal nodes. In this model, the external energy term is defined as

\[
E_{\text{ext}}(v(r,s,t)) = \omega f[I(v(r,s,t))] + \rho \frac{1}{|N(r,s,t)|} \sum_{p \in N(r,s,t)} \frac{1}{|v(r,s,t) - v(p)|} f[I(v(p))],
\]

where \( \omega \) and \( \rho \) are weights, \( I(v(r,s,t)) \) is the intensity value of the original image in the position \( v(r,s,t) \), \( \eta(v(r,s,t)) \) is the distance to the nearest point on the object surface, \( I(v) \) is a function associated to the image intensity and defined differently for both types of nodes.

On one hand, if the objects to detect are dark and the background is bright, the energy of an internal node will be minimum when it is on a point with a low grey level. On the other hand, the energy of an external node will be minimum when it is on a discontinuity and on a light point outside the object. In this situation, function \( f \) is defined as

\[
f[I(v)] = \begin{cases} 
\frac{\eta[I(v) - I]\max}{\max[I] - \min[I]} & \text{for internal nodes,} \\
\frac{\eta[I(v)] - \min[I]}{\max[I] - \min[I]} + \xi(\max[I] - \min[I]) & \text{for external nodes,}
\end{cases}
\]

where \( \xi \) is a weighting term, \( I_{\max} \) and \( I_{\min} \) are the maximum intensity values of image \( I \) and the gradient image \( G \), respectively, \( I(v) \) and \( G(v) \) are the intensity values of the original image and the gradient image in node position \( v(r,s,t) \), \( \overline{I(v)} \) is the mean intensity in a \( n \times n \times n \) cube, and \( h \) is an appropriate scaling function.

Otherwise, if the objects are bright and the background is dark, the energy of an internal node will be minimum when it is on a point with a high grey level and the energy of an external node will be minimum when it is on a discontinuity and on a dark point outside the object. In such a case, function \( f \) is defined as

\[
f[I(v)] = \begin{cases} 
\frac{\eta[I_{\min} - I]}{\max[I] - \min[I]} & \text{for internal nodes,} \\
\frac{\eta[I(v)] - \min[I]}{\max[I] - \min[I]} + \xi(\max[I] - \min[I]) & \text{for external nodes,}
\end{cases}
\]

where the symbols have the same meaning as in (4).
3. METHODOLOGY

The adjustment process of the TAV has several stages, as shown in Figure 2.

The first stage consists of placing the 3D structure in the image. The nodes cover the whole image and they are located in such a way that the internodal distance is the same in each dimension. This way, the model is able to detect an object in the image even if it were placed at different positions.

The energy minimisation is performed locally using a greedy algorithm. With this algorithm, the energy value for each node is computed in several positions of the 3D image (the current position and its 26 neighbour positions) at every step of the minimisation process and the best one is chosen as the next position of the node. The process finishes when the TAV reaches a stable situation, that is, when the energy of each node in the TAV is minimal.

Once the mesh reaches a stable situation, the number of nodes is recomputed in order to adjust the size of the TAV to the size of the object, that is, if the object is wider than high, the number of nodes in the x-axis will be larger than the number of nodes in the y-axis. After that, the mesh, which covered the whole image in the beginning, is centred around the object. The readjustment of the mesh allows the acquisition of the same distribution of nodes independently of the position of the object in the image [12]. Then the minimisation process is repeated.

The physical characteristics of the mesh do not enable the perfect adaptation of the nodes to the objects, so some kind of topological transformation on the TAVs are necessary to achieve a good adjustment. This way, the restrictions of a fixed topology can be avoided. The first topological change consists of the rupture of connections between external nodes which are wrongly placed in the image. This allows a perfect adjustment to the surfaces of the objects, the detection of several objects in the image, and the generation of external holes in the mesh. The second topological change is the creation of holes in the inside of the mesh in order to delimit the holes of the objects. The analysis of the internal nodes and the rupture of connections allow the second type of topological changes.

3.1. Rupture of connections

The TAVs have the ability to perform topological changes that give more flexibility to their structure and avoid the limitations of a fixed topology as we can see in Figure 3. The changes consist of removing connections between nodes. They allow to (a) increase the fitting accuracy on those zones where a greater definition is required, (b) detect two or more objects in the image through the generation of 2 or more sub-TAVs, and (c) model the found objects and adjust to external holes.

The process of breaking a connection takes place only between external nodes that are badly located after the minimisation process. This removal will affect the elementary structures, that is, the cubes the nodes belong to. Once the connection between two external nodes has disappeared, the nodes have a greater freedom of movement. Besides, the rupture of a link implies that some internal nodes will become external as Figure 4 shows. The new external nodes will allow a greater definition and adjustment in the area where the breaking takes place.

The breaking process has one restriction: the cubic structure of the model must be maintained (all the nodes and connections must belong to, at least, one cube) because TAV model performs, not only surface extraction, but also volume segmentation and so, isolated nodes, connections, or planes do not provide useful information. This is why connections will not be erased if the cubic structure is broken.

The breaking restriction can be avoided in many cases by performing several ruptures instead of one. When the algorithm finds a connection that cannot be broken, the cubes...
that share the connection are checked in order to find a combination of connection breakings that maintain the cubic structure. Nevertheless, a connection will only be broken if its side nodes are badly placed. Figure 5 depicts two cases of this particular situation. In Figure 5a, if connection $a$ is broken, connection $b$ will not belong to any cube as there are no cubes over it. Connection $a$ must be broken together with connection $b$. Notice that connection $b$ can be broken without connection $a$ because connection $a$ is part of another cube. Figure 5b shows an example where the connections $a$, $b$, $c$, and $d$ must be broken simultaneously in order to maintain the cubic structure.

After the minimisation process, the external nodes badly located are chosen by their gradient distance. Gradient distance is computed through an extension of midpoint algorithm [13] in 3D. A node will be badly located if its gradient distance is greater than a certain threshold, which is computed by Tchebycheff’s theorem [14]. This theorem identifies the outliers within a population. Thus, an external node $v_{\text{ext}}(r, s, t)$ will be badly located if its gradient distance, $\text{GD}_{v_{\text{ext}}}(r, s, t)$, verifies that

$$\text{GD}_{v_{\text{ext}}}(r, s, t) > \mu \text{GD}_{v_{\text{ext}}} + c \sigma \text{GD}_{v_{\text{ext}}},$$

where $\mu \text{GD}_{v_{\text{ext}}}$ and $\sigma \text{GD}_{v_{\text{ext}}}$ represent the average and the standard deviation of the gradient distance of the external nodes, respectively. $c$ is a constant so that $1/c^2$ is the percentage of correctly placed nodes. We have used a value of $c = 2$.

The gradient distance is computed as follows:

$$\text{GD}_{v_{\text{ext}}}(r, s, t) = \min \text{GD}^k_{v_{\text{ext}}}(r, s, t) \quad \forall k \in \kappa,$$

where $k$ is the set of possible directions defined by a 26-neighbourhood.

Once the external nodes badly placed are delimited, the connection to break is selected. It is the connection between the worst placed node and its worst neighbour. After breaking a connection and performing the topological changes, the TAV is minimised again. This process is repeated until there is no node which fulfils (6).

In order to guide the breaking process and avoid anarchical breakings, a cut priority is used. The cut priority of each node is initially 1 and is doubled every time a breaking takes place in the cube the node belongs to. The gradient distance is multiplied by the cut priority and the result is used instead of the gradient distance value in (6). This way, the gradient distance is weighted, which increases the probability that new breakings will be performed in the direction of previous breakings.

Figures 6a and 6b show the next breaking after the rupture of a connection in the upper cube. In these figures, the external nodes are labelled with the gradient distance and the gradient distance weighted by the cut priority, respectively. Without cut priority, the node with greater gradient distance and its worst placed neighbour share the connection to break, as Figure 6a shows. However, if cut priority is considered, after the first breaking, the nodes in the upper cube double its initial priority to 2 which weights the gradient distance, as Figure 6b shows. In this case, the next connection to break is in the cube where the first breaking has taken place.

Figures 6c and 6d depict the consequences of anarchical and organised breakings, that is, without or with cut priority. If the cut priority is not considered, the breakings will be chaotic and there could be groups of linked cubes badly placed whose connections cannot be broken (threads). When the cut priority is used, the probability of thread formation diminishes.

Advantages of the topological changes are shown in Figure 7. Figure 7a shows several steps in the detection of several objects in the image by means of a subdivision process, generating as many TAVs as target objects. Figure 7b shows how the model can detect and adjust correctly to holes, obtaining a good representation of their internal structure.
3.2. Hole creation

The topological information provided by the distribution of the internal nodes allows us to know the inner morphology of detected objects. This information can be used to detect inner holes in the objects.

If there is a hole in an object, the internal nodes will try to avoid it and there will be tension between them, so the value of the internal energy will be high (see (2)). The internal nodes will also have a high value of external energy due to the characteristics of the hole. For example, if the object is dark and the background is light, the holes in the object will be light, too. In this case, the energy value of an internal node is minimum when the node is in a dark voxel (see (4)). But the holes are light and then, if an internal node is inside a hole, its energy value will not be minimum.

Tchebycheff’s theorem [14] is also employed to find the internal nodes inside a hole. An internal node is supposed to be inside a hole if its energy value, $E_{v\text{in}}(r,s,t)$, verifies that

$$E_{v\text{in}}(r,s,t) > \mu E_{v\text{in}} + 3\sigma E_{v\text{in}},$$

where $\mu E_{v\text{in}}$ and $\sigma E_{v\text{in}}(r,s,t)$ represent the average and the standard deviation of the total energy, that is, the sum of the internal and external energies of the internal nodes.

Once the set of internal nodes inside the hole is known, the hole is created from a cube of internal nodes: one of them is the worst placed and the other ones are also badly placed nodes.

The hole must be composed of external nodes as they must be anchored to the surface of the hole. However, the aim of the external nodes is contracting around the object whereas the nodes inside a hole have to expand themselves towards the surfaces of the hole. So, there will be three kinds of nodes: internal nodes, external nodes outside the hole, and external nodes inside the hole. Each kind of node has its own external energy function, which replaces the function defined in (4). The new energy function is defined as

$$f[I(v)] = \begin{cases} h[I(v)_n] & \text{for internal nodes,} \\ h[I_{\text{max}} - I(v)_n] + \xi (G_{\text{max}} - G(v)) & \text{for external nodes,} \\ h[I_{\text{max}} - I(v)_n] + GD_{\text{min}}(v) & \text{for hole nodes,} \end{cases}$$

where $GD_{\text{min}}(v)$ represents the minimum gradient distance of the internal node $v(r,s,t)$.
Once the holes are created in the mesh, a process of rupture of connections and energy minimisations is performed as it has been explained in Section 3.1 but, in this case, the breakings would take place between external nodes inside the hole.

Figure 8 shows some examples of images with inner holes and the results of segmentation process. These 3D images are cylinders with an inner hole as shown in the 3D diagram in Figure 8a. Figure 8b shows a cross-section of the 3D image defined by the dotted lines in the image of Figure 8a. The point of view of these sections is shown by the arrow in Figure 8a. In these images, the internal adaptation of the TAV to the surface of the holes can be seen. Figures 8a and 8b show the 3D reconstruction of the external surface and the hole, respectively. In order to show the adjustment of the mesh, the reconstruction is based on the coordinates of the external nodes and no smooth technique was used to enhance the results.

4. RESULTS

This section shows the results obtained on artificial noisy images and CT medical images. A 3D image was obtained from a set of 2D CT images. In all the examples, the same image was used as external energy for both internal and external nodes. The parameters were empirically selected.
The model has been tested in images with different amounts of Gaussian noise. The parameters used in the segmentation of those images are shown in Table 2. Figure 10 shows the results of the segmentation process on a sample image. $\alpha$ controls the contraction of the mesh. When the noise is high, the $\alpha$ parameter increases its value in order to force the contraction of the mesh and avoid that the nodes be anchored to noise. $\omega$ is the parameter that weights the contribution of the current position of the node in the image. In noisy images, $\omega$ has to be reduced because there is a high probability that the node could be located over noise. On the other hand, the parameter that weights the contribution of the positions of the neighbouring nodes in the image, $\rho$, has to be increased since the analysis of several voxels in the image instead of only one reduces the probability of the noise influence. Finally, the contribution of the gradient, $\xi$, depends on the quality of the edge detector. The Canny filter, used in the noisy examples, produces very good results so $\xi$ can have a high value.

The model has been also tested with several sets of CT images. In this case, the Sobel filter was employed to obtain the gradient images.

The first example consists of 297 femur CT images with 255 grey levels. The parameters used were $\alpha = 4$, $\beta = 0.00001$, $\gamma = 0.00001$, $\omega = 4$, $\rho = 4$, and $\xi = 5$. The initial TAV had $25 \times 25 \times 8$ nodes and was readjusted to $18 \times 15 \times 21$ nodes. Figure 11 shows the results of the segmentation process.

The second example is a set of 342 tibia and fibula CT images with 255 grey levels. The parameters used were $\alpha = 2$, $\beta = 0.00001$, $\gamma = 0.00001$, $\omega = 5$, $\rho = 4$, and $\xi = 5$. The last row in Figure 9 represents the 3D reconstruction of the object from the TAV obtained after the segmentation process. It is performed from the cubes using the external nodes.

The model has been tested in images with different amounts of Gaussian noise. The parameters used in the segmentation process of the noisy examples, produces very good results so $\xi$ can have a high value.
5. CONCLUSIONS

This paper presents a new deformable model focused on segmentation and reconstruction tasks. The model consists of a volumetric structure and has the ability to integrate information based on discontinuities and regions. The model also allows the detection of two or more objects in the image and achieves a good adjustment to the surfaces and holes of the objects. This is due to the distinction of three classes of nodes: internal, external, and hole nodes. A complementary term of energy is assigned to each kind of node, which allows a different behaviour of internal, external, and hole nodes in analogous situations.

The model is fully automatic and does not need a user initialisation process like other deformable models. Once the TAV fits the object, the connections between the external nodes allow the definition of the surface of the object and its representation using any reconstruction method. On the other hand, the internal nodes show the spatial distribution inside the object that allows a topological analysis and helps find the inner holes of the objects.

The model was tested with artificial and CT images and a good adjustment to the objects was achieved. The model also gets good results with noisy images. A readjustment of the parameters is only necessary to obtain a better adaptation of the mesh to the objects in the image.

Future work includes the use of new basic structures in the mesh like triangular pyramids, and the introduction of graphical principles in nodes’ behaviour to obtain a better representation of the surfaces of the objects.

ACKNOWLEDGMENT

This paper has been partly funded by the Xunta de Galicia and the Ministerio de Ciencia y Tecnología through the Grant contracts PGIDIT03TIC0503PR and TIC2003-04649-C02-01, respectively.

REFERENCES


N. Barreira obtained the Bachelors degree in computer science in 2003 from the University of Coruña, Spain. She is currently a second-year Ph.D. student in the Computer Science Department, the University of Coruña. Her research interests lie in the fields of computer vision, pattern recognition, and biomedical image processing.

M. G. Penedo received the B.S. degree and the Ph.D. degree in physics from the University of Santiago de Compostela, Spain, in 1990 and 1997, respectively. From 1991 to 1999, he was an Associate Professor in the Computer Science Department, University of Coruña, Spain. Since 1999, he has been a Lecturer at the same department. His current research interests include computer vision, pattern recognition, biomedical image processing, and perceptual grouping.
EUSIPCO 2011
19th European Signal Processing Conference
August 29- September 2, 2011, Barcelona (Spain)

Preliminary call for papers
The 2011 European Signal Processing Conference (EUSIPCO-2011) is the nineteenth in a series of conferences promoted by the European Association for Signal Processing (EURASIP, www.eurasip.org). This year edition will take place in Barcelona, capital city of Catalonia (Spain), and will be jointly organized by the Centre Tecnològic de Telecomunicacions de Catalunya (CTTC) and the Universitat Politècnica de Catalunya (UPC).

EUSIPCO-2011 will focus on key aspects of signal processing theory and applications as listed below. Acceptance of submissions will be based on quality, relevance and originality. Accepted papers will be published in the EUSIPCO proceedings and presented during the conference. Paper submissions, proposals for tutorials and proposals for special sessions are invited in, but not limited to, the following areas of interest.

Areas of Interest
- Audio and electro-acoustics
- Design, implementation, and applications of signal processing systems
- Multimedia signal processing and coding
- Image and multidimensional signal processing
- Signal detection and estimation
- Sensor array and multi-channel signal processing
- Sensor fusion in networked systems
- Signal processing for communications
- Medical imaging and image analysis
- Non-stationary, non-linear and non-Gaussian signal processing

Submissions
Procedures to submit a paper and proposals for special sessions and tutorials will be detailed at www.eusipco2011.org. Submitted papers must be camera-ready, no more than 5 pages long, and conforming to the standard specified on the EUSIPCO 2011 web site. First authors who are registered students can participate in the best student paper competition.

Important Deadlines:

- Proposals for special sessions: 15 Dec 2010
- Proposals for tutorials: 18 Feb 2011
- Electronic submission of full papers: 21 Feb 2011
- Notification of acceptance: 23 May 2011
- Submission of camera-ready papers: 6 Jun 2011

Webpage: www.eusipco2011.org