



Effective homology of k -D digital objects (partially) calculated in parallel[☆]



Raúl Reina-Molina^{a,*}, Daniel Díaz-Pernil^b, Pedro Real^a, Ainhoa Berciano^c

^a Institute of Mathematics IMUS, University of Seville, Edificio Celestino Mutis, Primera planta, Campus de Reina Mercedes. Avda. Reina Mercedes s/n, 41012 Sevilla, Spain

^b CATAM Research Group - Dept. of Applied Mathematics I, University of Seville, E. T. S. de Ingeniería Informática, Av. Reina Mercedes s/n, 41012 Sevilla, Spain

^c Department of Didactic of Mathematics and Experimental Sciences, University of the Basque Country, Escuela Universitaria de Magisterio de Bilbao, B'Sarriena s/n, 48940 Leioa, Spain

ARTICLE INFO

Article history:

Available online 11 June 2016

Keywords:

Effective Homology
Digital Object
Parallel Algorithms
Chain Contraction
Discrete Morse Theory

ABSTRACT

In [18], a membrane parallel theoretical framework for computing (co)homology information of foreground or background of binary digital images is developed. Starting from this work, we progress here in two senses: (a) providing advanced topological information, such as (co)homology torsion and efficiently answering to any decision or classification problem for sum of k -xels related to be a (co)cycle or a (co)boundary; (b) optimizing the previous framework to be implemented in using GPGPU computing. Discrete Morse theory, Effective Homology Theory and parallel computing techniques are suitably combined for obtaining a homological encoding, called algebraic minimal model, of a Region-Of-Interest (seen as cubical complex) of a presegmented k -D digital image.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, to determine in a fast and accurate way topology-related information of technologically relevant mathematical structures has become a very important question. Within the context of Digital Imagery, computational topology has acquired a significant role in fields of applications like Computer Vision, Digital Image Processing or Medical Imaging. In [18], a series of massively parallel algorithms using membrane computing are developed in order to calculate a discrete Morse complex associated to the set of foreground k -xels of a k -D binary digital image. However, nowadays there is no device capable of executing such algorithms. Hence, an optimization is required to implement them in current computers. Although the algorithms developed in this paper are directly inspired by those developed in [18], the content of current paper can be read independently.

In order to avoid the mathematically ill-posed problems of segmentation and noise which are ubiquitous in most of the image processing tasks, we will work here with a pre-segmented k -D digital image represented by k -dimensional matrix with non-negative integer values. In fact, we can restrict our work to study the topology of the set of foreground (having value 1) k -xels of a

binary image. Working with cubical cell complexes representations of digital images, we rewrite in terms of explicit chain homotopy equivalences a non-negligent part of the computational machinery underlying in discrete Morse theory [7].

Effective homology (see [21]) deals with the encoding of advanced homology information of a cell complex K in terms of chain homotopy equivalences algebraically connecting K with a much smaller chain complex M . If the homology $H(K)$ of K is torsion-free, M is $H(K)$ (seen as trivial chain complex). Otherwise, minimality is reached with an Algebraic-Minimal model (see [9]). Anyway, the previous chain homotopy equivalences provides an almost automatic answer to any interrogation of homological nature or even about homotopy invariants [21].

The main contribution of this paper is the design of an algorithm for computing an AM-model of the foreground (seen as a cubical complex) of a k -D binary image.

Various DMT-based approaches to compute discrete functions defined on cells of a cell complex have been already used. [19,20] apply DMT methods to analyze vector fields. Bauer et al. [1] computed simplified two-dimensional scalar functions while ensuring that the input function is modified by no more than a threshold d and all surviving critical point pairs have persistence greater than $2d$. Cazals et al. [3] and Lewiner et al. [13] successfully employed Forman's discrete Morse theory to compute Morse complexes of piecewise-linear functions and show applications to segmentation, visualization, and mesh compression. Delgado-Friedrichs et al. [5] also used a DMT-based formulation to

[☆] This paper has been recommended for acceptance by Mihail Gaiuanu.

* Corresponding author: Tel.: +34 670 213 187.

E-mail addresses: raulm75@gmail.com (R. Reina-Molina), sbdani@us.es (D. Díaz-Pernil), real@us.es (P. Real), ainhoa.berciano@ehu.es (A. Berciano).

<http://dx.doi.org/10.1016/j.patrec.2016.05.034>

0167-8655/© 2016 Elsevier B.V. All rights reserved.

develop an efficient algorithm for computing Morse (more concretely, Morse–Smale) complexes of large 3D data. Reininghaus et al. [20] proposed an algorithm to compute the Morse complex of 2D and 3D gray-scale digital images modeled as discrete functions on cubical complexes. The algorithm computes the Morse complex with provable guarantees on its correctness with respect to the critical cells. In [14], an algorithm for computing homology using a recursive tree-based technique for generating Morse complexes is established. Peterka et al. [15] introduced a set of building blocks for implementing parallel algorithms, which leverage high performance computing clusters. In particular, they discuss a parallel implementation of the discrete Morse theory based on the algorithm proposed in [11]. Günther [10] design a memory efficient algorithm to compute the Morse–Smale complex for 3D data and use the complex to compute persistent homology groups. The discrete gradient field is computed using a parallel variant of the method proposed by Gyulassy et al. [11] followed by an efficient computation of the boundary map representing the Morse complex. Finally, the parallel algorithm for computing the discrete gradient field given in [22] is based on a novel description of the discrete Morse function followed by a two-step algorithm computing the cells of the Morse–Smale complex. The algorithms are implemented using a hybrid multi-core implementation.

This paper is structured as follows. First, some aspects related with the required algebraic topology background are reviewed. Then, the proposed algorithms are presented and discussed. Next, a simple example is shown to illustrate the algorithms. Finally, some conclusions are given.

2. Algebraic topology background

In this section we introduce the topological background needed for designing the algorithms. First of all, we review the required combinatorial structure for the topological spaces that are used. We mainly follow the process introduced in [12] with some minor changes. Specifically, Kaczyński et al. uses closed cubes as main combinatorial objects while we use open cubes.

A cubical cell σ is a finite product of intervals:

$$\sigma = I_1 \times \dots \times I_k \subset \mathbb{R}^k$$

where I_j is an interval with integer extremes $(m_j, m_j + 1)$ and length one or the singleton $\{m_j\}$, denoted as (m_j) , for each $j \in \{1, \dots, k\}$. The interval I_j is referred as the j th component of σ and denoted by $I_j(\sigma)$. The set of all cubical cells in \mathbb{R}^k is denoted by \mathcal{K}^k . The set of all cubical cells is

$$\mathcal{K} = \bigcup_{k=1}^{\infty} \mathcal{K}^k \quad (1)$$

Given a cubical cell σ in \mathbb{R}^k , its *embedding number* k is denoted by $\text{emb } \sigma$. The dimension of σ is defined as the number of unitary intervals in its expression as product of intervals and is denoted by $\text{dim } \sigma$. The set of all elementary cubes with dimension p is denoted by \mathcal{K}_p . The set of all elementary cubes in \mathbb{R}^k with dimension p is denoted by \mathcal{K}_p^k . Whenever the dimension of a cubical cell require to be made explicit, it is denoted as a superscript between parenthesis. Therefore, $\text{dim } \sigma^{(p)} = p$. We also explicitly indicate the dimension of a cell $\sigma^{(p)}$ writing that σ is a p -cell.

The closure¹ of a cubical cell can be decomposed into the union of lower-dimensional cubical cells. If δ and σ are two cubical cells in \mathbb{R}^k of any dimension and $\delta \subset \bar{\sigma}$, then δ is a *face* of σ and is denoted as $\delta \leq \sigma$. If δ is a face of σ and $\delta \neq \sigma$, then δ is a *proper face* of σ , denoted as $\delta < \sigma$. If δ is a face of σ and $\text{dim } \delta = \text{dim } \sigma -$

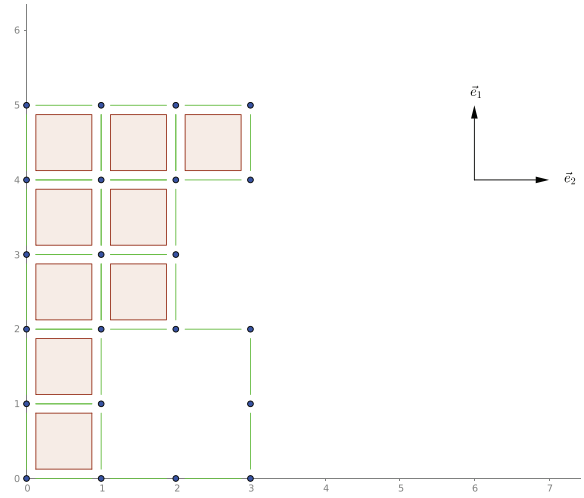


Fig. 1. Cubical complex K_0 .

1 then δ is a *primary face* of σ , denoted by $\delta \in \partial\sigma$. Therefore, the set of primary faces of a cell σ will be denoted as $\partial\sigma$

As an example, let σ be the open square $(0, 1) \times (3, 4)$, a 2-dimensional cubical cell in \mathbb{R}^2 . Its closure is the square $\bar{\sigma} = [0, 1] \times [3, 4]$ and is decomposed as follows

$$\begin{aligned} [0, 1] \times [3, 4] &= (0, 1) \times (3, 4) \cup \\ &\quad (0, 1) \times (3) \cup (0, 1) \times (4) \cup \\ &\quad (0) \times (3, 4) \cup (1) \times (3, 4) \cup \\ &\quad (0) \times (3) \cup (0) \times (4) \cup \\ &\quad (1) \times (3) \cup (1) \times (4) \end{aligned}$$

Namely, the interior of the square, its four edges and its four vertices. The set of primary faces is given by

$$\begin{aligned} \partial(0, 1) \times (3, 4) &= \{(0, 1) \times (3), (0, 1) \times (4), (0) \\ &\quad \times (3, 4), (1) \times (3, 4)\} \end{aligned}$$

A *cubical complex* is a collection K of cubical cells with the same embedding number and such that, for every cubical cell $\sigma \in K$, all of its primary faces are in the complex. We denote the set of p -cells in K as K_p . In Fig. 1 a cubical complex in \mathbb{R}^2 is showed.

As usual, we define the chain complex $C(K)$ as the graded \mathbb{Z} -module $\{C_p(K)\}_{p \in \mathbb{Z}}$ where $C_p(K)$ is the free abelian \mathbb{Z} -module generated by the cubical p -cells in K . Namely, a p -chain is a sum of cells multiplied by some integer number. The boundary map d in the chain complex is defined in [12]. For any cubical cell σ , $d\sigma$ is an alternating sum over its faces and such that $d \circ d = 0$. As usual, let $Z_p = \ker d_p$ the subgroup of cycles and $B_p = \text{Im } d_{p+1}$ the subgroup of boundaries. The p th cubical homology group of K is the quotient subgroup

$$H_p(K) = \frac{Z_p K}{B_p(K)}$$

Namely, a chain is a generator of homology if it is a cycle (has no boundary) and is not the boundary of any other cycle. The cubical homology of K is the collection $H_*(K) = \{H_p(K)\}_{p \in \mathbb{Z}}$. Recall that, for $p < 0$ or $p > \text{dim } K$, is $H_p(K) = 0$.

The cubical complex in Fig. 1 has homology groups given by:

$$H_0(K) \cong \mathbb{Z}, \quad H_1(K) \cong \mathbb{Z}, \quad H_n(K) \cong 0 \quad \text{for } n \notin \{0, 1\}$$

For a review in some algorithms for calculating homology groups, interested readers would read Chapter 3 in [12].

As an example, the chain

$$(4) \times (0, 1) - (3) \times (0, 1) + (3, 4) \times (0) - (3, 4) \times (1)$$

is a cycle but it is not a generator of homology as it is the boundary of $(3, 4) \times (0, 1)$.

¹ Given a set $A \subset \mathbb{R}^k$, its closure is $\bar{A} = \{p \in \mathbb{R}^k : \forall \epsilon > 0 \exists q \in A | d(p, q) < \epsilon\}$. Informally, the closure of a set is the set itself together with the points that are “infinitely” near to the set.

On the other hand, the chain

$$(0, 1) \times (1) + (1, 2) \times 1 + (2) \times (1, 2) + (2) \times (2, 3) - (1, 2) \times (3) - (0, 1) \times (3) - (0) \times (2, 3) - (0) \times (1, 2)$$

is a cycle and is not the boundary of any other chain, as it would be the boundary of

$$(0, 1) \times (1, 2) + (1, 2) \times (1, 2) + (0, 1) \times (2, 3) + (1, 2) \times (2, 3)$$

and this chain are not in $C_2(K)$ because its cells do not belong to K .

2.1. Discrete Morse theory

We need to recall now notions and results of discrete Morse theory. A real-valued smooth map defined over a compact k -manifold is a *Morse function* if all its critical points have non-singular Hessian matrix and no two critical points have the same function value [11]. Morse functions allow to endow the manifold with a cellular structure.

A CW complex is, roughly speaking, a topological space endowed with a decomposition into smaller pieces called cells, being homeomorphic to Euclidean spheres. A formal definition can be found in [7].

Discrete Morse theory introduced by Forman [7] adapts Morse theory to CW complexes instead of smooth manifolds. Although discrete Morse theory has been developed for finite regular CW complexes, we restrict our research to finite cubical complexes, which are a particular type of CW complexes.

Let K be a cubical complex and $f : K \rightarrow \mathbb{R}$ a function that assigns scalar values to every cell of K . f is a *discrete Morse function* if, for every cubical cell $\sigma^{(p)} \in K$ the following conditions hold:

- $\#\{\tau^{(p+1)} \in K | \tau > \sigma \wedge f(\tau) \leq f(\sigma)\} \leq 1$. I.e., there is at most one facet² of σ where f takes in it a lower value than it does on σ .
- $\#\{\mu^{(p-1)} \in K | \mu < \sigma \wedge f(\mu) \geq f(\sigma)\} \leq 1$. I.e., there is at most one face of σ where f takes in it a greater value than it does on σ .

where $\#A$ denotes the number of elements of the (finite) set A .

A discrete Morse function f can be thought as a discrete function that is increasing with respect to cell dimension except for, at most, two cells for each cell. Concretely, for each cell σ there is at most one face μ ($\mu \in \partial\sigma$) with $f(\mu) \geq f(\sigma)$ and there is at most one facet τ ($\sigma \in \partial\tau$) with $f(\tau) \leq f(\sigma)$.

A cubical cell $\sigma^{(p)}$ is *critical* if one of the following conditions hold:

- $\#\{\tau^{(p+1)} \in K | \tau > \sigma \wedge f(\tau) \leq f(\sigma)\} = 0$
- $\#\{\mu^{(p-1)} \in K | \mu < \sigma \wedge f(\mu) \geq f(\sigma)\} = 0$

Given a cubical cell σ its *barycenter* is the point

$$b(\sigma) = \left(\frac{\inf I_1(\sigma) + \sup I_1(\sigma)}{2}, \dots, \frac{\inf I_k(\sigma) + \sup I_k(\sigma)}{2} \right)$$

As the closure of a cubical cell is convex, its barycenter is an interior point.

We define a *discrete vector* as a pair of incident cells $\{\sigma^{(p)} < \tau^{(p+1)}\}$. Vectors are represented as arrows from the barycenter of the lower dimension cell to barycenter of the higher dimension cell. A *discrete vector field* V on K is a collection of vectors in K such that each cubical cell in the vector belongs to, at most, one pair of V .

Formally, a discrete vector field is a map $V: K \rightarrow K \cup \{0\}$ such that:

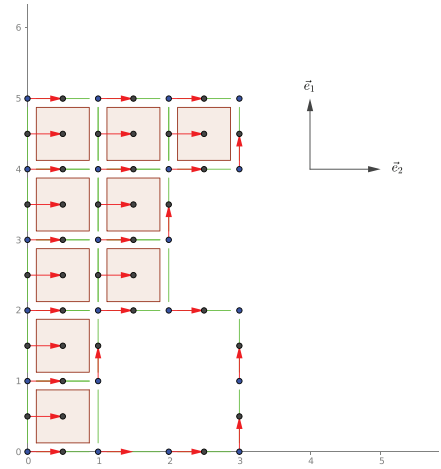


Fig. 2. Vector field V_0 in K_0 .

1. for each $\sigma \in K$, if $V(\sigma) \neq 0$ then $\dim V(\sigma) = \dim \sigma + 1$
2. for each $\sigma \in K_p$, either $V(\sigma) = 0$ or $\sigma \in \partial V(\sigma)$
3. if $\sigma \in \text{Im}(V)$ then $V(\sigma) = 0$
4. for each $\sigma \in K_p$

$$\#\{\mu^{(p-1)} \in K | V(\mu) = \sigma\} \leq 1$$

In Fig. 2, a vector field (red arrows) in the cubical complex in Fig. 1 is showed.

Given a discrete vector field V on K , a V -path of dimension p , γ , is a sequence of cubical p -cells $\sigma_0, \sigma_1, \sigma_2, \dots, \sigma_r$ such that

1. If $V(\sigma_i) = 0$, then $\sigma_{i+1} = \sigma_i$.
2. If $V(\sigma_i) \neq 0$, then $\sigma_{i+1} < V(\sigma_i)$ with $\sigma_{i+1} \neq \sigma_i$.

The set of V -paths is denoted as $\Gamma(V)$.

For example, $\gamma = (0) \times (0), (0) \times (1), (0) \times (2), (0) \times (3)$ is a V -path of dimension 0, where V is the vector field in Fig. 2.

A V -path $\gamma = \sigma_0, \sigma_1, \dots, \sigma_r$ is called *closed* if $\sigma_r = \sigma_0$ and is called *non-stationary* if $\sigma_1 \neq \sigma_0$. A vector field V is called *acyclic* if there is no non-stationary closed V -paths. As an example, the vector field in Fig. 2 is, indeed, an acyclic vector field.

Given a discrete Morse function, a special discrete vector field called *discrete gradient vector field* is defined so that $f(V(\sigma)) \leq f(\sigma)$. As shown in [7], a cubical complex can be transformed into another homotopically equivalent following a series of simplicial collapses³, where each of them collapses both cubical cells in each vector in the corresponding discrete gradient vector field.

Recall that vectors give some kind of dynamic to the cells. Namely, one can (imaginary) move from one cell σ to another one μ if $\mu \in \partial V(\sigma)$. In the previous example, one can move from cell $(0) \times (0)$ to $(0) \times (1)$ as $V((0) \times (0)) = (0) \times (0, 1)$ and $(0) \times (1)$ is in the boundary of the later.

We recall here one of the main results of discrete Morse theory.

Theorem 2.1 [7] 9.3. *A discrete vector field V is the gradient vector field of some discrete Morse function if and only if there are no non-stationary closed V -paths.*

A vector field V is extended to a graded group homomorphism $V : C_p(K) \rightarrow C_{p+1}(K)$ such that

$$V(\sigma^{(p)}) = \begin{cases} \pm \tau^{(p+1)} & \text{if } \{\sigma < \tau\} \text{ is a vector in } V \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where the sign is chosen so that $\langle \sigma, \partial V(\sigma) \rangle = +1$.

³ In [7], Forman works with simplicial complexes, however the mathematical scaffolding provided by DMT can be settled with no change to finite cubical complexes.

² A cell with σ in its boundary.

The reduced (discrete time) flow map, denoted as φ is defined by

$$\varphi = \text{id} - \partial \circ V \quad (3)$$

The reduced flow map associates a $(p+1)$ -chain $\varphi(\sigma)$ to any p -cell σ , such that $\partial V(\sigma) = \sigma - \varphi(\sigma)$. Hence, $\langle \sigma, \varphi(\sigma) \rangle = 0$. The map φ establishes a way of following the flow determined by the vector field. Concretely, if there is a V -path of length r from a p -cell σ to another p -cell σ' , then $\langle \varphi^r(\sigma), \sigma' \rangle \neq 0$. This fact can be proved by induction in the length of the V -path.

Please note that the reduced flow map encodes the dynamic idea presented before. If a cell μ is present in the chain $\varphi(\sigma)$, then σ, μ is a valid V -path.

Definition 2.2. Let K be a cubical complex and f be a Morse discrete function on K . Let $C_p(K)$ denote the p -chains on K and $M_p \subseteq C_p(K)$ the span⁴ of the critical p -cells. The graded group $M = \{M_p\}_{p \in \mathbb{Z}}$ is called the space of Morse chains.

In [7], Forman makes use of the set of V -paths to build the boundary map of the Morse complex associated to a given complex and an acyclic vector field. It is shown in the following result.

Theorem 2.3 [7] 7.1. *There are boundary maps $d_M : M_* \rightarrow M_{*+1}$ so that*

$$d_M \circ d_M = 0$$

and such that the differential complex

$$0 \xrightarrow{(d_M)_{k+1}} M_k \xrightarrow{(d_M)_k} M_{k-1} \xrightarrow{(d_M)_{k-1}} \dots \xrightarrow{(d_M)_2} M_1 \xrightarrow{(d_M)_1} M_0 \xrightarrow{(d_M)_0} 0$$

calculates de homology of K . That is, if we define

$$H_p(M, d_M) = \frac{\ker (d_M)_p}{\text{Im} (d_M)_{p+1}}$$

then, for each p

$$H_p(M, d_M) \cong H_p(K).$$

2.2. From DMT to effective homology

Our modus operandi for not only reaching homology torsion information but also for fully control it is to translate DMT to an algebraic context of chain contractions.

Definition 2.4 [6]. Given two chain complexes $C_* = \{C_q, d_q\}$ and $C'_* = \{C'_q, d'_q\}$, a chain map f is a family of homomorphisms $\{f_q : C_q \rightarrow C'_q\}$ such that $d'_q \circ f_q = f_{q-1} \circ d_q$.

A chain contraction from C_* to C'_* is a triple (π, ι, ϕ) of chain maps $\pi : C_* \rightarrow C'_*$ (projection), $\iota : C'_* \rightarrow C_*$ (inclusion) and $\phi : C_* \rightarrow C_{*+1}$ (chain homotopy or integral operator) that satisfy the following conditions:

- $\text{id}_C - \iota \circ \pi = d \circ \phi + \phi \circ d$
- $\pi \circ \iota = \text{id}_{C'}$
- $\pi \circ \phi = 0$
- $\phi \circ \iota = 0$
- $\phi \circ \phi = 0$

This is a classical notion in homological algebra and algebraic topology: see e.g. Section 12 of [6] and comments on the terminology and applications in [9], p.86. Note that because of the condition (a), if there exists a chain contraction from C_* to C'_* then their homology modules are isomorphic.

Given two chain contractions $(\pi_1, \iota_1, \phi_1) : C_*^1 \rightarrow C_*^2$ and $(\pi_2, \iota_2, \phi_2) : C_*^2 \rightarrow C_*^3$, the composition

$$(\pi, \iota, \phi) = (\pi_2, \iota_2, \phi_2) \circ (\pi_1, \iota_1, \phi_1) : C_*^1 \rightarrow C_*^3$$

is given by:

$$\pi = \pi_2 \circ \pi_1 \quad (4)$$

$$\iota = \iota_1 \circ \iota_2 \quad (5)$$

$$\phi = \phi_1 + \iota_1 \circ \phi_2 \circ \pi_1 \quad (6)$$

In the proof of Theorem 7.3 in [7], the argument used to demonstrate that the homology groups of a cell complex and its Morse complex are isomorphic consists in the creation of a chain contraction from the chain complex (associated to the initial cell complex) to the Morse complex (associated to a given acyclic vector field V). This chain contraction is explicitly given by:

$$\phi = \left(V \circ \sum_{p=0}^N \varphi^p \right) : C_* \rightarrow C_{*+1} \quad (7)$$

$$\pi = [(\text{id}_C - d \circ \phi - \phi \circ d) \cdot \chi_{\text{crit}_V}] : C_* \rightarrow M_* \quad (8)$$

$$\iota = (\text{id}_C - d_C \circ \phi - \phi \circ d_C) : M_* \rightarrow C_* \quad (9)$$

where N is a non negative integer number such that, $\forall n > N : V \circ \varphi^n = 0$ and χ_{crit_V} is the characteristic function⁵ of critical cells of V extended by linearity to the chain complex. Recall that the set of critical p -chains M_p is a subset of the set of chains C_p for any $p \in \mathbb{Z}$.

Please note that the Morse complex is the set of critical chains endowed with the Morse differential given by

$$d_M = \pi \circ d_C \circ \iota \quad (10)$$

The classical method for calculating the homology of a cell complex involves reducing the matrices of the differential morphisms to its Smith normal form, which consists in finding two bases which respect to them the matrix of the differential is 0 outside the main diagonal and the main diagonal is $a_{11}|a_{22}| \dots |a_{pp}. 0, \dots, 0$ where $|$ means “divides to” and $a_{ii} \neq 0$. The later matrix is called the Smith normal form of the differential and is represented as SNF for short.

Definition 2.5 (AM Model, [9]). An algebraic minimal model, or an AM model for short, of a cell complex K is a chain contraction from $C_*(K)$ to a chain complex (M, d') such that each M_p is a free \mathbb{Z} -module and all the non-zero elements in the SNF of each d_q are non-invertible.

Recall that, in \mathbb{Z} the only invertible elements are ± 1 .

An AM model exists for every cell complex, and any two AM models for the same complex are isomorphic. An algorithm computing an AM-model for a given cell complex is designed in [9]. An efficient sequential implementation of an AT-model and AM-model for a given cubical complex is shown in [16].

3. Algorithm

Our algorithm consists on several stages. All of them are presented below.

Stage 1 Build the cubical complex associated to the foreground of a binary image. Recall that a binary image can be represented as a function $I : \mathbb{Z}_n^k \rightarrow \{0, 1\}$. The cubical complex is built using the following fact:

A cell $(a_1, b_1) \times (a_2, b_2) \times \dots \times (a_k, b_k)$ with $0 \leq b_i - a_i \leq 1$ is in the cell complex if and only if

$$\{a_1, b_1\} \times \{a_2, b_2\} \times \dots \times \{a_k, b_k\} \subseteq I^{-1}(1)$$

Namely, all the vertices in the cubical cell belongs to points in the image (usually represented as black pixels).

⁴ Linear combinations of the critical cells.

⁵ The characteristic function χ_A of a given set A is the function that values 1 in A and 0 outside.

This step can be executed in parallel in one computation step. Concretely, all the possible cells in the cell complex, which is finite like the source image, are all checked at the same time.

Algorithm 1 Acyclic cubical vector field for a cubical complex.

```

1: function ACYCLICCUBICALVECTORFIELD( $K$ )
2:    $V \leftarrow (0 : \mathbb{Z}[K_*] \rightarrow \mathbb{Z}[K_{*+1}])$ 
3:   for  $p \leftarrow n, 1$  do
4:     for all  $\sigma \in K$  parallel do
5:        $\tau \leftarrow \sigma \oplus \bar{e}_p$ 
6:       if  $\tau \neq \sigma \wedge \tau \in K$  then
7:          $q \leftarrow \dim \sigma$ 
8:          $dom_\sigma \leftarrow \sum_{\mu \in K_{q+1}} |V_q[\mu, \sigma]|$ 
9:          $dom_\tau \leftarrow \sum_{\mu \in K_{q+2}} |V_{q+1}[\mu, \tau]|$ 
10:         $img_\sigma \leftarrow \sum_{\mu \in K_q} |V_q[\sigma, \mu]|$ 
11:         $img_\tau \leftarrow \sum_{\mu \in K_{q+1}} |V_{q+1}[\tau, \mu]|$ 
12:        if  $dom_\sigma + dom_\tau + img_\sigma + img_\tau = 0$  then
13:           $V_q(\sigma) \leftarrow \langle \sigma, d(\tau) \rangle \tau$ 
14:        end if
15:      end if
16:    end for
17:  end for
18:  return  $V$ 
19: end function

```

- Stage 2 Create an acyclic vector field V within the cubical complex K . This stage finishes in $embK$ parallel steps. See Algorithm 1. This stage is the only one requiring that K needs to be a cubical complex. If other kind of cell complex is given, e.g. simplicial or CW, the construction of the acyclic vector field is the only step to be changed. The rest of the stages remain unchanged as they do not depend on the geometry of the cell complex but on its associated chain complex, which is an algebraic object.
- Stage 3 Calculate the chain contraction $(\pi_V, \iota_V, \phi_V) : C(K) \rightarrow M_V(K)$ associated to the vector field V . This stage requires, at most, $\lceil \log_2 \lambda_V \rceil$ steps (where λ_V is the length of the longest V -path).
- Stage 4 Create another acyclic vector field in the Morse complex of the previous vector field.
- Stage 5 Repeat the two previous stages until the differential of the Morse complex associated to the simplified vector field has no invertible elements in its matrix.
- Stage 6 If the Morse complex of the last reduced vector field has some non zero element in its matrix representation, calculate the AM-model for it.
- Stage 7 Compose all the chain contractions associated to the vector fields previously calculated. This composition represents, in fact, an effective homology of cubical complex K .

3.1. Acyclic vector field on a cubical complex

Let K be a cubical complex. We suppose, without loose of generality, that there exists a non negative integer n , such that for every cell σ in K

$$0 \leq \inf I_p(\sigma) \leq \sup I_p(\sigma) < n \quad (1 \leq p \leq emb K) \quad (11)$$

Let \bar{e}_p denote the p th vector in the orthonormal canonical base of $\mathbb{R}^{emb K}$. We define the *right shift* of a cell $\sigma = I_1 \times I_2 \times \dots \times I_{emb K}$ by $\sigma \oplus \bar{e}_p = I'_1 \times I'_2 \times \dots \times I'_{emb K}$ where

$$I'_k = \begin{cases} (\inf I_p, 1 + \inf I_p) & \text{if } k = p \wedge \inf I_p = \sup I_p \\ I_k & \text{otherwise} \end{cases}$$

For example, $(1, 2) \times (2, 3) \times (2) \oplus \bar{e}_2 = (1, 2) \times (2, 3) \times (2)$ as the second interval is non degenerated, and $(1, 2) \times (2, 3) \times (2) \oplus \bar{e}_3 = (1, 2) \times (2, 3) \times (2, 3)$. Roughly speaking, shifting a cell in a (principal) direction consists on extending the cell along this direction.

Recall that $\mathbb{Z}[A]$ is the free \mathbb{Z} -module with basis the elements of the set A and $V[\mu, \sigma]$ is the coefficient of the chain $V(\sigma)$ corresponding to cell μ . Please note that, essentially, a morphism between two free modules can be represented as a matrix, hence the later notation gives direct access to the elements of that matrix.

The Algorithm 1 works as follows. In line 2 a null chain map of degree +1 is created. The “for” loop in line 3 sequentially iterates over the vectors belonging to the canonical orthonormal base for the embedding space of the cubical complex. The parallel “for” in line 4 iterates in parallel over all the cells in the cubical complex. In this situation, at most at theoretical level, the algorithm will treat every cell at the same time. In line 5 the right shift of every cell along the current vector in the canonical base is computed. The condition at line 6 ensures that the shift is well defined (i.e., it returns a facet inside the cubical complex). In lines 8 through 11, we calculate if σ or τ belongs to the domain or the image of the (previously calculated) vector field. The condition at line 12 ensures that the vector field is well defined, i.e. every cell belongs, at most, to only one vector in the vector field. Finally, line 13 creates the corresponding vector with the correct sign. Recall that the incidence of σ in $V(\sigma)$ is the same as the incidence of σ in τ .

Recall that all the vectors in the vector field follow one of the positive directions in the canonical orthonormal base in $\mathbb{R}^{emb K}$. This ensures that the vector field is acyclic as we avoid the possibility of turning back and, hence, preventing the generation of cycles.

The vector field showed in Fig. 2 is computed using Algorithm 1.

3.2. Chain contraction from vector field

Let V be an acyclic vector field over a chain complex (C_*, d_*) . We will compute the chain contraction associated to V defined in 2.2 using Algorithm 2, whose correctness is based on the Proposition 3.1.

Algorithm 2 Chain contraction associated to an acyclic vector field.

```

1: function CHAINCONTRACTION( $V$ )
2:    $\varphi \leftarrow id - d \circ V$ 
3:    $S \leftarrow id + \varphi$ 
4:    $P \leftarrow \varphi$ 
5:   while  $V \circ P \neq 0$  do
6:      $S \leftarrow (S - id) \circ (P + id) + id$ 
7:      $P \leftarrow P \circ P$ 
8:   end while
9:    $\phi \leftarrow V \circ S$ 
10:   $\pi \leftarrow id - d \circ \phi - \phi \circ d$ 
11:   $\iota \leftarrow (id : Crit_* \rightarrow C_*)$ 
12:  return  $(\pi, \iota, \phi)$ 
13: end function

```

$\triangleright Crit_p$ is the spanning of critical p -cells

Proposition 3.1. *Let φ be the reduced flow of an acyclic vector field in a chain complex (C_*, d_*) . The integral operator $\phi = \sum_{q \geq 0} (V \circ \varphi^q)$ can be calculated as $\phi = V \circ S^{(N)}$ for N higher enough, where*

$$S^{(m)} = \sum_{q=0}^{2^m} \varphi^q \quad (12)$$

Furthermore, a “logarithmic” way of calculating ϕ is given by the following recursive equation:

$$\begin{cases} S^{(0)} = \text{id} + \varphi \\ S^{(m+1)} = (S^{(m)} - \text{id}) \circ (\varphi^{2^m} + \text{id}) + \text{id} \end{cases} \quad (13)$$

Proof. Let N' be an integer satisfying Theorem 7.2 in [7]. Then, for all $m \geq N'$, is $\sum_{q=0}^m (V \circ \varphi^q) = \sum_{q=0}^{N'} (V \circ \varphi^q)$.

Furthermore, it is straightforward to prove that

$$S^{(m+1)} = S^{(m)} + \varphi^{2^m} \circ (S^{(m)} - \text{id}) = (S^{(m)} - \text{id}) \circ (\varphi^{2^m} + \text{id}) + \text{id}$$

Last equation allows us to calculate ϕ in logarithmic time with respect to the number of summands. It is enough to take $N = \lceil \log_2(N') \rceil$. \square

The geometric–combinatorial approach of Forman to Morse theory using acyclic vector fields, allows us to represent the vector field as arrows and the flow of the vector field as “path” of cells following those arrows. This point of view ensures that the maximum number of summands in the ordinary expression of ϕ is, at most, the length of the longest V -path. Hence, previous result leads us to ensure that the integral operator can be calculated in $\lceil \log_2(\lambda) \rceil$ iterations where λ is the length of the longest V -path.

3.3. Vector field simplification

The acyclic vector field built in Algorithm 1 is far from being combinatorially optimal⁶. In this stage we follow the reasoning in Section 11 of [7] and Section 9 of [8]. Both sections are devoted to the simplification of a given acyclic vector field in order to reduce the amount of critical cells.

Theorem 3.2 (11.1 [8]). *Suppose f is a discrete Morse function on K such that $\beta^{(p+1)}$ and $\alpha^{(p)}$ are critical, and there is exactly one gradient path from the boundary β to α .*

Then there is another Morse function g on K with the same critical cells except that α and β are no longer critical. Moreover, the gradient vector field associated to g is equal to the gradient vector field associated to f except along the unique gradient path from the boundary β to α .

Recall that $\text{dom}V = \{c \in C_* : V(c) \neq 0\}$ and χ_A is the characteristic function of the set A , i.e.

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Note also that function $\text{IsCRITICAL}(V, \sigma)$ returns true whenever σ is a critical cell in V .

The validation of Algorithm 3 relies on the following observation: condition at line 13 checks if there is only one V -path from the boundary of a critical cell τ to another critical cell σ . This instruction is in agreement with Theorem 3.2. Please note that loop at line 6 computes the sum of powers of h , and h can be considered as the adjacency matrix of the graph whose vertices are cells of the cell complex and whose edges are determined by the following relation: there is an edge between two different nodes σ and σ' if and only if $\langle \sigma', \varphi(\sigma) \rangle \neq 0$.

4. Example

In this section we present a simple example to illustrate the behavior of the previous set of algorithms. Let us start with the cubical complex K_0 in Fig. 1 in page 2. Algorithm 1 computes an acyclic vector field V_0 in K_0 as showed in Fig. 2 in page 3. Please

⁶ Recall that an acyclic vector field is combinatorially optimal if there is no other acyclic vector field with less critical cells.

Algorithm 3 Vector field simplification.

```

1: function SIMPLIFIEDVECTORFIELD(V)
2:   V' ← V
3:   h ← φ · χdomV
4:   t ← h
5:   M ← MORSECOMPLEX(V)
6:   while V ◦ t ≠ 0 do
7:     h ← (idC + t) ◦ h
8:     t ← t ◦ t
9:   end while
10:  for τ ∈ K(M) do
11:    for σ ∈ ∂M(τ) do
12:      PC ← ∑μ ∈ ∂C(τ) ⟨h(μ), σ⟩
13:      if PC = 1 ∧ |(dM(τ), σ)| = 1 ∧ IsCRITICAL(V', σ) ∧
IsCRITICAL(V', τ) then
14:        V'(σ) ← τ
15:      end if
16:    end for
17:  end for
18:  return V'
19: end function

```

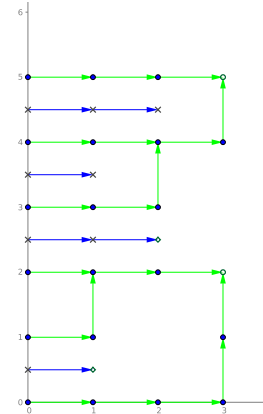


Fig. 3. Reduced flow φ_0 for V_0 .

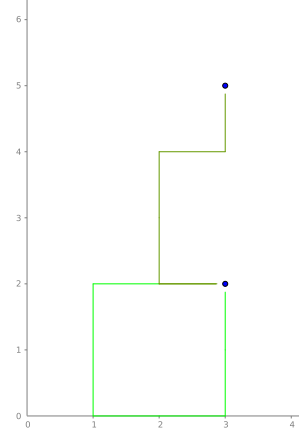


Fig. 4. Morse complex M_1 .

note how the vectors follow the positive directions provided by the canonical orthonormal basis.

An useful way of describing the dynamic of the vector field is by its reduced flow φ . This fact is illustrated in Fig. 3, where only 0-cells and 1-cells are represented, as 2-cells have no facets in \mathbb{R}^2 . In this figure we represents cells as vertices of a directed graph. Critical cells $(2) \times (3)$, $(5) \times (3)$, $(0, 1) \times (1)$ and $(2, 3) \times (2)$ are distinguished.

The matrix

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

represents the homomorphism φ_0 .

The Morse complex K_1 associated to V_0 is presented in Fig. 4, where both critical 0-cells and both critical 1-cells are shown. This figure also shows that the cell in K_1 corresponding to $(0, 1) \times (1)$ represents a cycle which is not a boundary (an homology generator representative) and the cell corresponding to $(2, 3) \times (2)$ has boundary given by $(5) \times (3) - (2) \times (3)$.

Therefore, we have some entries in the differential⁷ of the Morse complex that are invertible in \mathbb{Z} and we must simplify the vector field. The result is illustrated in Fig. 5.

The final complex which combinatorially represents the homology of K_0 is given by K_2 , the Morse complex of V_1 and is showed in Fig. 6, where we can see the expected result: one generator for H_0 (the only connected component) and one generator for H_1 (the only hole).

Finally, in Fig. 7 the composition of the corresponding chain contractions is represented as an optimized acyclic vector field in K_0 .

5. Conclusions and further work

In this paper, the bio-inspired theoretical framework of [18] is optimized in order to be implemented using GPGPU computing. Taking as input a ROI K of a pre-segmented k -D digital image, the final output of this effective homological approach of discrete Morse theory is an AM-model for a cubical complex version of the ROI. To efficiently compute advanced topological information (both homological and homotopical) from this AM-model is possible thanks to its parallel algebraic (as a chain contraction) and

combinatorial (as an acyclic vector field) encoding. Effective homology encodings of digital objects allows to describe the global topology phenomenology of them in terms of homology generators and relations between them [2].

In fact, such AM-model representations shows how the original digital object can be algebraically transformed into a simpler one. This last cell complex presents one cell for each torsion-free homology generator and two cells for each homology torsion generator. AM-models also allow to efficiently answer to any decision or classification problem for a sum of k -xels at (co)boundary, (co)cycle or (co)homology level.

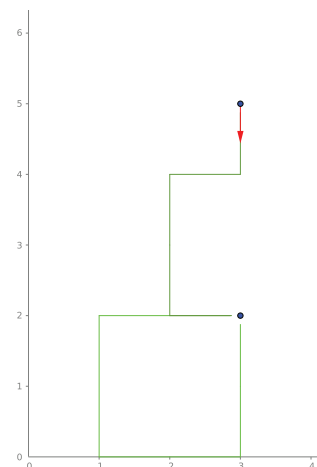


Fig. 5. Simplified vector field V_1 in K_1 .

⁷ Concretely, the matrix of the differential $(d_{K_1})_1$ is $\begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix}$

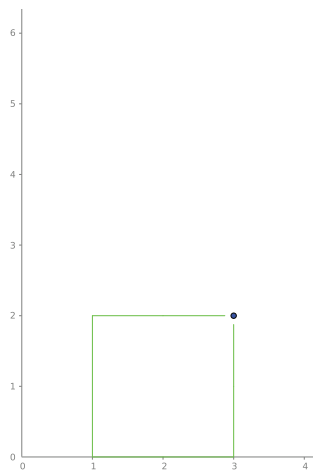


Fig. 6. Morse complex M_2 combinatorially representing the homology of K_0 .

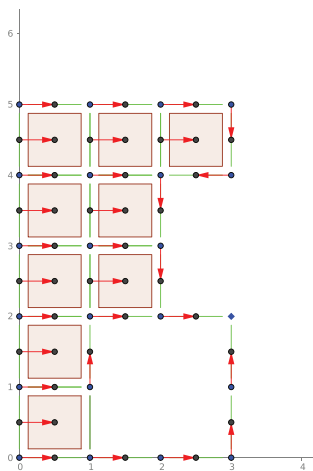


Fig. 7. Optimized acyclic vector field in K_0 .

The algorithm presented here can be used in many image processing applications that involves counting or identifying holes at any dimension. For example, counting cavities in trabecular bone can be used as a measure of osteoporosis level (see [23]), or topological trees or region-adjacency-graphs are used for segmentation of skin lesions in dermatoscopic images (see [4]).

An implementation of this algorithmic theoretical study have already been done using Python and NVidia CUDA (see [17] for source code). However, the memory consumption grows exponentially with image dimension, enormously limiting the computation power of the software.

In a near future, some algorithms must be developed in order to break the original image in smaller pieces and, after com-

puting an AM-model for every piece, appropriately “glue” all that information for calculating an AM-model of the whole original image.

References

- [1] U. Bauer, C. Lange, M. Wardetzky, Optimal topological simplification of discrete functions on surfaces, *Discrete Computat. Geom.* 47 (2012) 347–377.
- [2] A. Berciano, H. Molina-Abril, P. Real, Searching high order invariants in computer imagery, *Appl. Algebra Eng. Commun. Comput.* 23 (2012) 17–28.
- [3] F. Cazals, F. Chazal, T. Lewiner, Molecular shape analysis based upon the morse-smale complex and the connolly function, *Proceedings of the Nineteenth Annual Symposium on Computational Geometry, ACM, 2003*, pp. 351–360.
- [4] R. Cucchiara, C. Grana, S. Seidenari, G. Pellacani, Exploiting color and topological features for region segmentation with recursive fuzzy c-means, *Mach. Graph. Vis.* 11 (2002) 169–182.
- [5] O. Delgado-Friedrichs, V. Robins, A. Sheppard, Skeletonization and partitioning of digital images using discrete morse theory, *Pattern Anal. Mach. Intell. IEEE Trans.* 37 (2015) 654–666.
- [6] S. Eilenberg, S. MacLane, On the groups $h(\pi, n)$, ii: Methods of computation, *Ann. Math.* (1954) 49–139.
- [7] R. Forman, Morse theory for cell complexes, *Adv. Math.* 134 (1998) 90–145.
- [8] R. Forman, A user's guide to discrete morse theory, *Sém. Lothar. Combin.* 48 (2002) 35pp.
- [9] R. Gonzalez-Diaz, P. Real, et al., Computation of cohomology operations of finite simplicial complexes, *Homol. Homotopy Appl.* 5 (2003) 83–93.
- [10] D. Günther, J. Reininghaus, I. Hotz, H. Wagner, Memory-efficient computation of persistent homology for 3d images using discrete morse theory, *Proceedings of the 24th SIBGRAPI Conference on Graphics, Patterns and Images (Sibgrapi)*, IEEE, 2011, pp. 25–32.
- [11] A. Gyulassy, P.T. Bremer, B. Hamann, V. Pascucci, A practical approach to morse-smale complex computation: Scalability and generality, *Vis. Comput. Graph. IEEE Trans.* 14 (2008) 1619–1626.
- [12] T. Kaczynski, K.M. Mischaikow, M. Mrozek, *Computational Homology*, 157, Springer Science & Business Media, 2004.
- [13] T. Lewiner, H. Lopes, G. Tavares, Applications of forman's discrete morse theory to topology visualization and mesh compression, *Vis. Comput. Graph. IEEE Trans.* 10 (2004) 499–508.
- [14] H. Molina-Abril, P. Real, Homological optimality in discrete morse theory through chain homotopies, *Pattern Recognit. Lett.* 33 (2012) 1501–1506.
- [15] T. Peterka, R. Ross, A. Gyulassy, V. Pascucci, W. Kendall, H.W. Shen, T.Y. Lee, A. Chaudhuri, Scalable parallel building blocks for custom data analysis, *Proceedings of the IEEE Symposium on Large Data Analysis and Visualization (LDAV)*, IEEE, 2011, pp. 105–112.
- [16] P. Pilarczyk, P. Real, Computation of cubical homology, cohomology, and (co)homological operations via chain contraction, *Adv. Comput. Math.* 41 (2015) 253–275.
- [17] R. Reina-Molina, in: Python and Nvidia CUDA code for effective homology of cubical complexes, 2016. https://munkres.us.es/wiki/pages/a5V9r2M9H/SOURCE_CODE.html
- [18] R. Reina-Molina, D. Díaz-Pernil, P. Real, A. Berciano, Membrane parallelism for discrete morse theory applied to digital images, *Appl. Algebra Eng. Commun. Comput.* 26 (2015) 49–71.
- [19] J. Reininghaus, I. Hotz, Combinatorial 2d vector field topology extraction and simplification, *Topological Methods in Data Analysis and Visualization*, Springer, 2011, pp. 103–114.
- [20] J. Reininghaus, C. Löwen, I. Hotz, Fast combinatorial vector field topology, *Visual. Comput. Graph. IEEE Trans.* 17 (2011) 1433–1443.
- [21] F. Sergeraert, The computability problem in algebraic topology, *Adv. Math.* 104 (1994) 1–29.
- [22] N. Shivashankar, V. Natarajan, et al., Parallel computation of 2d morse-smale complexes, *Vis. Comput. Graph. IEEE Trans.* 18 (2012) 1757–1770.
- [23] F.W. Wehrli, B.R. Gomberg, P.K. Saha, H.K. Song, S.N. Hwang, P.J. Snyder, Digital topological analysis of in vivo magnetic resonance microimages of trabecular bone reveals structural implications of osteoporosis, *J. bone Mineral Res.* 16 (2001) 1520–1531.