

Homological methods in feature extraction of multidimensional images

Marian Mrozek*, Marcin Żelawski†, Andrzej Krajniak†, Andrzej Gryglewski‡ and Sejin Han§

*Institute of Computer Science, Jagiellonian University, 30-348 Kraków, Poland
and Chair of Computational Mathematics, WSB-NLU,
33-300 Nowy Sacz, Poland

Email: Marian.Mrozek@ii.uj.edu.pl, Marian.Mrozek@wsb-nlu.edu.pl

†Institute of Computer Science, Jagiellonian University,
30-348 Kraków, Poland

Emails: zelawski@ii.uj.edu.pl, krajniak@ii.uj.edu.pl

‡Department of Anatomy, Collegium Medicum, Jagiellonian University
Kraków, Poland,

Email: msgrygle@cyf-kr.edu.pl

§Physics Department, University of Maryland,
College Park, MD, 20742, USA,

Email: sejinhan@umd.edu

Abstract—We show that the newly developed homology algorithms are helpful in imaging problems on the example of an algorithm extracting one dimensional features from a noisy image. We indicate that in some situations the global nature of this algorithm may become advantageous when compared with the standard algorithms based on skeletonization and pruning. The algorithm works in every dimension.

I. INTRODUCTION

The topological technique commonly used in image segmentation is the extraction of connected components of a black and white image. In the language of homology theory [1] this may be viewed as constructing the zeroth homology group of the image. Higher homology groups, which measure the presence of cycles, tunnels, and cavities in the image, are also used in computer vision and image processing [2], [1], [3], [4], [5], [6] but so far their usage is limited in part because the cycles, tunnels and cavities are not so frequent in images and in part because the classical algorithm for higher homology groups has cubical complexity which is in contrast to the linear time needed to construct connected components [7].

In this paper we show that higher homology groups are useful in image analysis even if the image itself and/or the features to be extracted contain no tunnels or holes. This is because holes may appear when the image is superimposed over some pattern (mask) and the appearance (or lack of appearance) of holes may be used to test certain features of the image via the study of homology generators. Moreover, image analysis based on higher homology groups may be performed quickly due to the recently developed reduction homology algorithms [8], [9], [10], which offer speed comparable to the speed of the algorithms constructing connected components.

The first four authors supported by Polish MNSzW, Grant N201 037 31/3151.

The implementations of these algorithms are available from [11] (see also [12], [13]).

To present the method we discuss a sample problem, which consists in the extraction of linear structures from the image under the presence of other features and noise. We became interested in the problem when studying two concrete issues: the analysis of 2D colonoscope images of blood vessels in colon mucosa [14], [15] and the analysis of 3D reconstituted confocal microscopy images of type I collagen fibrils [16]. Similar issues are studied in the analysis of blood vessel images in various other settings [17], [18], [19]), as well as in the detection of roads on satellite images ([20], [21]) and many other problems.

II. THE MODEL PROBLEM

To keep things simple, in this paper we discuss the problem in a model form. For a broader presentation of the problem in the context of concrete applications we refer the reader to [22]. By a linear structure we mean a feature of a binary image consisting of thin threads which may branch. We assume that a two or three dimensional image containing linear structures as well as some other features and noise is given. Moreover, we assume that the linear features are disconnected into several small pieces due to the noise. The goal is the separation of the linear structures from the other features and noise. This is a necessary step for further processing discussed in [22], which consists in finding the number and location of branching points together with the length of branches.

The standard approach to such problems is based on skeletonization of a binary image obtained by some thresholding (see [23] and the reference therein). However, because of the presence of noise separating the linear structure into many pieces, immediate skeletonization does not make sense.

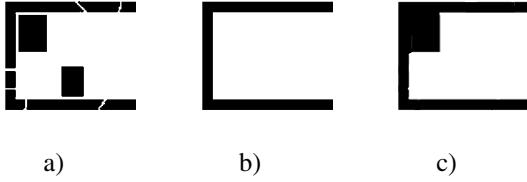


Fig. 1. An oversimplified model problem

Instead, some gluing, for example by means of morphological operations must be performed first. This may cause that some artifacts may be glued to linear structures which do not belong to them. Sometimes such artifacts may be later removed by trimming the short branches of the skeleton and disregarding the features which do not intersect the trimmed skeleton. However, in some situations this method does not work. For instance, consider the image presented in Figure 1a. The linear structure contained in this image consists of three bars presented in Figure 1b. Unfortunately, the standard approach, for example based on the algorithm presented in [23], at the best leads to the result presented in Figure 1c. More precisely, what happens is as follows. After applying wrapping to image in Figure 1a in order to glue together the pieces of the linear structure, the two artifacts in the form of rectangles are also attached to the linear structure. The contour and skeleton of the resulting image obtained by applying the method introduced in [23] are presented in Figure 2. Now we trim short branches. The question to be answered is: how short the branches should be to be trimmed. If we trim only the shortest branches, which do not branch themselves, the resulting trimmed skeleton will still intersect both rectangles, which do not belong to the linear structure. But even if we decide to trim all the branches leaving the skeleton consisting of one line only, we will get rid only of the rectangle in region 2, because the trimmed skeleton will still intersect the rectangle in region 1. As the consequence the standard method applied to this problem will always cause the effect of the presence of the undesired rectangle in the extracted image.

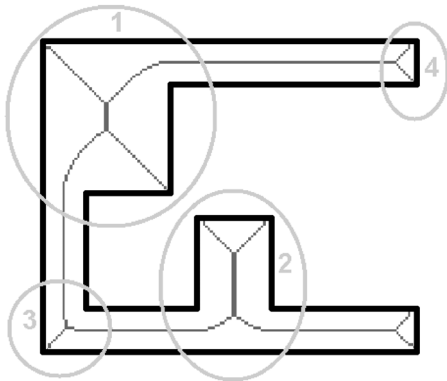


Fig. 2. Contour and skeleton of the image in Figure 1

Of course, in the case of the oversimplified image in Figure 1 it is easy to detect the undesired rectangular features directly. However, in the more realistic problems the components of the noisy linear structure may resemble in shape and size the components of the undesired features as in the image in Figure 3a. The linear structure to be extracted is presented in Figure 3b. The technique based on skeletonization and pruning leads to the image in Figure 3c. Also in this case an undesired effect is present in the circled region of Figure 3c. It takes the form of a thickening caused by an artifact glued to it. However, in this case it is not possible to get rid of the artifact by means of analyzing its shape as in our oversimplified example.

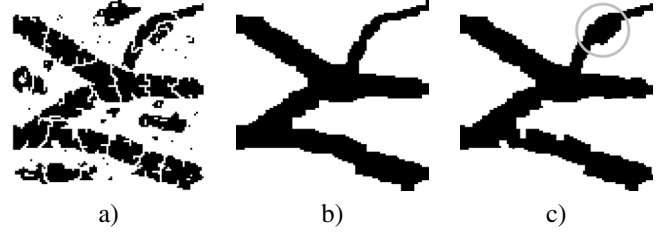


Fig. 3. A model problem

In the approach we propose the presence or lack of presence of a component in the final extracted image is determined by checking if the component is crucial in constructing long threads. For instance, the artifact which caused the unwanted thickening in the circled region of Figure 3c will not be present in the image extracted by our algorithm, because it is not needed to build long threads.

The method we use for determining if a component is needed in the final extracted image is based on the construction of first homology generators of a modified image. We describe it in the next section.

III. THE ALGORITHM

Consider the binary image in Figure 3a and the problem of extracting from this image the feature visible in Figure 3b consisting of four threads glued at three branching points.

The extra tool we propose is based on the detection of one dimensional holes in a modified image. We first explain the algorithm on the example based on processing the image in Figure 3a. We begin with constructing the morphological dilation X' (marked light gray in Figure 4) of the original image X in order to glue all the nearby connected components. Then we compute the generators of the relative homology $H(X', X)$ and add their supports (marked black in Figure 5 and enlarged in Figure 4 to X). In the next step we add a mask consisting of a collection of parallel hyperplanes (lines in the case of a two-dimensional image marked dark gray in Figure 5) to ensure that the branching points form holes and we compute the one-dimensional homology generators of the resulting set (marked dark and light gray in Figure 6). We restrict our attention only to generators which have non-empty intersection with the hyperplanes in the mask (marked dark gray in Figure 6). These generators surround holes formed

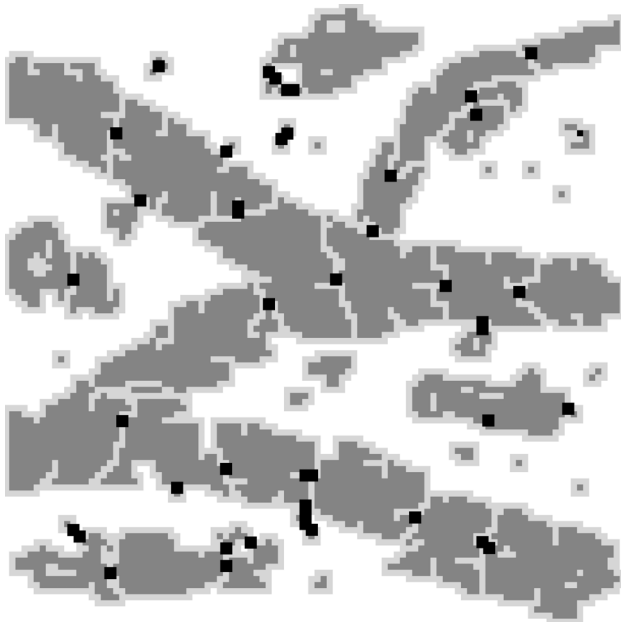


Fig. 4. Image with morphological dilation and with added supports of generators of $H(X', X)$

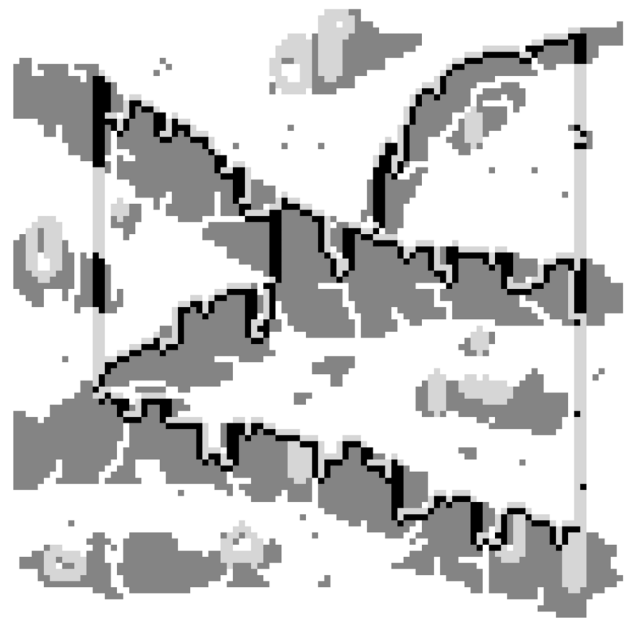


Fig. 6. Image with marked generators from step 6 of the algorithm

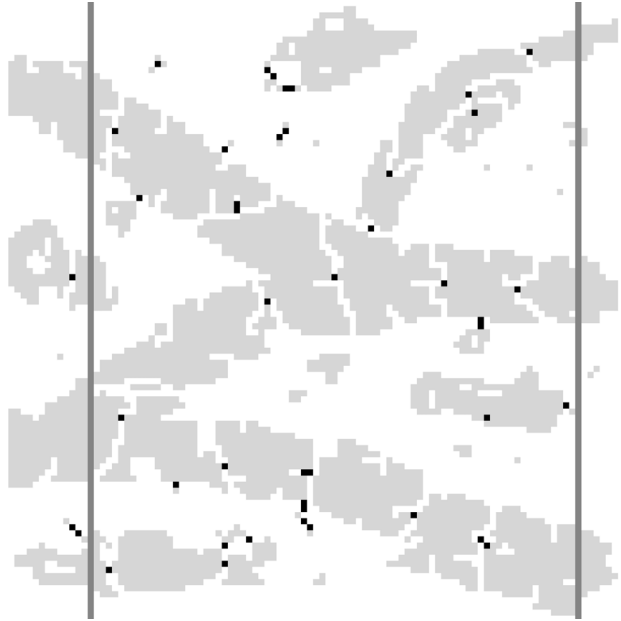


Fig. 5. Image with added vertical mask

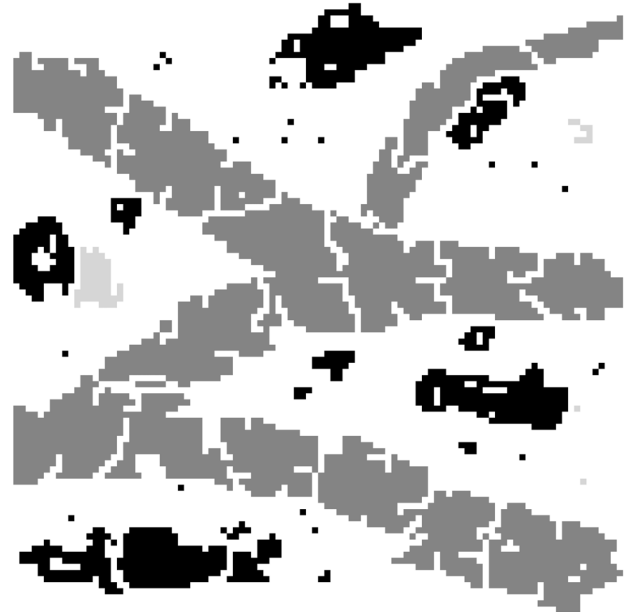


Fig. 7. Original image with discarded parts marked

by the superimposition of the mask over the original image, because the mask itself has no holes. Therefore, the parts of the generators disjoint from the mask must extend along the threads and branching points in the image. However, they are unlikely to intersect the other features of the image which may have been glued to the set in the dilation process. Now we add the support of the selected generators to the original image and disregard the connected components of the resulting image which have empty intersection with the selected generators

(marked black in Figure 7). Finally we discard the remaining small components (marked light gray in Figure 7) and apply smoothing based on the closing process to obtain the image in Figure 3b.

This procedure may be summarized in the following algorithm, in which X stands for a black-and-white image considered as a subset of a rectangular grid of points in the Euclidean space and for a homology generator c the notation $|c|$ represents the smallest subset of the image entirely covering the support of c .

Algorithm 1: Extracting the linear structure

1. Input: image X , integers m and a
2. $X' :=$ the morphological dilation of X
3. $A :=$ the union of the supports
of the generators of $H_1(X', X)$
4. $Y := X \cup A$
5. $M :=$ collection of grid points belonging to
parallel hyperplanes separated by distance m
6. $\mathcal{C} :=$ homology generators of $H_1(Y \cup M)$
7. $D := \emptyset$
8. for each $c \in \mathcal{C}$
if $|c| \cap M \setminus X \neq \emptyset$ then $D := D \cup (|c| \cap X)$
9. $\mathcal{W} :=$ the family of connected components of $D \cup X$
which have nonempty intersection with D
10. Output: $\bigcup \{W \in \mathcal{W} \mid \text{card } W \geq a\}$

The parameter a is a threshold value for selecting large connected components. If the linear features are relatively large when compared to the other features and artifacts present in the image, then selecting a good value for a is not a problem even if originally the linear features are broken by noise into several small connected components. This is caused by the fact that the broken components are already glued when a is used in the last step of the algorithm.

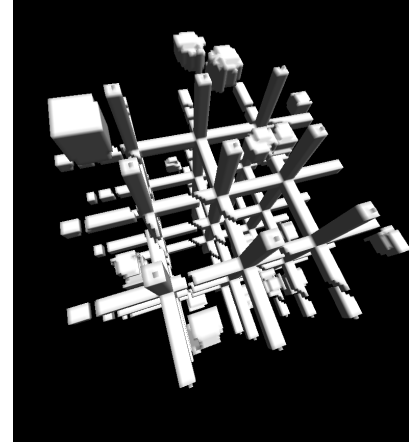
The parameter m influences the size of branches detected by the algorithm. The larger m , the larger are the branches discarded by the algorithm. Even in the case of a large m a short branch coming from an artifact may not be discarded by the algorithm. This will happen when a hyperplane of the mask will pass close to the artifact. The chances for that are small and may be even further decreased by running the algorithm a few times with shifted location of the hyperplanes and taking the intersection of the outputs.

The direction of the hyperplanes may also influence the quality of the output. If there is no natural choice of the direction, one can run the algorithm for several choices of the direction and then take the union of the outputs.

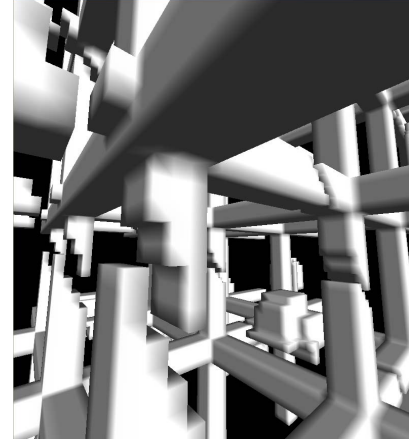
The strong advantage of the method is that the proposed algorithm may be used in higher dimensions without any change. For instance, consider the image in Figure 8a which contains a three-dimensional grid of broken bars organized in a regular grid with some artifacts. A magnification of a part of this image is presented in Figure 8b and the outcome of the algorithm is presented in Figure 8c. Another three-dimensional example is presented in Figure 9.

IV. CONCLUSION

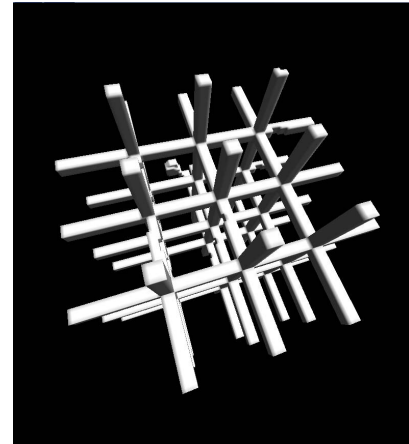
The presented method shows that the algorithms constructing homology generators are helpful in solving imaging problems in which the standard methods based on skeletonization and pruning fail. Unlike the standard methods, the algorithm on which the method is based does not require any additional levels of processing (construction of graph, graph analysis) needed in the standard approach. This is because the algorithm operates directly on the black-and-white image. The method is dimension independent and has been successfully tested



a)



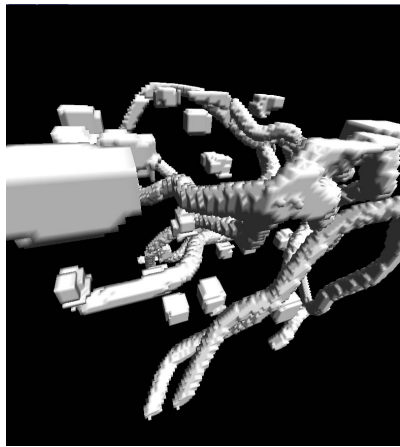
b)



c)

Fig. 8. An artificial three dimensional example based on grid of bars: a) input b) zoom c) output

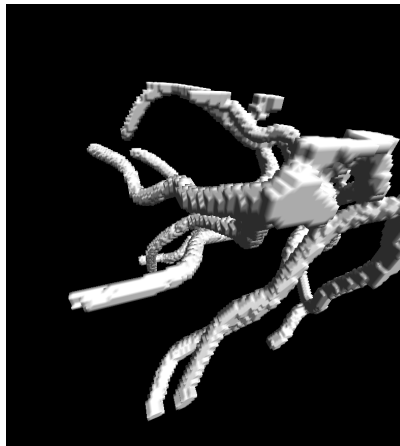
on two and three dimensional images constructed artificially as well as coming from concrete two and three dimensional problems (see [22]). Moreover, the method may be easily adapted to the extraction of branching surface-like features from a noisy three dimensional image in the presence of other features.



a)



b)



c)

Fig. 9. A three dimensional example consisting of bending threads: a) input b) zoom c) output

Although the presented method is restricted to black-and-white images, a more sophisticated method to analyse gray scale images may be obtained by an analogous topological method based on a decomposition into a filtration of black and white images. The filtration may be then be used to investigate the image via studying the homology of the inclusion maps

and/or persistent homology (cf. [24]). Similar approach is also possible for color images. The gray scale and color case are left for future investigation.

REFERENCES

- [1] T. Kaczynski, K. Mischaikow and M. Mrozek, *Computational Homology*, Springer-Verlag, Appl. Math. Sci. Series, Vol. 157, New York, 2004.
- [2] M. Allili, K. Mischaikow and A. Tannenbaum, Cubical homology and the topological classification of 2d and 3d imagery, in *Proceedings of International Conference Image Processing*, Vol. 2, pp. 173–176, 2001.
- [3] M. Niethammer, A.N. Stein, W.D. Kalies, P. Pilarczyk, K. Mischaikow and A. Tannenbaum, Analysis of blood vessel topology by cubical homology, *Proceedings of the International Conference on Image Processing*, Vol. 2, pp. 969–972, 2002.
- [4] A. Sędziwy and M. Żelawski, Detecting Pathologies with Homology Algorithms in 3D MRI/CT Images of Brain, in preparation.
- [5] M. Żelawski, Pattern Recognition Based on Homology Theory, *Machine Graphics & Vision*, Vol. 14, No. 3, pp. 309–324, 2005.
- [6] M. Żelawski, Detecting Pathologies with Homology Algorithms in Magnetic Resonance Images of Brain, *Machine Graphics & Vision*, accepted.
- [7] M. Sonka, V. Hlavac and R. Boyle, *Image Processing, Analysis and Machine Vision*, Brooks/Cole Publishing Company, 1999.
- [8] M. Mrozek, P. Pilarczyk and N. Żelazna, Homology Algorithm Based on Acyclic Subspace, *Computers and Mathematics with Applications*, Vol. 55, No. 11, pp. 2395–2412, 2008.
- [9] M. Mrozek and B. Batko, Coreduction homology algorithm, *Discrete and Computational Geometry*, Vol. 41, pp. 96–118, 2009.
- [10] S. Peltier, S. Alayrangués, L. Fuchs and J.O. Lachaud, Computation of homology groups and generators, *Computers and Graphics*, Vol. 30, pp. 62–69, 2006.
- [11] M. Mrozek, *Homology Software*, <http://www.ii.uj.edu.pl/~mrozek/software/homology.html>, 2006.
- [12] *Computer Assisted Proofs in Dynamics*, <http://capd.wsb-nlu.edu.pl>.
- [13] *Computational Homology Project*, <http://www.math.rutgers.edu/~chompl/>.
- [14] M.A. Konerding, E. Fait and A. Baumann, 3D microvascular architecture of pre-cancerous lesions and invasive carcinomas of the colon, *Brit J Cancer*, Vol. 84, pp. 1354–1362, 2001.
- [15] W. McLaren, P. Anikijenko, S. Thomas, P. Delaney and R. King, In vivo detection of morphological and microvascular changes of the colon in association with colitis using fiberoptic confocal imaging, *Dig Dis Sci*, Vol. 47, pp. 2424–2433, 2002.
- [16] S. Han, D.J. McBride, W. Losert, and S. Leikin, Segregation of Type I collagen Homo- and Heterotrimers in Fibrils, *Journal of Molecular Biology*, Vol. 383, No. 1, pp. 122–132, 2008.
- [17] C. Kirbas and F. Quek, A review of vessel extraction techniques and algorithms, *ACM Comput. Surv.*, Vol. 36, No. 2, pp. 81–121, 2004.
- [18] K. Bühler, P. Felkel and A. La Cruz, Geometric methods for vessel visualization and quantification – a survey, in *Geometric Modelling for Scientific Visualization*, G. Brunnett, B. Hamann and H. Müller (Eds.), pp. 421–399, 2003.
- [19] P. Felkel, R. Wegenkittl and A. Kanitsar, Vessel Tracking in Peripheral CTA Datasets An Overview, in *SCCG 2001: Proceedings of the 17th Spring conference on Computer graphics*, Washington, DC, USA: IEEE Computer Society, pp. 232–239, 2001.
- [20] S. Hinz, A. Baumgartner, Automatic extraction of urban road networks from multi-view aerial imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, No. 1–2, pp. 83–98, 2003.
- [21] F. Tupin, B. Houshmand, M. Datcu, Road detection in dense urban areas using SAR imagery and the usefulness of multiple views, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, No. 11, pp. 2405–2414, 2002.
- [22] M. Mrozek, M. Żelawski, A. Gryglewski, S. Han, A. Krajniak, Extraction and analysis of linear features in multidimensional images by homological methods, in preparation.
- [23] X. Bai, L.J. Latecki and W. Liu, Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution, *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 29, No. 3, pp. 449–462, 2007.
- [24] H. Edelsbrunner, D. Letscher and A. Zomorodian, Topological persistence and simplification, *Discrete Comput. Geom.*, Vol. 28, pp. 511–533, 2002.