

From the Zones of Influence of Skeleton Branch Points to Meaningful Object Parts

L. Serino, C. Arcelli, and G. Sanniti di Baja

Institute of Cybernetics “E. Caianiello”, CNR, Naples, Italy
{l.serino,c.arcelli,g.sannitidibaja}@cib.na.cnr.it

Abstract. A 3D object decomposition method is presented, which is based on the decomposition of the linear skeleton guided by the zones of influence. These are the connected components of voxels obtained by applying the reverse distance transformation to the branch points of the skeleton. Their role is to group sufficiently close branch points and to detect perceptually meaningful skeleton branches that are in a one-to-one relation with the object parts.

1 Introduction

According to the structural approach to shape analysis, an object can be interpreted as constituted by a number of perceptually meaningful parts and its description can be given in terms of the description of the various parts and of their spatial relationships. This approach has been inspired by the behavior of the human visual system, as discussed in [1-4]. One of the advantages of such a structured representation is a greater robustness under changes in viewing conditions.

The skeleton is a tool often employed to achieve a structural analysis of the object it represents [5-12]. In fact, the skeleton is a linear subset of the object reflecting the topological and geometrical features of the object and such that each skeleton branch is in correspondence with one of the parts understood as constituting the object. Thus, a decomposition of an object into its constituting parts can be guided by a decomposition of the skeleton into its constituting branches.

Parts associated with skeleton branches meeting in common points, the branch points of the skeleton, overlap with each other. If decomposition into disjoint parts is preferred, care is necessary to deal with the overlapping regions.

We have suggested shape decomposition methods guided by skeleton decomposition for both 2D and 3D objects. In the 2D case, we decomposed the skeleton into its constituting branches and obtained object decomposition into partially overlapping parts [5,6]. In the 3D case, we favored decomposition into disjoint parts and to this purpose we suggested a suitable partition of the skeleton [10,11]. In particular in [11], we partitioned the skeleton into three types of components, respectively called complex sets, simple curves and single points, which correspond to three types of object parts, respectively called kernels, simple regions and bumps. Simple regions and bumps protrude from the kernels. In turn, kernels can

be interpreted as sort of main bodies of the object. Kernels were identified in correspondence with the positions where different skeleton branches meet.

In this paper, we work with 3D objects and propose a decomposition method that is inspired by our work in 2D as concerns the decomposition of the skeleton into its constituting branches, and shares with our 3D method the fact that the object is decomposed into disjoint parts. The current method will face the problem of identifying skeleton branches corresponding to meaningful object parts, and will deal with the assignment of each overlapping region to only one of the proper object parts overlapping each other.

The object parts obtained as described in this paper are somehow analogous to the simple regions and bumps of the decomposition method in [11]. In both cases, object decomposition is into disjoint parts. An important difference between the two methods is that with the current method no region merging is necessary to achieve an object decomposition in accordance with human intuition. Other analogies and differences will be discussed in Section 4.

2 Preliminaries

We consider objects rid of cavities in binary voxel images in cubic grids. The 26-connectedness is used for the object and the 6-connectedness for the background. The neighbors of a voxel p are the 26 voxels sharing with p a face, an edge, or a vertex.

The distance between two voxels p and q is defined as the length of a minimal discrete path linking p to q . The three integer weights $w_f=3$, $w_e=4$ and $w_v=5$ are used to measure the unit moves from a voxel towards its face-, edge- and vertex-neighbors along the path, respectively. This choice of weights is motivated by the fact that the so obtained $\langle 3,4,5 \rangle$ weighted distance provides a reasonably good approximation to the Euclidean distance [13].

According to the model proposed by Blum [14], the skeleton of an object is a subset of the object consisting of points symmetrically placed within the object, having the same topology of the object, and such that each skeleton point is associated with its distance from the background. The value of a skeleton point can be interpreted as the radius of a ball that, centered on the point, is bi-tangent to the object's boundary and is included in the object. The object can be recovered starting from its skeleton by computing the envelope of the balls associated to its points.

For 3D objects, the above model originates a surface skeleton, which consists of the union of surfaces and curves. The surface skeleton of objects rid of cavities can be furthermore reduced to originate a skeleton exclusively consisting of curves. Only partial object recovery from such a skeleton is possible, unless the object consists of parts with tubular shape. In fact, only in such a case the symmetry points are mostly aligned along symmetry axes, while in the general case symmetry points are placed along symmetry planes and axes.

In the digital space, the skeleton of a 3D object can be computed according to the model of Blum by identifying the centers of maximal balls in the distance transform of the object. The distance transform DT is a multi-valued replica of the object, where

each voxel is labeled with its distance from the background. Thus, each voxel in DT can be interpreted as the center of a ball with radius equal to the corresponding distance label. In particular, a voxel whose associated ball is not included by any other single ball in the object is called center of maximal ball CMB. The object can be recovered by the union of the balls associated to the CMBs of the object. The ball associated to any distance labeled voxel p can be obtained by applying to p the reverse distance transformation [15].

In this paper, we compute DT by using the $\langle 3,4,5 \rangle$ weighted distance. There, to establish whether a voxel p is a CMB it is enough to compare the distance label of p with the distance labels of its 26 neighbors, by suitably taking into account the weights w_f , w_e and w_v [16]. As for the skeleton used to guide object decomposition, we refer to the linear skeleton obtained by the algorithm for DT based skeletonization suggested in [17]. We are aware that when using a linear skeleton a difference generally exists between the input object and the union of the only balls associated with the voxels of its skeleton. Thus, our decomposition method is completely effective only in case of objects that are perceived as consisting of parts with tubular shape, since in this case the above difference is negligible. This is the object domain considered in the following.

3 The Method

Let S be the skeleton of the object at hand. A voxel p of S is an end point when it has only one neighbor in S , is a normal point when it has two neighbors in S , and is a branch point when it has more than two neighbors in S . A skeleton branch is a curve of S entirely consisting of normal points, except for the two extremes of the curve that are end points or branch points.

Balls associated to a set of distance labeled skeleton voxels by the reverse distance transformation may overlap and merge into connected components. Let us consider the balls associated with all the branch points of S . Each group of these balls forming a connected component is called *zone of influence* of the branch points it includes. Branch points that are neighbors of each other or are closer to each other than the sum of their associated distance labels are included in the same zone of influence.

Intersecting object parts are mapped into branches of an ideal continuous skeleton that meet in a single branch point. In turn, more than one single branch point is generally found in the digital skeleton S in correspondence with intersecting object parts. However, in any such a case the branch points are likely to be very close to each other, so that they are included in a unique zone of influence. Thus, the zones of influence can be used to group branch points of S actually corresponding to a single branch point configuration of the skeleton that would ideally represent the object at hand. Obviously, the number of zones of influence may be smaller than the number of branch points of S .

In the following, the zones of influence are used to correctly identify the configurations where skeleton branches meet with each other. They are also used to count the number of perceptually meaningful branches of the skeleton and, hence, the

number of object decomposition parts. To this aim, connected component labeling is accomplished on the zones of influence, as well as on the set that is obtained from S by removing from it all voxels that result to be included in any zone of influence.

Short branches may exist in S , whose voxels are all included in a zone of influence, e.g., see Fig. 1. These short branches do not correspond to meaningful object parts and, hence, should not be counted. This is guaranteed since connected component labeling of S is done after removal of the skeleton voxels placed in zones of influence.

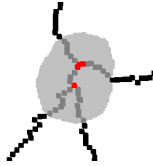


Fig. 1. A zone of influence (gray) including a short skeleton branch

3.1 Detecting Skeleton Components

To illustrate our object decomposition method, let us refer to the object “horse”, shown in Fig. 2 top left, which will be used as running example. The skeleton S computed by the algorithm [17] is shown in Fig. 2 top right, where red, black and green are used for branch points, normal points and end points, respectively. We note that S includes six branch points, six end points, and nine skeleton branches. The two zones of influence resulting after applying the reverse distance transformation to the six branch points of S are shown in Fig. 2 bottom left; finally, the result obtained by applying connected component labeling to the zones of influence and to the voxels outside them is shown in Fig. 2 bottom right, where different colors represent different identity labels.

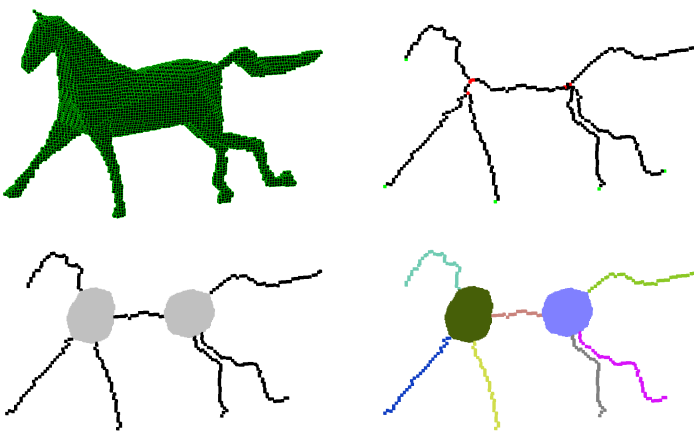


Fig. 2. From top left to bottom right: the object “horse”, its skeleton, the zones of influence and the result of connected component labeling

Still with reference to Fig. 2, we note that out of the nine skeleton branches initially detected in the skeleton S , only seven branches having at least one voxel outside the zones of influence are identified as having perceptual significance. These skeleton branches correspond to the object parts of which the object can be interpreted as constituted. The seven object parts intersect with each other in correspondence of the zones of influence. However, as it will be explained in the following, the actual overlapping region among the object parts in correspondence of a given zone of influence is likely to be larger than the zone of influence itself.

3.2 Identifying Object Decomposition Components

Ideally, by subtracting from the input object the zones of influence we should obtain a number of connected components of object voxels equal to the number of skeleton branches counted by connected component labeling. However, notwithstanding the fact that the image domain considered in this work includes only objects perceived as consisting of parts with almost tubular shape, a difference unavoidably exists between the volume of the input object and the volume of the object that could be obtained by applying to S the reverse distance transformation. Each zone of influence is certainly adjacent somewhere to the original background, but such an adjacency does not regard the whole surface delimiting the zone of influence. Thus, the number of achieved connected components of object voxels is generally smaller than expected. See Fig. 3 top left, where a section (black voxels) of the set that would be obtained by subtracting from “horse” the zones of influence (gray voxels) is shown. We observe that only one connected component of object voxels would remain after subtraction, while our aim is to decompose that part of the object into the four regions (torso, tail and the two back legs) that are in correspondence with the four detected perceptually meaningful skeleton branches.

To solve the above problem, we need to expand the zones of influence, so as to identify the proper overlapping regions. We aim at overlapping regions such that the cuts resulting in the object, when subtracting the overlapping regions from the input object, are in correspondence with significant curvature changes along the boundary of the object. To avoid both a too little expansion that would not produce the desired separation of the object parts, and an excessive expansion that would originate parts with unnatural separation cuts, we exploit distance information available in DT. In practice, the voxels of the zones of influence are labeled with the distances pertaining to them in DT. Then, the expansion is achieved by applying the reverse distance transformation to the so obtained distance labeled zones of influence.

The above process guarantees that the surfaces of the overlapping regions have a high adjacency degree with the original background. At the same time, reverse distance transformation also guarantees that unnatural cuts are not produced when the overlapping regions are subtracted from the input object. In fact, the zones of influence are nearly convex regions and this geometric property is preserved when the zones of influence are expanded. Thus, the overlapping regions cannot extend beyond the curvature minima along the boundary of the object. The overlapping regions obtained for the running example are shown in Fig. 3 top right.

We remark that the number of object parts has already been determined by counting the number of connected components of skeleton voxels outside the zones of influence. Thus, care should be taken to avoid diminishing such a number due to a possible fusion of the regions obtained when applying the reverse distance transformation to the distance labeled zones of influence. To this aim, a topology preserving reverse distance transformation is taken into account so as to avoid fusion. In practice, topology is maintained by setting to the background value the voxels that, though reached by the expansion of the zones of influence, result to be at the same distance from more than one zone of influence.

The overlapping regions are subtracted from the input object and the identity labels assigned to the voxels of the perceptually meaningful skeleton branches are finally used to label the connected components of object voxels they belong to. See Fig.3 bottom left, where the components of object voxels are colored as the corresponding perceptually meaningful skeleton branches. In this way, the preliminary object decomposition shown in Fig. 3 bottom right is obtained, where each overlapping region, colored as the influence zone from which has been obtained, still has to be ascribed to the proper component of object voxels.

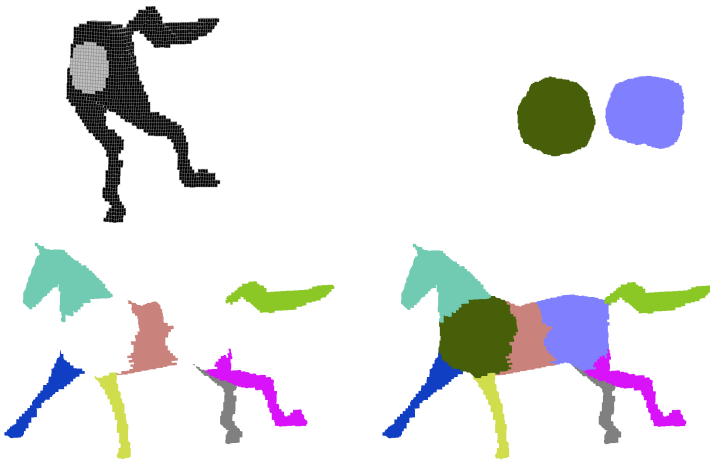


Fig. 3. From top left to bottom right: section showing in gray the voxels of the zones of influence; the overlapping regions obtained by applying the topology preserving reverse distance transformation to the zones of influence; the components of object voxels obtained by subtracting from the object the overlapping regions; and the preliminary decomposition of the object before ascribing the overlapping regions to the proper object parts

For the sake of completeness, we point out that when computing the set difference between the input object and the overlapping regions, some connected components composed by a small number of object voxels may exist, which do not include any skeletal voxel. Such components are likely to exist since their voxels were not reached when applying the reverse distance transformation to the zones of influence due to the fact that S does not include all the CMBs of the object. These components

do not correspond to meaningful object parts and their voxels are assigned to the overlapping regions they are adjacent to.

3.3 Ascribing the Overlapping Regions

According to our model, the object should be decomposed into a number of disjoint parts equal to the number of perceptually meaningful skeleton branches. To reach this goal, a decision must be taken to ascribe each overlapping region to only one of the disjoint components of object voxels adjacent to it.

By observing an overlapping region, say OR_k , and the adjacent components of object voxels, say P_1, P_2, \dots, P_n , we may note that while OR_k is an almost convex set, P_1, P_2, \dots, P_n show a different degree of concavity in correspondence with the positions where they result to be adjacent to OR_k . Let A_k denote the area of the surface bounding the overlapping region OR_k , i.e., the number of voxels of OR_k having at least one face-neighbor outside OR_k . Moreover, let $A_k(P_i)$ denote the portion of the area of the surface bounding the overlapping region OR_k that is adjacent to P_i , i.e., the number of voxels of OR_k having at least one face-neighbor in P_i . Then, we roughly evaluate how much OR_k intrudes into the adjacent component P_i , by computing the ratio $R = A_k(P_i)/A_k$. We ascribe the overlapping region OR_k to the adjacent component of object voxels that maximizes the ratio R . The choice of this criterion is due to the fact that in our opinion the more OR_k intrudes in a given component P_i , the more the shape of P_i benefits if the overlapping region is ascribed to it.

To accomplish the assignment of the overlapping regions in a computationally convenient manner, an adjacency matrix is built having as many rows as many are the overlapping regions and a number of columns equal to the number of components of object voxels plus one for the background. By inspecting the array where the preliminary decomposition of the object is stored, each time that a voxel of OR_k is met having at least one face neighbor outside OR_k , the proper element of the k -th row of the matrix is increased by one. In this way, once the matrix has been built, the value at row k and column j measures the portion of surface of OR_k in common with the background or with one of the adjacent components of object voxels. Then, we can easily decide to which component of object voxels to assign each overlapping region. In the rare case in which for an overlapping region an identical ratio is obtained for more than one adjacent component, the overlapping region is assigned to any of such components.

The adjacency matrix for the running example is shown in Table 1, where the two overlapping regions OR_1 and OR_2 are respectively those colored in green and blue in Fig. 3 bottom right.

Table 1. The adjacency matrix for the running example

	background	leg1	leg2	leg3	leg4	neck	torso	tail
OR_1	1517	44	33	0	0	125	411	0
OR_2	1622	0	0	70	53	0	414	5

The decomposition obtained for the running example is shown in Fig. 4 from two different view points.



Fig. 4. The obtained decomposition seen from two different view points

4 Experimental Results and Discussion

We have tested our method on 3D objects taken from publicly available shape repositories, e.g., [18], as done by most of the researchers in this field. The obtained results are generally satisfactory. In particular, for some objects used also in [7, 9, 12] to show the performance of the corresponding decomposition methods, our results seem to be qualitatively better. In Fig. 5 a few examples are given to show the performance of our decomposition method.

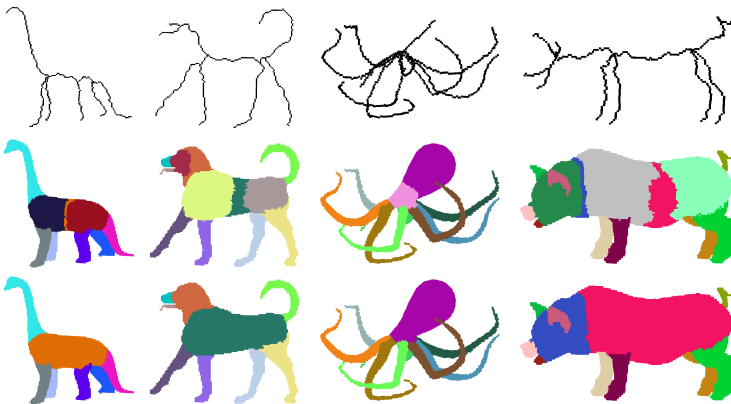


Fig. 5. From top to bottom: The skeletons of various input objects, the preliminary decompositions, and the resulting object decompositions

The algorithm runs on a Pentium 4 (3 GHz, 2 GB RAM) personal computer and its computational cost is $O(N)$, where N is the number of voxels in the image. The decomposition method is simple to implement, is rather fast and is completely automatic since it does not require any threshold. Of course, the quality of the results is influenced by the quality of the starting skeleton. In this respect, the skeletonization algorithm [17] has a positive impact on the decomposition method. In fact, due to the

use of the $\langle 3,4,5 \rangle$ distance that provides a good approximation of the Euclidean distance, the skeleton is rather stable under object rotation and for scale changes. Thus, stability also characterizes object decomposition. Moreover, the skeletonization algorithm includes a clever pruning step that, if the same object is presented in a different pose, allows us to obtain a skeleton with mostly the same structure. Actually, in this respect a key role for a satisfactory object decomposition is played by the zones of influence. In fact the skeleton of the same object in different poses may be characterized by a different number of branch points. However, due to the zones of influence we detect the same number of branch point configurations in all cases and, hence, the same number of perceptually significant skeleton branches.

Stability of the decomposition method with respect to pose/size changes can be appreciated by referring to Fig. 6, for the object “armadillo”. We observe that the main parts of the armadillo (the torso, the four limbs, the tail, the ears and the muzzle) are detected in all poses as individual decomposition parts. This is due to the detection of the zones of influence, which identifies the same number of branch point configurations in all cases, and to the criterion adopted to assign the overlapping regions to the proper adjacent components of object voxels. In turn, small peripheral parts, such as the toes, are not always individually detected as object parts since they are not individually mapped into skeleton branches.

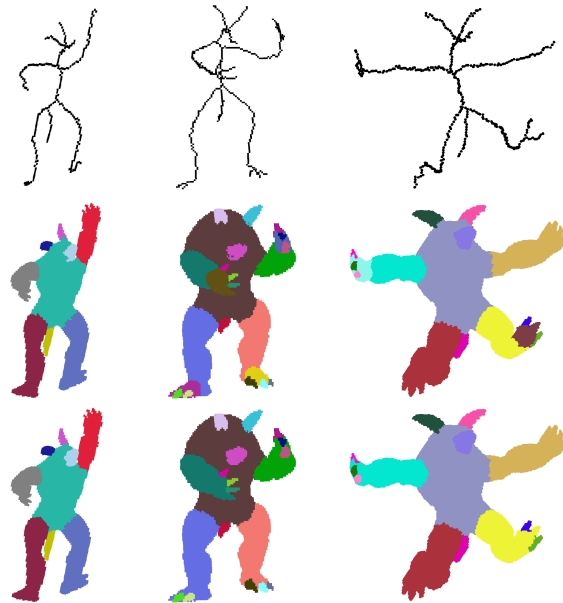


Fig. 6. The skeleton of “armadillo” in different poses/sizes, top, the preliminary decomposition, middle, and the decomposition after assignment of the overlapping regions, bottom

We have also tested stability of decomposition when the object is deformed, e.g., by stretching it without tearing it apart or sticking distinct parts together. For example, see Fig. 7, where the decomposition of a deformed version of “horse” is shown. We note that also in this case two overlapping regions are found, which are still assigned to the torso.

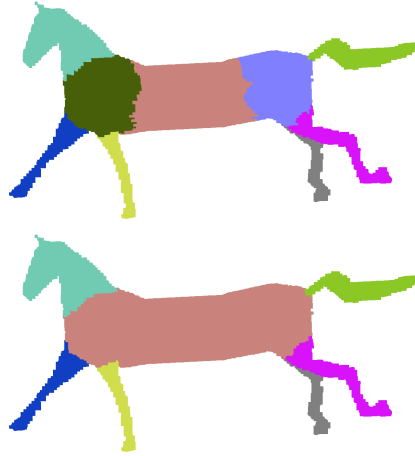


Fig. 7. The preliminary decomposition of a deformed version of “horse”, top, and the decomposition into the same number of disjoint parts as that achieved for the non deformed “horse”, bottom

As pointed out in the introduction, this decomposition method has some analogies with our previous decomposition method [11]. Both methods decompose the skeleton and generate object decompositions into disjoint parts. Moreover, the overlapping regions and the components of object voxels obtained by subtraction of the overlapping regions from the input have some analogy with the kernels and with the simple regions and bumps detected in [11]. However, the two methods are rather different, both as concerns the model and as concerns the computational cost.

As concerns the model, in [11] we give a prominent role to the kernels, while in this paper the key role is played by the object parts that correspond to the perceptually meaningful skeleton branches. The one-to-one correspondence between the skeleton components identified during the skeleton decomposition process and the parts into which the object is decomposed is maintained by the current process, which produces as many object parts as many are the detected perceptually significant skeleton branches. In turn, with the method in [11] the one-to-one correspondence is maintained only if no merging phase is accomplished; on the other hand, merging is almost always necessary in order to obtain a decomposition more in accordance with intuition.

As regards the computational cost, the current method is noticeably cheaper. The object parts are simply obtained by subtraction from the input of the overlapping regions, while a 2-step more sophisticated and expensive process was used in [11] to build the various regions starting from the skeleton components. Moreover, a concavity filling algorithm had to be used in [11] to move a number of voxels from any kernel to the adjacent regions, so as to have almost planar separation cuts where significant changes of curvature occurred along the boundary of the object. In turn, this is no longer necessary in the new method, due to the criterion adopted to identify the overlapping regions. Finally, merging was a necessary step in [11], which implies

an additional computational effort and the use of merging thresholds, while such a step is not necessary in the current work.

As already said, our method effectively works for objects perceived as composed by the superposition of parts with tubular shape, possibly characterized by different width. According to our method, the overlapping regions cannot individually exist, nor are divided among the adjacent components of object voxels. The component to which an overlapping region is ascribed is the one whose shape appears as completed by the assignment of the overlapping region. Such a component is characterized by width comparable to the width of the adjacent overlapping region, while the remaining adjacent components have smaller width.

To extend the applicability of the method to an image domain wider than that including only objects that are perceived as composed by parts with tubular shape, some considerations on the assignment of the overlapping regions can be done. For example, think of a rounded pincushion from which a number of pins come out. We can assume that a unique large overlapping region exists, which exhausts the portion of the 3D space occupied by the cushion. If such an overlapping region is assigned to one of the pins, the shape of the so obtained object part would have no perceptual evidence. The boundary of the compound part (cushion plus pin), in fact, would not be characterized by that good continuity that a human observer would consider adequate to perceive that compound object part as a whole. A feature that would certainly characterize the adjacency matrix built for the pincushion is that the ratio R remains always rather small, since the cushion intrudes very little within the pins. Thus, a way to extend the decomposition method to a wider domain is to introduce a threshold on the minimal value that the ratio R should have in order a compound region (overlapping region plus adjacent component of object voxels) can be reasonably perceived as a whole. Selecting the proper value for such a threshold will be argument of future investigations. If for all components of object voxels adjacent to a given overlapping region the ratio R is below the threshold value, the overlapping region is not assigned to any component and is taken as an individual decomposition part. Obviously, the method would be no longer fully automatic, since a threshold has to be set, the one-to-one correspondence between perceptually meaningful skeleton branches and object parts would be no longer guaranteed, since overlapping regions may be individual decomposition parts, but the method would have larger applicability.

5 Concluding Remarks

In this work we have introduced a 3D object decomposition method based on skeleton decomposition. The objects of interest are understood as constituted by parts with tubular shape and possibly different width. Starting from the linear skeleton of the object, the zones of influence, i.e., the regions where different skeleton branches meet, are identified. The zones of influence are used to group branch points sufficiently close to each other, and to identify the perceptually meaningful skeleton branches. Then, the overlapping regions, i.e., the object regions where object parts

intersect, are identified and components of object voxels are obtained by subtraction of the overlapping regions from the input object. The overlapping regions are finally ascribed to the adjacent components of object voxels that better benefit of such an assignment.

References

1. Palmer, S.E.: Hierarchical structure in perceptual representation. *Cognitive Psychology* 9, 441–474 (1977)
2. Marr, D., Nishihara, H.K.: Representation and recognition of three-dimensional shapes. *Proc. Royal Society of London: Series B* 200, 269–294 (1978)
3. Hoffman, D.D., Richards, W.A.: Parts of recognition. *Cognition* 18, 65–96 (1984)
4. Biederman, I.: Recognition-by-components: A theory of human image understanding. *Psychological Review* 94, 115–147 (1987)
5. Sanniti di Baja, G., Thiel, E.: The path-based distance skeleton: a flexible tool to analyse silhouette shape. In: *Proc. 12th ICPR*, vol. II, pp. 570–572 (1994)
6. Sanniti di Baja, G., Thiel, E.: (3,4)-weighted skeleton decomposition for pattern representation and description. *Pattern Recognition* 27, 1039–1049 (1994)
7. Cornea, N.D., Silver, D., Yuan, X., Balasubramanian, R.: Computing hierarchical curve-skeletons of 3D objects. *The Visual Computer* 21(11), 945–955 (2005)
8. Lien, J.-M., Geysler, J., Amato, N.M.: Simultaneous shape decomposition and skeletonization. In: *Proc. 2006 ACM Symposium on Solid and Physical Modeling*, pp. 219–228 (2006)
9. Reniers, D., Telea, A.: Skeleton-based hierarchical shape segmentation. In: *Proc. IEEE Int. Conf. on Shape Modeling and Applications*, pp. 179–188 (2007)
10. Serino, L., Sanniti di Baja, G., Arcelli, C.: Object decomposition via curvilinear skeleton partition. In: *Proc. 20th ICPR*, pp. 4081–4084 (2010)
11. Serino, L., Sanniti di Baja, G., Arcelli, C.: Using the Skeleton for 3D Object Decomposition. In: Heyden, A., Kahl, F. (eds.) *SCIA 2011*. LNCS, vol. 6688, pp. 447–456. Springer, Heidelberg (2011)
12. Chiang, P.-Y., Jay Kuo, C.-C.: Voxel-based shape decomposition for feature-preserving 3D thumbnail creation. *J. Vis. Commun. Image R.* 23, 1–11 (2012)
13. Borgefors, G.: On digital distance transform in three dimensions. *CVIU* 64(3), 368–376 (1996)
14. Blum, H.: Biological shape and visual science. *J. Theor. Biol.* 38, 205–287 (1973)
15. Nyström, I., Borgefors, G.: Synthesising Objects and Scenes Using the Reverse Distance Transformation in 2D and 3D. In: Braccini, C., Vernazza, G., DeFloriani, L. (eds.) *ICIAP 1995*. LNCS, vol. 974, pp. 441–446. Springer, Heidelberg (1995)
16. Svensson, S., Sanniti di Baja, G.: Using distance transforms to decompose 3D discrete objects. *Image and Vision Computing* 20, 529–540 (2002)
17. Arcelli, C., Sanniti di Baja, G., Serino, L.: Distance driven skeletonization in voxel images. *IEEE Trans. PAMI* 33(4), 709–720 (2011)
18. Shilane, P., Min, P., Kazhdan, M., Funkhouser, T.: The Princeton Shape Benchmark, Shape Modeling International, Genova, Italy (June 2004)