

Incremental-Decremental Algorithm for Computing AT-Models and Persistent Homology^{*}

Rocio Gonzalez-Diaz¹, Adrian Ion^{2,3},
Maria Jose Jimenez¹, and Regina Poyatos¹

¹ Applied Math Dep. (I), School of Computer Engineering,
University of Seville

{rogodi,majiro,rpoyatos}@us.es

² Pattern Recognition and Image Processing Group,
Vienna University of Technology

ion@prip.tuwien.ac.at

³ Institute of Science and Technology, Austria

Abstract. In this paper, we establish a correspondence between the incremental algorithm for computing AT-models [8,9] and the one for computing persistent homology [6,14,15]. We also present a decremental algorithm for computing AT-models that allows to extend the persistence computation to a wider setting. Finally, we show how to combine incremental and decremental techniques for persistent homology computation.

Keywords: Persistent homology, AT-model for computing homology, cell complex.

1 Introduction

Homology is a topological invariant i.e. it is a property of an object which does not change under continuous (elastic) transformations of the object. Homology characterizes “holes” in any dimension (e.g. connected components, tunnels, cavities, etc.) by means of *cycles*. Given a combinatorial object made up by basic building blocks called *cells* (vertices, edges, faces, etc.), a cycle is a set of cells that “surround” a hole or a part of the object (e.g. a closed path in 2D, a closed path or a closed surface in 3D). Intuitively a *homology class* collects all cycles that “surround” the same hole (a precise mathematical definition is given later). The homology of a given object is then fully characterized by a basis of independent homology classes, which in turn is characterized by identifying one cycle, called *representative cycle*, for each of these classes.

Persistent homology studies homology classes and their lifetimes (persistence). Notice that while homology characterizes an object, persistent homology characterizes a sequence of growing object-instances i.e. an object together with an ordering for the cells (called a *filtration*). In recent years, persistent homology has found its way to applications, where it is mainly used to identify salient

^{*} Partially supported by the Austrian Science Fund under P20134-N13.

features of an object in the presence of noise. E.g. find relevant local maxima without smoothing, compute the similarity of two objects as the similarity of their persistence information [5,3]. What all these applications have in common is that the object under study is fixed (e.g. one picture [3], one set of 3D sample points [4], one scan of a bone [7], etc.).

Current sensor and recording technologies provides not just one such recording but whole sequences over the temporal domain. Video has become ubiquitous, with decent to good quality recordings being produced by most mobile devices (phones, PDAs) and low priced webcams. Medical imaging is also moving from single 2D/3D image capture to recordings over a certain time period (e.g. sequences of ultrasound images). One way to deal with sequences is to take each frame separately and do all computation independently. However, due to temporal continuity, the same object can look very similar in consecutive frames and the overlap in the image can be high. Moreover, temporal information present in the sequence can help to identify salient features, as ones having a long lifetime not just over the filtration in the same frame, but also over multiple frames. There is no guarantee of inclusion or growing of objects in consecutive frames, some parts might “disappear” and some might “appear”. A theory of decremental persistence, and incremental-decremental algorithms are needed for such arbitrary changes in the object.

In this paper we describe a further step in providing such a theory. First, we establish a correspondence between the incremental algorithm for computing AT-models [8,9] and the one for computing persistent homology [6,14,15]. Then, we provide a decremental algorithm for computing AT-models, suitable for extending the computation of persistence with the combination of an incremental-decremental technique.

2 Background

We consider $\mathbf{Z}/2$ as the ground ring throughout the paper.

Roughly speaking, a *cell complex* is a general topological structure by which a space is decomposed into basic elements (*cells*) of different dimensions, which are *glued* together by their boundaries (see a formal definition of CW-complex in [12]). If the building blocks (cells) of a cell complex are convex polytopes (vertices, edges, polygons, polyhedra, ...) then the cell complex is a *polyhedral cell complex*. Given a (polyhedral) cell complex K , a *proper face* of $\sigma \in K$ is a face of σ whose dimension is strictly less than the one of σ . A *facet* of σ is a proper face of σ of maximal dimension. A *maximal cell* of K is a cell of K which is not a proper face of any other cell of K .

For any graded set $S = \{S_p\}_p$ (subscript is used to denote the dimension of the elements), one can consider formal sums of elements of $S_p = \{s_p^1, \dots, s_p^{m_p}\}$, for a fixed p , which are called *p-chains*, and which form an abelian group, denoted by $C_p(S)$, with respect to the component-wise addition (mod 2). Therefore, a *p-chain* c is $c = \sum_{i=1}^m a^i s_p^i$, where $a^i \in \mathbf{Z}/2$ for $i = 1, \dots, m$. This way, $s_p^i \in c$ if $a^i = 1$. The collection of all the chain groups associated to S , $\{C_p(S)\}_p$, is called also

chain group, for simplicity. A *chain complex* is a collection $\mathcal{C}(S) = \{C_p(S), \partial_p^S\}_p$, where $\partial^S = \{\partial_p^S : C_p(S) \rightarrow C_{p-1}(S)\}$ is a square zero homomorphism (i.e., $\partial_{p-1}\partial_p \equiv 0$) called the *boundary operator*. The boundary of a q -cell is the formal sum of all its facets. It is extended to q -chains by linearity. A homomorphism $f = \{f_p : C_p(S) \rightarrow C_p(S')\}_p$ is a *chain map* if $f_{p-1}\partial_p^S \equiv \partial_p^{S'}f_p$, for all p . For simplicity, we sometimes write $f : C(S) \rightarrow C(S')$ instead of $f = \{f_p : C_p(S) \rightarrow C_p(S')\}_p$ and $f(\sigma)$ instead of $f_p(\sigma)$. A p -chain $a \in C_p(S)$ is called a p -*cycle* if $\partial_p^S a = 0$. If $a = \partial_{p+1}^S b$ for some $b \in C_{p+1}(S)$ then a is called a p -*boundary*. We say that two p -cycles a and b are *homologous* if there exists a $(p + 1)$ -chain $c \in C_{p+1}(S)$ such that $a = b + \partial_{p+1}^S c$. Define the p -*th homology group* to be the quotient group of p -cycles mod p -boundaries denoted by $H_p(S)$. Each element $[a]$ of $H_p(S)$ is a quotient class obtained by adding each p -boundary to a given p -cycle a called a *representative cycle* of the homology class $[a]$. The *homology* of S is the chain group $\mathcal{H}(S) = \{H_p(S)\}_p$. See [13] for further details.

3 AT-Model

An algebraic topological (AT) model (implicitly used in [8] and first defined in [9]) for a given cell complex K not only permits to compute homology but also finer topological invariants such as cohomology or the cohomology ring.

An *AT-model* for a cell complex K is an algebraic set (f, g, ϕ, K, H) , where:

- K is the cell complex.
- $H \subseteq K$ describes the homology of K , in the sense that it contains a distinct p -cell for each p -homology class of a basis, for all p . The cells in H are called *surviving cells*. Since $\partial_p^H \equiv 0$ for all p , $\mathcal{C}(H)$ is simply the chain group $\{C_p(H)\}_p$. Moreover, $\mathcal{C}(H)$ (the chain group generated by H) and $\mathcal{H}(K)$ (the homology of K) are isomorphic [8,9]. Therefore, every cell of H corresponds to a homology class generator.
- $g = \{g_p : C_p(H) \rightarrow C_p(K)\}_p$ is a chain map that maps each p -cell h in H to one representative cycle $g_p(h)$ of the corresponding class $[g_p(h)]$ in $H_p(K)$.
- $f = \{f_p : C_p(K) \rightarrow C_p(K)\}_p$ is a chain map that maps each p -cell x in K to a sum of surviving cells, satisfying that if $a, b \in C_p(K)$ are two homologous p -cycles then $f_p(a) = f_p(b)$. Moreover, $fg(x) = x$ for any $x \in H$.
- $\phi = \{\phi_p : C_p(K) \rightarrow C_{p+1}(K)\}_p$ is a *chain homotopy* (see [13]) from gf to the identity homomorphism of $\mathcal{C}(K)$. Intuitively $\phi(\sigma)$ returns the cells needed to be contracted to “bring” σ to a surviving cell.

Fig. 1 is an example of an AT-model for a single pixel codified as a cubical complex.

In order to establish a connection between the existing algorithms for computing AT-models [9] and the theory of persistent homology [6,14,15], we add two **extra-conditions** which ensure that the basis for $\mathcal{H}(K^i)$ is maintained implicitly through the cells in H^i (see Section 5 for more details):

- (P1) Annihilation: $\phi_{p+1}\phi_p \equiv 0$, $f_{p+1}\phi_p \equiv 0$ and $\phi_p g_p \equiv 0$.
- (P2) If $h \in K$ is a surviving p -cell then $f_p(h) = h$ and $\phi_p(h) = 0$.

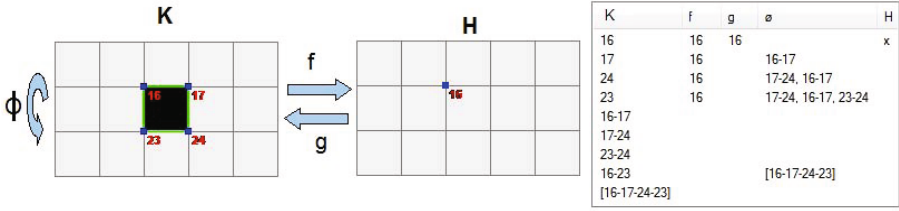


Fig. 1. AT-model for a single pixel

Lemma 1. *The extra-conditions (P1) and (P2) are satisfied by the incremental and decremental algorithm for computing AT-models presented in this paper.*

4 Algorithms for Computing AT-Models

Now, we give an intuitive approach to the incremental algorithm for computing AT-models given in [8] in order to establish an interpretation in terms of persistence, and present a decremental algorithm for computing AT-models with the aim of extending the computation of persistent homology to a more general setting.

Incremental AT-model: Let K be a cell complex with a full ordering of its cells, $\{\sigma^1, \dots, \sigma^n\}$, satisfying that if σ^i is a face of σ^j then $i < j$. Consider a filtration of K , i.e. a nested sequence of subcomplexes, $\emptyset = K^0 \subseteq K^1 \subseteq K^2 \subseteq \dots \subseteq K^n$ such that $K^i = \{\sigma^1, \dots, \sigma^i\}$. All the proper faces of σ^i are in K^{i-1} .

First, define an AT-model $(f^1, g^1, \phi^1, K^1, H^1)$ for K^1 : $H^1 := \{\sigma^1\}$, $f^1(\sigma^1) := \sigma^1$, $g^1(\sigma^1) := \sigma^1$ and $\phi^1(\sigma^1) := 0$. Second, successively add a new cell σ^i , for $i = 2, \dots, n$, computing a new AT-model $(f^i, g^i, \phi^i, K^i, H^i)$ for $K^i = K^{i-1} \cup \{\sigma^i\}$, as follows: Initially, $H^i := H^{i-1}$; $f^i(\mu) := f^{i-1}(\mu)$ and $\phi^i(\mu) := \phi^{i-1}(\mu)$, for any $\mu \in K^i$ different from σ^i ; $g^i(h) := g^{i-1}(h)$, for any $h \in H^i$. Consider $f^{i-1}\partial(\sigma^i)$ to detect if σ^i will create or destroy a homology class:

- (1) If $f^{i-1}\partial(\sigma^i) = 0$ (a new homology class is created) then, $H^i := H^{i-1} \cup \{\sigma^i\}$, $f^i(\sigma^i) := \sigma^i$, $g^i(\sigma^i) := \sigma^i + \phi^{i-1}\partial(\sigma^i)$ and $\phi^i(\sigma^i) := 0$.
- (2) If $f^{i-1}\partial(\sigma^i) \neq 0$ (a homology class is destroyed), let j be the largest index such that $\sigma^j \in f^{i-1}\partial(\sigma^i)$. Then, observe that $j < i$ and $\dim \sigma^j = \dim \sigma^i - 1$. Then, $H^i := H^{i-1} \setminus \{\sigma^j\}$, $f^i(\sigma^i) := 0$, $\phi^i(\sigma^i) := 0$.

Besides, two operations are applied to all cells $x \in K^i$ such that $\sigma^j \in f^{i-1}(x)$:

- Update f : $f^i(x) := f^{i-1}(x) + f^{i-1}\partial(\sigma^i)$. Intuitively, it “propagates” over σ^i the information for f of the cells in $\partial\sigma^i$ and cancels out σ^j .
- Update ϕ : $\phi^i(x) := \phi^{i-1}(x) + \sigma^i + \phi^{i-1}\partial(\sigma^i)$. Roughly speaking, $\sigma^i + \phi^{i-1}\partial(\sigma^i)$ is a connection between the old and the new surviving cell(s).

Relation to persistent homology: The algorithm for computing persistent homology that appears in [6,14,15], marks a k -cell σ^i as positive if it belongs to a k -cycle in $C(K^i)$ (σ^i creates a new homology class at time i) and negative otherwise (σ^i destroys the homology class created before by σ^j for $j < i$ and,

in this case, σ^i is paired with σ^j). The following lemmas show the equivalence between these concepts and the incremental AT-model above.

Lemma 2. σ^i belongs to a k -cycle in $C(K^i)$ if and only if $f^{i-1}\partial(\sigma^i) = 0$.

Proof. If $f^{i-1}\partial(\sigma^i) = 0$, then $\partial\sigma^i = \partial\phi^{i-1}(\partial\sigma^i)$. Therefore, $\sigma^i + \phi^{i-1}\partial(\sigma^i)$ is a k -cycle and σ^i belongs to it. Conversely, if σ^i belongs to a k -cycle a in $C(K^i)$, then $a = \sigma^i + b$ where b is a k -chain in $C(K^{i-1})$. Since $\partial a = 0$ then $f^{i-1}(\partial\sigma^i) = f^{i-1}(\partial b)$. Since $f^{i-1}(\partial b) = \partial f^{i-1}(b) = 0$ then $f^{i-1}(\partial\sigma^i) = 0$. \square

By Lemma 2, the fact of marking a cell σ^i as positive is equivalent to holding condition $f^{i-1}\partial(\sigma^i) = 0$ in the incremental algorithm for computing AT-models.

Following the theory of persistent homology, a *canonical cycle* c^i is a non-bounding cycle that contains σ^i but no other positive cell.

Lemma 3. If σ^i is positive, then $g^i(\sigma^i) = \sigma^i + \phi^{i-1}\partial(\sigma^i)$ is a canonical cycle.

At time i , the youngest cell $\sigma^j \in \Gamma(\partial\sigma^i)$ is paired with σ^i , identifying σ^i as the destroyer, σ of the homology class created by σ^j . Once one has the pairing of positive and negative cells, computing the persistent Betti numbers is trivial. To measure the life-time of a non-bounding cycle, one has to find when the cycle's homology class is created by a positive cell and destroyed by a negative cell. To detect these events, the collection of positive k -cells $\Gamma(d)$ for a given cycle d such that d and $\sum_{\sigma^g \in \Gamma(d)} c^g$ are homologous, is obtained using the incremental method for computing AT-models as follows:

Lemma 4. Any k -cycle d in $C(K^{i-1})$ is homologous to $g^{i-1}f^{i-1}(d)$. Moreover, $\Gamma(d) = f^{i-1}(d)$ and $g^{i-1}f^{i-1}(d) = \sum_{\sigma^g \in \Gamma(d)} c^g$.

Decremental AT-model: Let (f, g, ϕ, K, H) be an AT-model for a cell complex K satisfying the extra-conditions (P1) and (P2). Let σ be a maximal cell of K . Then an AT-model for $K' = K \setminus \{\sigma\}$ is constructed as follows:

Algorithm 1. Initially, $H' := H$, $g'(h) := g(h)$ for all $h \in H'$, $f'(x) := f(x)$ and $\phi'(x) := \phi(x)$ for all $x \in K'$.

- (1) If there exists $\beta \in H$ such that $\sigma \in g(\beta)$ then σ destroys the homology class $[g(\beta)]$ created before by β . Therefore, $H' := H \setminus \{\beta\}$; $f'(x) := f(x) + \beta$ if $\beta \in f(x)$ and $x \in K'$; $g'(h) := g(h) + g(\beta)$ if $\sigma \in g(h)$ and $h \in H'$; $\phi'(y) := \phi(y) + g(\beta)$ if $\sigma \in \phi(y)$ and $y \in K'$.
- (2) Otherwise, since $\sigma \in \phi\partial\sigma$, there exists $\mu \in \partial\sigma$, $\mu \notin H$, such that $\sigma \in \phi(\mu)$. Then μ creates a new homology class $[g'(\mu)]$. Therefore: $H' := H \cup \{\mu\}$; $g'(\mu) := gf(\mu) + \partial\phi(\mu)$. $f'(x) := f(x) + \mu + f(\mu)$ and $\phi'(x) := \phi(x) + \phi(\mu)$ if $\sigma \in \phi(x)$ and $x \in K'$.

In order to satisfy the following proposition, the formulas for f' and ϕ' in step (2) above are different to the ones given in [9].

Proposition 1. The output of Alg. 1, (f', g', ϕ', K', H') , is an AT-model for $K' = K \setminus \{\sigma\}$ satisfying the extra-conditions (P1) and (P2).

Proof. Proof of step (1) in Alg. 1 is given in [9]. The verification of the rest of the properties follows a similar strategy and is left to the reader. \square

5 Incremental-Decremental Algorithm for Computing AT-Models and Persistent Homology

Now, let $\emptyset = K^0 \leftrightarrow K^1 \leftrightarrow \dots \leftrightarrow K^n$ be a sequence of cell complexes (that we call a *zig-zag filtration*), such that every two consecutive complexes differ by a single cell, i.e. either $K^i = K^{i-1} \cup \{\sigma\}$ or $K^i = K^{i-1} \setminus \{\sigma\}$. Let $\{\sigma^1, \dots, \sigma^m\}$, $m \leq n$, be the ordered set of all the cells added in a given zig-zag filtration such that if $i < j$ then σ^i was added before σ^j to the filtration. Then, one can consider the sequence of homology groups $H(K^0) \leftrightarrow H(K^1) \leftrightarrow \dots \leftrightarrow H(K^n)$ where the connecting homomorphisms are induced by inclusion.

Incremental-decremental algorithm: Initially, $H^1 := \{\sigma^1\}$, $f^1(\sigma^1) := \sigma^1$, $g^1(\sigma^1) := \sigma^1$ and $\phi^1(\sigma^1) := 0$. At time i , a cell σ is added or removed. Then, use the incremental or decremental algorithm presented here, respectively, for computing the AT-model $(f^i, g^i, \phi^i, K^i, H^i)$. Two cases can occur:

- (1) A homology class is created by a positive cell μ . If $K^i = K^{i-1} \cup \{\sigma\}$ then $\mu := \sigma$. If $K^i = K^{i-1} \setminus \{\sigma\}$, then $\mu := \sigma^j$ where σ^j is the youngest cell in $\partial\sigma$ such that $\sigma \in \phi^{i-1}(\sigma^j)$. The cell μ is added to H^{i-1} to get H^i .
- (2) A homology class represented by a positive cell σ^j is destroyed by a negative cell σ^k , $j < k \leq i$. If $K^i = K^{i-1} \cup \{\sigma\}$ then $\sigma^k := \sigma$ and σ^j is the youngest cell in $f^{i-1}\partial(\sigma)$. If $K^i = K^{i-1} \setminus \{\sigma\}$, then $\sigma^k := \sigma$ and σ^j is the youngest cell in H^{i-1} such that $\sigma^k \in g^{i-1}(\sigma^j)$. The cell σ^j is removed from H^{i-1} to get H^i .

If a cell μ creates a homology class at time j and it is destroyed at time i , $j < i$, then a horizontal line from (j, ℓ) to (i, ℓ) is added to the corresponding barcode (see [2]); If a cell μ creates a homology class at time i and it survives along the process then a horizontal line from (i, ℓ) to (∞, ℓ) is added, where ℓ is the index of the cell μ in the given ordering of the cells. See examples of barcodes using incremental-decremental algorithm for computing AT-models in Fig. 2, 3, 4. See Fig. 5 as an example of application with digital images.

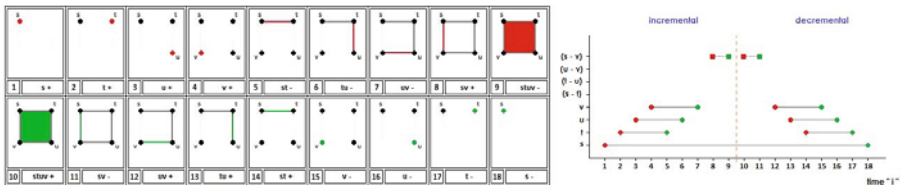


Fig. 2. Symmetric zig-zag filtration and barcode

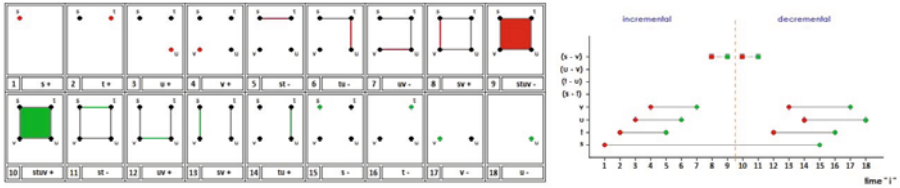


Fig. 3. Non-symmetric zig-zag filtration and barcode

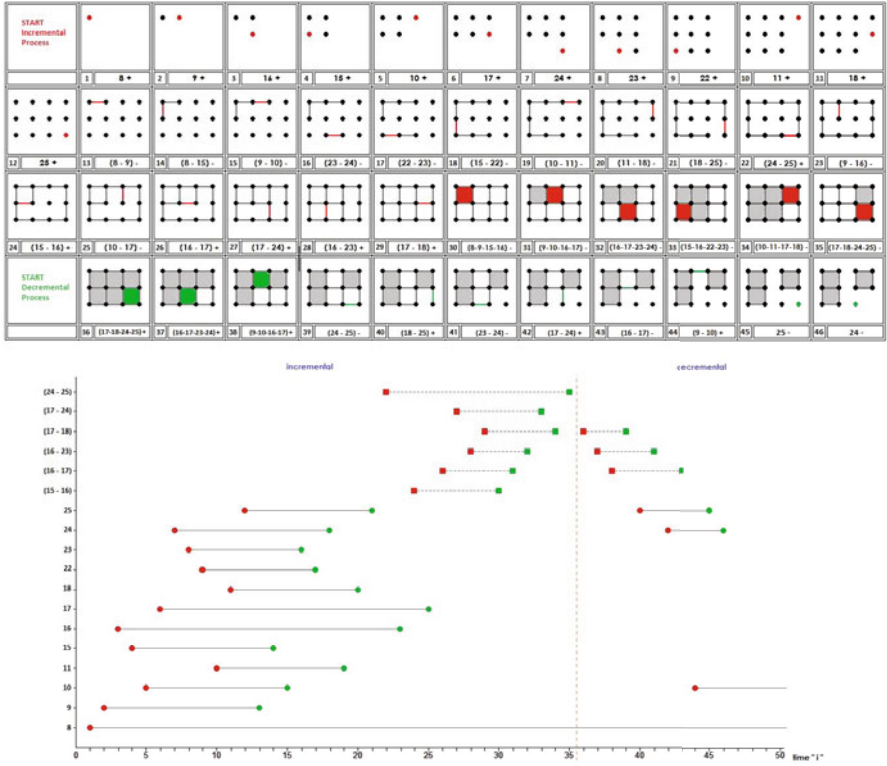


Fig. 4. Example of zig-zag filtration and barcode

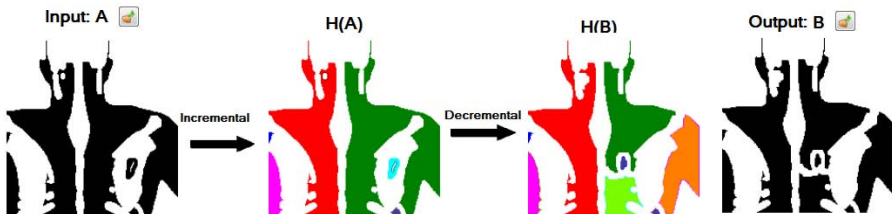


Fig. 5. Example of application. Each color refers to one connected component

6 Future Work

The proposed method is able to deal with a general filtration, allowing randomly adding or removing a cell. This is different from both standard persistence and zigzag persistence [2], which compute the filtration from a single scalar function. A correspondence between the algorithms presented here and the one given in [2] is left as a future work.

The presented algorithm is valid for any dimension but with $\mathbf{Z}/2$ domain; how can persistence for integer homology be defined? The results given in [10] may be used to try to find the answer. Since the computation of AT-models allows the computation of finer invariants than homology such as the cohomology ring [8], how could we deal with persistence of other (finer) topological invariants? We also plan to deal with the problem of extending persistence to other geometrical operations such as face removal, simplicial collapse and edge contractions using AT-models by means of the initial results given in [11].

References

1. Biasotti, S., Giorgi, D., Spagnuolo, M., Falcidieno, B.: Size functions for comparing 3D models. *Pattern Recogn.* 41(9), 2855–2873 (2008)
2. Carlsson, G., de Silva, V., Morozov, D.: Zigzag persistent homology and real-valued functions. In: *SoCG 2009*, pp. 247–256. ACM, New York (2009)
3. Chazal, F., Cohen-Steiner, D., Guibas, L., Mémoli, F., Oudot, S.: Gromov-Hausdorff Stable Signatures for Shapes using Persistence. *Comput. Graph. Forum* 28(5), 1393–1403 (2009)
4. Chazal, F., Guibas, L., Oudot, S., Skraba, P.: Analysis of scalar fields over point cloud data. In: *Proc. of SODA 2009*, pp. 1021–2030 (2009)
5. Cerri, A., Ferri, M., Giorgi, D.: Retrieval of trademark images by means of size functions. *Graph. Models* 68(5), 451–471 (2006)
6. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. In: *FOCS 2000*, pp. 454–463. IEEE Computer Society, Los Alamitos (2000)
7. Ferri, M., Stanganelli, I.: Size functions for the morphological analysis of melanocytic lesions. *Journal of Biomedical Imaging* 5, 1–5 (2010)
8. Gonzalez-Diaz, R., Real, P.: On the cohomology of 3D digital images. *Discrete Applied Math.* 147(2-3), 245–263 (2005)
9. Gonzalez-Diaz, R., Medrano, B., Real, P., Sanchez-Pelaez, A.: Simplicial perturbation techniques and effective homology. In: Ganzha, V.G., Mayr, E.W., Vorozhtsov, E.V. (eds.) *CASC 2006*. LNCS, vol. 4194, pp. 166–177. Springer, Heidelberg (2006)
10. Gonzalez-Diaz, R., Jimenez, M., Medrano, B., Real, P.: A tool for integer homology computation: Lambda-AT model. *Image and Vision Computing*, 1–9 (2008)
11. Gonzalez-Diaz, R., Jimenez, M., Medrano, B., Molina, H., Real, P.: Integral operators for computing homology generators at any dimension. In: Ruiz-Shulcloper, J., Kropatsch, W.G. (eds.) *CIARP 2008*. LNCS, vol. 5197, pp. 356–363. Springer, Heidelberg (2008)
12. Hatcher, A.: *Algebraic Topology*. Cambridge University Press, Cambridge (2002)
13. Munkres, J.: *Elements of Algebraic Topology*. Addison-Wesley Co., Reading (1984)
14. Zomorodian, A.: *Topology for Computing*. Cambridge Monographs on Applied and Computational Mathematics (16) (2005)
15. Zomorodian, A., Carlsson, G.: Computing persistent homology. *Discrete and Computational Geometry* 33(2), 249–274 (2005)