A global method for reducing multidimensional size graphs*

Andrea Cerri[1], Patrizio Frosini[2], Walter G. Kropatsch[1], and Claudia Landi[3]

[1] Vienna University of Technology, Faculty of Informatics,
Pattern Recognition and Image Processing Group, Austria
{acerri,krw}@prip.tuwien.ac.at
[2] Università di Bologna,
ARCES and Dipartimento di Matematica, Italy
frosini@dm.unibo.it
[3] Università di Modena e Reggio Emilia,
Dipartimento di Scienze e Metodi dell'Ingegneria, Italy
clandi@unimore.it

Abstract. This paper introduces the concept of discrete multidimensional size function, a mathematical tool studying the so-called size graphs. These graphs constitute an ingredient of Size Theory, a geometrical/topological approach to shape analysis and comparison. A global method for reducing size graphs is presented, together with a theorem stating that size graphs reduced in such a way preserve all the information in terms of multidimensional size functions. This approach can lead to simplify the effective computation of discrete multidimensional size functions.

1 Introduction

In the last twenty years, Size Theory has revealed to be a suitable geometrical/topological approach to shape analysis and comparison, which are probably two of the main issues in the fields of Computer Vision, Computer Graphics and Pattern Recognition. In this context, the main tool proposed by Size Theory is the concept of *size function*, a shape descriptor able to capture the qualitative aspects of a shape, and describing them quantitatively. More precisely, the basic idea is to model a shape by means of a topological space $\mathcal{M}$ and a continuous function $\varphi : \mathcal{M} \to \mathbb{R}$, called *measuring function*. The role of the measuring functions is to describe those properties that are considered relevant for the shape comparison or the shape analysis problem at hand. In this setting, the size function $\ell_{(\mathcal{M},\varphi)}$ encodes part of the topological changes occurring in the lower level sets induced on $\mathcal{M}$ by φ. In this way, the starting problem of comparing shapes modeled by pairs $(topological\ space, measuring\ function)$ can then be recast into the one of comparing the associated size functions. For details and more references about Size Theory the reader is referred to [3].

More recently, similar ideas have been re-proposed by Persistent Homology from the homological point of view [14, 15]. More precisely, the notion of size function coincide with the one of 0th persistent homology group.

Since their introduction, size functions have been extensively used especially in the fields of Computer Vision [8, 13, 25], where the objects under study are images, and Computer Graphics, comparing, e.g, 3D-models [4]. The success of size functions in such applicative contexts is due to the fact that they admit a

* Research supported by the Austrian Science Fund (FWF) grant no. P20134-N13.

very simple and compact representation [17], and they are stable with respect to a suitable distance [10, 11]. Moreover, size functions show resistance to noise and modularity [16]: In particular, they inherit their invariance properties directly from the considered measuring functions. For example, in [8] an effective system for content-based retrieval of figurative images, based on size functions, is presented. Three different classes of image descriptors are introduced and integrated, for a total amount of 25 measuring functions. The evaluation of this fully automatic retrieval system has been performed on a benchmark database of more than 10,000 real trademark images, supplied by the United Kingdom Patent Office. Comparative results have been performed, showing that the proposed method actually outperforms other existing whole-image matching techniques.

As the previous considerations suggest, a common scenario in applications is when two or more properties characterize the objects under study. Moreover, sometimes it could be desirable to consider properties of shapes that are intrinsically multidimensional, such as the coordinates of a point into the 3-dimensional space or the representation of color in the RGB model. These motivations recently drove the attention to extending Size Theory to a multidimensional context [2, 6] (see [7] for the multidimensional version of Persistent Homology). Here the term multidimensional (or, equivalently, k-dimensional) refers to the fact that the measuring functions take values in $\mathbb{R}^k$ and has no reference with the dimensionality of the objects under study. Therefore, such an enlarged setting leads to model a shape as a pair $(\mathcal{M}, \boldsymbol{\varphi})$, with $\boldsymbol{\varphi} : \mathcal{M} \to \mathbb{R}^k$, and consequently to consider the so called *multidimensional size functions*.

Even in this multidimensional setting Size Theory gave encouraging results when applied to shape analysis and comparison problems, see, e.g., [2, 5]. In those papers it has been shown that, besides enabling the study of multidimensional properties of the objects under study, the advantage of working with multidimensional measuring functions is that shapes can be simultaneously investigated by k different 1-dimensional measuring functions. In other words, k different functions cooperate to produce a single shape descriptor. The higher discriminatory power of multidimensional size functions in comparison to 1-dimensional ones has been formally proved in [2].

Motivations and contributions of the paper. Obviously, dealing with applications involves the development of a discrete counterpart of the theory. In this perspective, a shape can be discretized by a graph $G = (V(G), E(G))$ endowed with a function $\boldsymbol{\varphi} : V(G) \to \mathbb{R}^k$, being $V(G)$ the set of vertices of G. This leads to consider pairs of the type $(G, \boldsymbol{\varphi})$, called *size graphs*. In this mathematical setting, *discrete k-dimensional size functions* count the number of connected components in $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$ containing at least one vertex of $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{x} \rangle$ where, for $\boldsymbol{t} \in \mathbb{R}^k$, $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{t} \rangle$ is defined as the subgraph of G obtained by erasing all vertices of G at which φ_i takes a value strictly greater than t_i, for at least one index $i \in \{1, \ldots, k\}$, and all the edges connecting those vertices to others.

Therefore, in computing discrete k-dimensional size functions, we have to count the connected components of particular subgraphs of a size graph. It is reasonable to argue that, the greater the dimension k, the higher the discriminatory power of k-dimensional size functions. On the other hand, the smaller the graph, the faster the computation. Moreover, in applications we often have to deal with big graphs, implying high computational costs. According to these considerations, it follows that the problem of reducing a size graph without changing the associated discrete k-dimensional size function is a desirable target.

In previous works ([12, 18]), it has been proved that, in the case $k = 1$, a size graph can be reduced by means of a global method (its application requires the knowledge of all the size graph) and a local method (it requires only a local knowledge of a size graph), obtaining a very simple structure.

In this paper, we present a first attempt for a reduction procedure for size graphs in the case $k > 1$. More precisely, we provide a global reduction method for size graphs, together with a theorem stating that reduced size graphs preserve all the information in terms of k-dimensional size functions.

Related works. The ideas underlying the concept of size functions are partly shared in introducing the so-called *maximally stable extremal regions (MSERs)* [21]. MSERs are image elements useful in wide-baseline matching. Given a gray-level image I, the basic intuition is to study the evolution of the thresholded image I_t, varying the parameter t. The formal definition of MSERs is then derived by considering the set of all connected components of all thresholded images I_t. These image elements are characterized by a number of nice properties, such as the invariance to affine transformation of image intensities and stability.

Besides being related to [12] and [18], the present work fits in the current research and interest in strategies for reducing data structures preserving some topological/homological information, motivated by Pattern Recognition and data analysis problems. For example, in [24] the authors propose a method for computing homology groups and their generators of a 2D pixel image, by using a hierarchical structure called irregular graph pyramid. Their method is based on two operations, preserving the homology information contained in each region of an image, but progressively simplifying the starting graph representing the image, and constituting the base level of the pyramid. The desired homological information is then computed at the top of the pyramid. This approach finds its roots in a more general framework firstly introduced in [20]. Motivated by problems coming from discrete dynamics, in [19] the authors propose an algorithm for computing homology of a finitely generated chain complex. Such an algorithm is based on reducing the size of the complex preserving homology information in each step of the reduction. Computing the homology of the chain complex is then postponed until the complex is acceptably small. The same philosophy leads the authors of [23] to provide a reduction algorithm for simplifying the computation of homology information for cubical sets and polytopes.

The remainder of the paper is organized as follows. In Section 2 we introduce the standard facts and some basic definitions about discrete multidimensional size functions. Section 3 is devoted to present our main result, together with some experiments. Some discussions in Section 4 conclude the paper.

2 Basic definitions

In this section we provide some basic definitions about discrete k-dimensional size functions. The following relations are introduced in $\mathbb{R}^k$: for every $\boldsymbol{x} = (x_1, \ldots, x_k)$ and $\boldsymbol{y} = (y_1, \ldots, y_k)$, we shall say $\boldsymbol{x} \preceq \boldsymbol{y}$ (resp. $\boldsymbol{x} \succ \boldsymbol{y}$, $\boldsymbol{x} \succeq \boldsymbol{y}$) if and only if $x_i \leq y_i$ (resp. $<$, $x_i \geq y_i$) for every index $i = 1, \ldots, k$. Moreover, we shall write $\boldsymbol{x} \not\preceq \boldsymbol{y}$ (resp. $\boldsymbol{x} \prec \boldsymbol{y}$) when the relation between $\boldsymbol{x}$ and $\boldsymbol{y}$ expressed by the operator $\preceq$ (resp. $\succeq$) is not verified. Finally, we recall that Δ^+ is defined as the open set $\{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^k \times \mathbb{R}^k : \boldsymbol{x} \prec \boldsymbol{y}\}$.

Definition 1 (Size Graph). *Let $G = (V(G), E(G))$ be a finite, ordered simple graph with $V(G)$ set of vertices and $E(G)$ set of edges. Assume that a function $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_k) : V(G) \to \mathbb{R}^k$ is given. The pair $(G, \boldsymbol{\varphi})$ is called a* size graph.

Definition 2. *For every* $\boldsymbol{y} = (y_1, \ldots, y_k) \in \mathbb{R}^k$*, we denote by* $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$ *the subgraph of* G *obtained by erasing all vertices* $v \in V(G)$ *such that* $\boldsymbol{\varphi}(v) \not\preceq \boldsymbol{y}$*, and all the edges connecting those vertices to others. If* $v_a, v_b \in V(G)$ *belong to the same connected component of* $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$*, we shall write* $v_a \cong_{G\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle} v_b$*.*

We are now ready to introduce discrete k-dimensional size functions.

Definition 3 (discrete k-dimensional size function). *We shall call* discrete k-dimensional size function *of the size graph* $(G, \boldsymbol{\varphi})$ *the function* $\ell_{(G,\boldsymbol{\varphi})} : \Delta^+ \to \mathbb{N}$ *defined by setting* $\ell_{(G,\boldsymbol{\varphi})}(\boldsymbol{x}, \boldsymbol{y})$ *equal to the number of connected components in* $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$ *containing at least one vertex of* $G\langle \boldsymbol{\varphi} \preceq \boldsymbol{x} \rangle$*.*

Example 1. Figure 1 shows an example of size graph, together with the related discrete 1-dimensional size function. We remark that in the case $k = 1$ the symbols $\boldsymbol{\varphi}, \boldsymbol{x}, \boldsymbol{y}$ are replaced by φ, x, y respectively. As can be seen, in the 1-dimensional case the domain Δ^+ of $\ell_{(G,\varphi)}$ is a subset of the real plane. In each region of Δ^+, the value of $\ell_{(G,\varphi)}$ in that region is displayed.

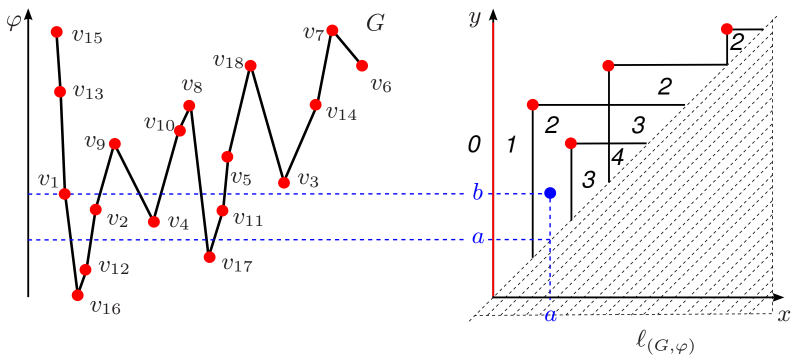

Fig. 1. A size graph and the associated discrete size function.

For example, to compute the value of $\ell_{(G,\varphi)}$ at the point (a, b), it is sufficient to count how many of the three connected components in the subgraph $G\langle \varphi \le b \rangle$ contain at least one vertex of $G\langle \varphi \le a \rangle$: It can be easily checked that $\ell_{(G,\varphi)}(a, b) = 2$.

In what follows, we will assume that the set of vertices $V(G)$ of the graph G is a subset of a Euclidean space.

3 A global method for reducing $(G, \boldsymbol{\varphi})$: the $\mathcal{L}$-reduction

As stressed before, our goal is to reduce a size graph $(G, \boldsymbol{\varphi})$ without changing the related discrete k-dimensional size function: This can be done by erasing all those vertices of G that do not contain, in terms of discrete k-dimensional size functions, "meaningful information". Indeed, in order to compute the discrete k-dimensional size function of $(G, \boldsymbol{\varphi})$, we are only interested in capturing the "birth" of new connected components and the "death", i.e. the merging, of the existing ones: As will be shown, these events are strongly related to particular vertices of G, that can be seen, in some sense, as "critical points" of the function $\boldsymbol{\varphi}$ with respect to the relation $\preceq$. The proposed reduction method allows us to detect these particular vertices and to introduce the concept of *$\mathcal{L}$-reduction of*

$(G,\boldsymbol{\varphi})$, a new size graph $(G_{\mathcal{L}},\boldsymbol{\varphi}_{\mathcal{L}})$ that is obtained by considering *only* such vertices instead of the entire set $V(G)$. The importance of the $\mathcal{L}$-reduction is shown by our main result, stated in Theorem 1, which will be formally proved at the end of this section and can be rephrased as follows:

Theorem 1 (rephrased). The k-dimensional size functions of $(G_{\mathcal{L}},\boldsymbol{\varphi}_{\mathcal{L}})$ and $(G,\boldsymbol{\varphi})$ coincide.

From now on, we assume that a size graph $(G,\boldsymbol{\varphi})$ is given. Moreover, for every $v_i \in V(G)$ we define A_i as the set of the "lower adjacent vertices" for v_i, i.e. $A_i = \{v_j : (v_i, v_j) \in E(G), \boldsymbol{\varphi}(v_j) \preceq \boldsymbol{\varphi}(v_i)\} \cup \{v_i\}$.

Definition 4 (Single step descent flow operator). *Let $L : V(G) \to V(G)$ be the function defined as follows: For every $v_i \in V(G)$ let $B_i \subseteq A_i$ be the set whose elements are the vertices $w \in A_i$ for which the Euclidean norm $\|\boldsymbol{\varphi}(w) - \boldsymbol{\varphi}(v_i)\|$ takes the largest value. Finally, we choose the vertex $v_h \in B_i$ for which the index h is minimum. Then, we set $L(v_i) = v_h$. We shall call L the* single step descent flow operator.

From the definition of L and the finiteness of $V(G)$, if follows that for every $v \in V(G)$ there must exist a minimum index $m(v) \leq |V(G)|$ such that $L^{m(v)}(v) = L^{m(v)+1}(v)$ (if $L(v) = v$ we will set $m(v) = 0$).

Definition 5 (Descent flow operator). *For every $v \in V(G)$ we set $\mathcal{L}(v) = L^{m(v)}(v)$. We shall call the function $\mathcal{L} : V(G) \to V(G)$ the* descent flow operator.

In other words, the descent flow operator takes each vertex $v_i \in V(G)$ to a sort of "local minimum" $v_j = \mathcal{L}(v_i)$ of the function $\boldsymbol{\varphi}$, with respect to the relation $\preceq$. This implies that, starting from v_j we are not able to reach a vertex w adjacent to it with $\boldsymbol{\varphi}(w) \preceq \boldsymbol{\varphi}(v_j)$, strictly decreasing the value of at least one component of $\boldsymbol{\varphi}$. During the descent, indexes are used to univocally decide the path in case the set B_i contains more than one vertex.

Example 2. Figure 2 shows some possible cases arising from the action of the operators L and $\mathcal{L}$ when $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$. As can be seen, the vertex v_1 is taken by the operator L to $v_4 = L^5(v_1)$. Since it is not possible to reach another vertex from v_4 decreasing the values of both φ_1 and φ_2, we shall set $v_4 = \mathcal{L}(v_1)$. Analogously, we have $v_5 = \mathcal{L}(v_2)$. The last considered case is represented by the vertex v_3: it can be seen as a fixed point with respect to the operator L, i.e. it holds that $L(v_3) = v_3$, so we shall set $\mathcal{L}(v_3) = v_3$.

Definition 6 (Minimum vertex). *Each vertex v for which $\mathcal{L}(v) = v$ will be called a* minimum vertex *of $(G,\boldsymbol{\varphi})$. Call M the set of minimum vertices of $(G,\boldsymbol{\varphi})$.*

We point out that M is the set of all those vertices representing the "birth" of new connected components in $(G,\boldsymbol{\varphi})$: Indeed, by increasing the values of $\varphi_1, \ldots, \varphi_k$, such an event occurs only when the values labeling a minimum vertex are reached.

The following two definitions characterize the "death-points" of existing connected components of $(G,\boldsymbol{\varphi})$.

Definition 7 (Ridge pair). *Let $v_{j_1}, v_{j_2} \in V(G)$ be two distinct minimum vertices of $(G,\boldsymbol{\varphi})$. Suppose v_{i_1}, v_{i_2} are two adjacent vertices of G, such that $\{\mathcal{L}(v_{i_1}), \mathcal{L}(v_{i_2})\} = \{v_{j_1}, v_{j_2}\}$; we shall call $\{v_{i_1}, v_{i_2}\}$ a* ridge pair adjacent to the minimum vertices *v_{j_1} and v_{j_2}.*

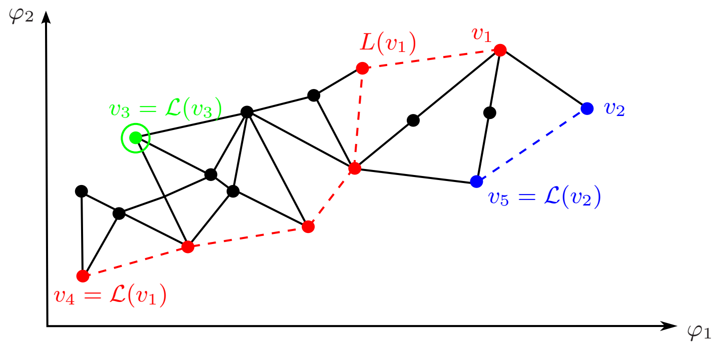

Fig. 2. The operators L and $\mathcal{L}$ in action: Some examples when $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$.

Definition 8 (Main saddle). *Let $v_{j_1}, v_{j_2} \in V(G)$ be two distinct minimum vertices of $(G, \boldsymbol{\varphi})$. Let also $\{v_{i_1}, v_{i_2}\}$ be a ridge pair adjacent to the minimum vertices v_{j_1}, v_{j_2} such that the following statements hold:*

1. *there does not exist another ridge pair $\{v_{i_3}, v_{i_4}\}$ adjacent to the minimum vertices v_{j_1}, v_{j_2} with*

$$(1) \begin{cases} \max\{\varphi_h(v_{i_3}), \varphi_h(v_{i_4})\} \leq \max\{\varphi_h(v_{i_1}), \varphi_h(v_{i_2})\}, \ h = 1, \ldots, k; \\ \exists \bar{h} : \max\{\varphi_{\bar{h}}(v_{i_3}), \varphi_{\bar{h}}(v_{i_4})\} < \max\{\varphi_{\bar{h}}(v_{i_1}), \varphi_{\bar{h}}(v_{i_2})\}, \ \bar{h} \in \{1, \ldots, k\}; \end{cases}$$

2. *if $\{v_{i_3}, v_{i_4}\}$ is another ridge pair adjacent to the minimum vertices v_{j_1}, v_{j_2} with*

$$(2) \qquad \max\{\varphi_h(v_{i_3}), \varphi_h(v_{i_4})\} = \max\{\varphi_h(v_{i_1}), \varphi_h(v_{i_2})\}, \ h = 1, \ldots, k,$$

then (i_1, i_2) precedes (i_3, i_4) in the lexicographic order. We shall call the set $\{v_{i_1}, v_{i_2}\}$ the main saddle adjacent to the minimum vertices v_{j_1}, v_{j_2} *and S the set of main saddles of $(G, \boldsymbol{\varphi})$.*

Roughly speaking, the set of ridge pairs of $(G, \boldsymbol{\varphi})$ can be partially ordered by means of the relation $\preceq$. In this sense, the main saddles will be the lowest ridge pairs.

Example 3. Figure 3(a), 3(b), 3(c) shows some examples of ridge pairs and main saddles, when function $\boldsymbol{\varphi}$ takes values in $\mathbb{R}^2$.

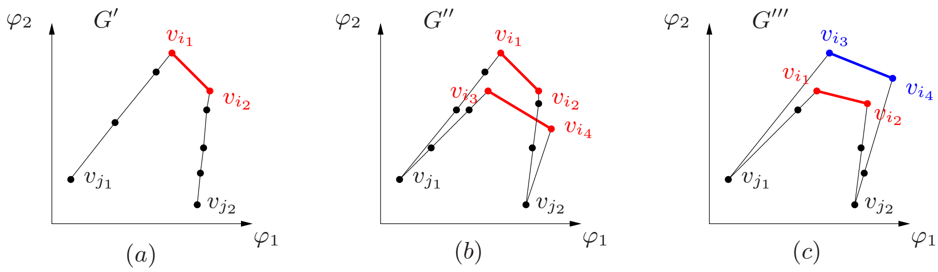

Fig. 3. Some examples of ridge pairs and main saddles.

To clarify the role of main saddles, let us study the changing in the number of connected components of the subgraphs $G'\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$, $G''\langle \boldsymbol{\varphi} \preceq \boldsymbol{y} \rangle$ and $G'''\langle \boldsymbol{\varphi} \preceq$

$\boldsymbol{y}\rangle$, with $\boldsymbol{y} \in \mathbb{R}^2$, just for $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{j_1}), \varphi_1(v_{j_2})\}, \max\{\varphi_2(v_{j_1}), \varphi_2(v_{j_2})\})$: Indeed, we want to capture the merging of the connected components arising from the minimum vertices v_{j_1} and v_{j_2} in the three instances. According to this consideration, by means of the chosen assumption on $\boldsymbol{y}$ we ensure that both v_{j_1} and v_{j_2} belong to the subgraphs $G'\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$, $G''\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$ and $G'''\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$.

In Figure 3(a) a main saddle adjacent to the minimum vertices v_{j_1} and v_{j_2} is displayed. In this setting, by varying the values taken by $\boldsymbol{y}$ under the assumption $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{j_1}), \varphi_1(v_{j_2})\}, \max\{\varphi_2(v_{j_1}), \varphi_2(v_{j_2})\})$, it holds that for $\boldsymbol{y} \not\succeq (\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\}, \max\{\varphi_2(v_{i_1}), \varphi_2(v_{i_2})\})$ the subgraph $G'\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$ consists of the two connected components arising from v_{j_1} and v_{j_2}, reducing to a unique one when $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\}, \max\{\varphi_2(v_{i_1}), \varphi_2(v_{i_2})\})$.

Figure 3(b) shows an example of two ridge pairs that can be considered "uncomparable", due to the fact $\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\} < \max\{\varphi_1(v_{i_3}), \varphi_1(v_{i_4})\}$, while $\max\{\varphi_2(v_{i_1}), \varphi_2(v_{i_2})\} > \max\{\varphi_2(v_{i_3}), \varphi_2(v_{i_4})\}$. Thus, both $\{v_{i_1}, v_{i_2}\}$ and $\{v_{i_3}, v_{i_4}\}$ will be main saddles. In this case, when $\boldsymbol{y}$ varies under the assumption $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{j_1}), \varphi_1(v_{j_2})\}, \max\{\varphi_2(v_{j_1}), \varphi_2(v_{j_2})\})$, the number of the connected components in the subgraph $G''\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$ decreases (from 2 to 1) when the relation $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\}, \max\{\varphi_2(v_{i_1}), \varphi_2(v_{i_2})\})$ (or, alternatively, the relation $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{i_3}), \varphi_1(v_{i_4})\}, \max\{\varphi_2(v_{i_3}), \varphi_2(v_{i_4})\})$) becomes true.

Finally, Figure 3(c) shows two comparable ridge pairs, hence the "lower" one, that is $\{v_{i_1}, v_{i_2}\}$, will be a main saddle, while the other will be not. Consider $G'''\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$, assuming that $\boldsymbol{y}$ varies according to the restriction $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{j_1}), \varphi_1(v_{j_2})\}, \max\{\varphi_2(v_{j_1}), \varphi_2(v_{j_2})\})$: It consists of two connected components arising from v_{j_1} and v_{j_2}, merging into a unique one as soon as the relation $\boldsymbol{y} \succeq (\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\}, \max\{\varphi_2(v_{i_1}), \varphi_2(v_{i_2})\})$ becomes true.

As Example 3 suggests, S is the set of all those couples of vertices representing the "death", i.e. the merging, of existing connected components in the given size graph $(G, \boldsymbol{\varphi})$.

We are now ready to introduce the concept of $\mathcal{L}$-reduced size graph:

Definition 9 ($\mathcal{L}$-reduced size graph). *Let $G_\mathcal{L} = (V(G_\mathcal{L}), E(G_\mathcal{L}))$ be the graph with $V(G_\mathcal{L}) = M \cup S$ and $E(G_\mathcal{L})$ defined as follows: $(u, v) \in E(G_\mathcal{L})$ (and hence u and v are adjacent) if and only if either u or v is a minimum vertex and the other is a main saddle adjacent to it (in the sense of Definition 8). Let also $\boldsymbol{\varphi}_\mathcal{L} : V(G_\mathcal{L}) \to \mathbb{R}^k$ be a function defined in this way: $\boldsymbol{\varphi}_\mathcal{L}(v) = \boldsymbol{\varphi}(v)$ if $v \in M$ and $\boldsymbol{\varphi}_\mathcal{L}(u) = (\max\{\varphi_1(v_{i_1}), \varphi_1(v_{i_2})\}, \ldots, \max\{\varphi_k(v_{i_1}), \varphi_k(v_{i_2})\})$ if $u = \{v_{i_1}, v_{i_2}\} \in S$. The size graph $(G_\mathcal{L}, \boldsymbol{\varphi}_\mathcal{L})$ will be called the* $\mathcal{L}$-reduction of *$(G, \boldsymbol{\varphi})$.*

Remark 1. We stress that each main saddle $\{v, w\}$ of a size graph $(G, \boldsymbol{\varphi})$ will be represented, in the $\mathcal{L}$-reduced size graph, by a *unique* vertex labeled by the k-tuple $(\max\{\varphi_1(v), \varphi_1(w)\}, \ldots, \max\{\varphi_k(v), \varphi_k(w)\})$.

Remark 2. The global reduction method we have just defined is strongly related to the concept of Pareto-Optimality, a well-known topic in Economy, especially in the field of Multi-Objective Optimization. For a detailed treatment about Pareto-Optimality, the reader is referred to [22]. Another related notion is the one of *pseudocritical point* studied in [9]

The importance of the $\mathcal{L}$-reduction is shown by our main result, stating that discrete k-dimensional size functions are invariant with respect to this global reduction method.

Theorem 1. *For every* $(\boldsymbol{x},\boldsymbol{y}) \in \Delta^+$*, it holds that* $\ell_{(G,\boldsymbol{\varphi})}(\boldsymbol{x},\boldsymbol{y}) = \ell_{(G_{\mathcal{L}},\boldsymbol{\varphi}_{\mathcal{L}})}(\boldsymbol{x},\boldsymbol{y})$.

In order to prove Theorem 1, we need the following lemma.

Lemma 1. *Let* v_1*,* v_2 *be two minimum vertices of* $(G,\boldsymbol{\varphi})$*. Then, for every* $\boldsymbol{y} \in \mathbb{R}^k$*, it holds that* $v_1 \cong_{G\langle\boldsymbol{\varphi}\preceq\boldsymbol{y}\rangle} v_2$ *if and only if* $v_1 \cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle} v_2$.

Proof. Suppose that $v_1 \cong_{G\langle\boldsymbol{\varphi}\preceq\boldsymbol{y}\rangle} v_2$. Then, by definition there exists a sequence $(v_1 = v_{j_1}, v_{j_2}, \ldots, v_{j_{m-1}}, v_{j_m} = v_2)$ such that $(v_{j_n}, v_{j_{n+1}}) \in E(G)$ for every $n = 1,\ldots,m-1$, and $v_{j_n} \in G\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$ for every $n = 1,\ldots,m$. Consider the sequence $(\mathcal{L}(v_1) = v_1, \mathcal{L}(v_{j_2}), \ldots, \mathcal{L}(v_{j_{m-1}}), \mathcal{L}(v_2) = v_2)$ of minimum vertices. Substituting each subsequence of equal consecutive vertices by a unique vertex representing such a subsequence, we obtain a new sequence $(v_1 = w_1, w_2, \ldots, w_{s-1}, w_s = v_2)$ (in other words, the sequence $(u_1, u_1, \ldots, u_1, u_2, u_2 \ldots, u_2, \ldots, u_n, u_n, \ldots, u_n)$ is substitute with $(u_1, u_2, \ldots, u_n)$). It is easy to prove that, for every index $j < s$, there exists at least one main saddle σ_j adjacent to w_j and w_{j+1}, such that $\sigma_j \in G\langle\boldsymbol{\varphi} \preceq \boldsymbol{y}\rangle$. Then, consider the sequence $(w_1, \sigma_1, v_2, \sigma_2, \ldots, w_{s-1}, \sigma_{s-1}, w_s)$: such a sequence proves that $v_1 \cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle} v_2$.

Conversely, suppose that $v_1 \cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle} v_2$. By definition there exists a sequence $(v_1 = w_1, \sigma_1, w_2, \sigma_2, \ldots, w_{s-1}, \sigma_{s-1}, w_s = v_2)$ of vertices of $G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}} \preceq \boldsymbol{y}\rangle$ such that every vertex w_j is a minimum vertex and every σ_j is a main saddle adjacent to w_j and w_{j+1}. Therefore, we can modify such a sequence in order to obtain the following one: for every index $j < s$, between w_j and $\sigma_j = \{v_{i_j}, v_{n_j}\}$ insert the sequence $(L^{m(v_{i_j})-1}(v_{i_j}), L^{m(v_{i_j})-2}(v_{i_j}), \ldots, L^2(v_{i_j}), L(v_{i_j}))$, while between σ_j e w_{j+1} insert the sequence $(L(v_{n_j}), L^2(v_{n_j}), \ldots, L^{m(v_{n_j})-2}(v_{n_j}), L^{m(v_{n_j})-1}(v_{n_j}))$ (we are assuming $w_j = \mathcal{L}(v_{i_j})$ and $w_{j+1} = \mathcal{L}(v_{n_j})$). Finally, by substituting the vertices v_{i_j} e v_{n_j} (taken in this order) for every main saddle σ_j, we obtain a new sequence proving that $v_1 \cong_{G\langle\boldsymbol{\varphi}\preceq\boldsymbol{y}\rangle} v_2$.

Now we are ready to prove Theorem 1.

Proof. Let $(\boldsymbol{x},\boldsymbol{y}) \in \Delta^+$. We have to prove that there exists a bijection $F : G\langle\boldsymbol{\varphi} \preceq \boldsymbol{x}\rangle/\cong_{G\langle\boldsymbol{\varphi}\preceq\boldsymbol{y}\rangle} \to G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}} \preceq \boldsymbol{x}\rangle/\cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle}$. For every equivalence class $C \in G\langle\boldsymbol{\varphi} \preceq \boldsymbol{x}\rangle/\cong_{G\langle\boldsymbol{\varphi}\preceq\boldsymbol{y}\rangle}$ we choose a minimum vertex $v_C \in C$. Obviously, v_C is also a vertex of $G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}} \preceq \boldsymbol{x}\rangle$. Therefore in $G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}} \preceq \boldsymbol{x}\rangle/\cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle}$ there exists an equivalence class D containing v_C. We shall set $F(C) = D$. From Lemma 1 it follows that F is equivalence class in $G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}} \preceq \boldsymbol{x}\rangle/\cong_{G_{\mathcal{L}}\langle\boldsymbol{\varphi}_{\mathcal{L}}\preceq\boldsymbol{y}\rangle}$ contains at least one minimum vertex of $G\langle\boldsymbol{\varphi} \preceq \boldsymbol{x}\rangle$.

Remark 3. The $\mathcal{L}$-reduction of a size graph $(G,\boldsymbol{\varphi})$ is not unique. In particular, changing the ordering of the set $V(G)$ can produce different, non-isomorphic $\mathcal{L}$-reduced size graphs. On the other hand, Theorem 1 shows that we will always obtain $\mathcal{L}$-reductions of $(G,\boldsymbol{\varphi})$ with the same associated discrete k-dimensional size function.

Therefore, Theorem 1 allows us to evaluate the discrete k-dimensional size function of a size graph $(G,\boldsymbol{\varphi})$ directly dealing with one of its $\mathcal{L}$-reductions.

3.1 Experimental results

Table 1 shows how our global reduction method can facilitate the computation of $\ell_{(G,\boldsymbol{\varphi})}$, simplifying the structure of (G,φ) but preserving the same information in terms of discrete k-dimensional size functions. We considered four graphs obtained from as many triangle meshes (available at [1]) by taking the 0-dimensional simplexes as vertices and the 1-dimensional simplexes as edges. For each graph, we considered the 2-dimensional measuring function $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ taking each vertex v of coordinates (x,y,z) to the pair $\boldsymbol{\varphi}(v) = (|x|,|y|)$.

	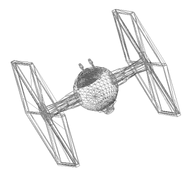	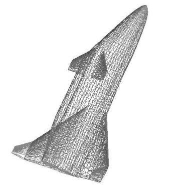	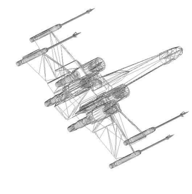	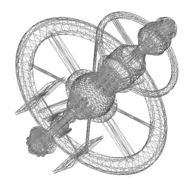
	tie	space_shuttle	x_wing	space_station
$\|V(G)\|$	2014	2376	3099	5749
$\|E(G)\|$	5944	6330	9190	15949
$\|V(G_{\mathcal{L}})\|$	588	262	571	1935
$\|E(G_{\mathcal{L}})\|$	826	328	838	2778
$V\% - E\%$	29.2% - 13.9%	11% - 5.2%	18.4% - 9.2%	33.66% - 17.42%

Table 1. Some experimental results.

Table 1, from row 1 to 4, shows respectively the number of vertices $|V(G)|$ and edges $|E(G)|$ for each considered size graph $(G, \boldsymbol{\varphi})$ and for the associated $\mathcal{L}$-reduction $(G_{\mathcal{L}}, \boldsymbol{\varphi}_{\mathcal{L}})$ (i.e. $|V(G_{\mathcal{L}})|$ and $|E(G_{\mathcal{L}})|$). In particular, if M and S are respectively the set of minimum vertices and of main saddles for a size graph $(G, \boldsymbol{\varphi})$, it easily follows from Definition 9 that $|V(G_{\mathcal{L}})| = |M \cup S|$ and $|E(G_{\mathcal{L}})| = 2|S|$. For each considered $(G, \boldsymbol{\varphi})$, the last row of Table 1 shows respectively the ratios $V\% = |V(G_{\mathcal{L}})|/|V(G)|$ and $E\% = |E(G_{\mathcal{L}})|/|E(G)|$, expressing them in percentage points. In other words, the lower those ratios, the higher the reduction rate. As can be seen, these experiments gave encouraging results, enabling the reduction of a size graph up to the 11% of the starting number of vertices and the 5.2% of the starting number of edges (space_shuttle case).

To conclude, we report the most salient part, at least in our perspective, of the algorithm we implemented to obtain our experimental results, i.e. the computation of the descent flow operator introduced in Definition 5. All the rest can be easily derived from the theoretical setting discussed in the previous sections, combined with what follows. The symbol $SSDFO(v)$ denotes the single step descent flow operator computed at a vertex v.

Algorithm 1 Computation of the descent flow operator $\mathcal{L}(v)$ for $v \in V(G)$

$L(v) \leftarrow SSDFO(v)$
while $L(v) \neq v$ **do**
  $v \leftarrow L(v)$;
  $L(v) \leftarrow SSDFO(v)$;
end while
$\mathcal{L}(v) \leftarrow L(v)$

4 Conclusion

In this paper we presented a global method for reducing size graphs together with a theorem, stating that discrete multidimensional size functions are invariant with respect to this reduction method. This result can lead us to easily and fast compute discrete multidimensional size functions for applications, as highlighted by some experiments showing the feasibility of the proposed reduction scheme. This work can be seen as a contribution in finding reduction methods for data structure encoding multidimensional information of shapes, in a way that the topological/homological information carried with them is preserved. For the next future, it could be interesting to study the existence of a local reduction method for k-dimensional size graphs preserving the information in terms of multidimensional size functions.

References

1. http://gts.sourceforge.net/samples.html
2. Biasotti, S., Cerri, A., Frosini, P., Giorgi, D., Landi, C.: Multidimensional size functions for shape comparison. J. Math. Imaging Vision 32(2), 161–179 (2008)
3. Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., Papaleo, L., Spagnuolo, M.: Describing shapes by geometrical-topological properties of real functions. ACM Comput. Surv. 40(4), 1–87 (2008)
4. Biasotti, S., Giorgi, D., Spagnuolo, M., Falcidieno, B.: Size functions for comparing 3d models. Pattern Recogn. 41(9), 2855–2873 (2008)
5. Biasotti, S., Cerri, A., Giorgi, D.: k-dimensional size functions for shape description and comparison. In: Cucchiara, R. (ed.) ICIAP. pp. 795–800. IEEE Computer Society (2007)
6. Cagliari, F., Di Fabio, B., Ferri, M.: One-dimensional reduction of multidimensional persistent homology. Proc. Amer. Math. Soc. 138(8), 3003–3017 (2010)
7. Carlsson, G., Zomorodian, A.: The theory of multidimensional persistence. Discrete & Computational Geometry 42(1), 71–93 (2009)
8. Cerri, A., Ferri, M., Giorgi, D.: Retrieval of trademark images by means of size functions. Graph. Models 68(5), 451–471 (2006)
9. Cerri, A., Frosini, P.: Necessary conditions for discontinuities of multidimensional size functions. Tech. Rep. no. 2565, Università di Bologna (2008), Available at http://amsacta.cib.unibo.it/2565/
10. d'Amico, M., Frosini, P., Landi, C.: Using matching distance in size theory: A survey. Int. J. Imag. Syst. Tech. 16(5), 154–161 (2006)
11. d'Amico, M., Frosini, P., Landi, C.: Natural pseudo-distance and optimal matching between reduced size functions. Acta. Appl. Math. 109, 527–554 (2010)
12. d'Amico, M.: A New Optimal Algorithm for Computing Size Functions of Shapes. In: CVPRIP Algorithms III. pp. 107–110. Atlantic City (2000)
13. Dibos, F., Frosini, P., Pasquignon, D.: The use of size functions for comparison of shapes through differential invariants. J. Math. Imaging Vis. 21(2), 107–118 (2004)
14. Edelsbrunner, H., Harer, J.: Computational Topology: An Introduction. American Mathematical Society (2009)
15. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. Discrete & Computational Geometry 28(4), 511–533 (2002)
16. Frosini, P., Landi, C.: Size theory as a topological tool for computer vision. Pattern Recognition and Image Analysis 9, 596–603 (1999)
17. Frosini, P., Landi, C.: Size functions and formal series. Appl. Algebra Engrg. Comm. Comput. 12(4), 327–349 (2001)
18. Frosini, P., Pittore, M.: New methods for reducing size graphs. Intern. J. Computer Math (70), 505–517 (1999)
19. Kaczynski, T., Mrozek, M., Slusarek, M.: Homology computation by reduction of chain complexes. Computers & Mathematics with Applications 35(4), 59–70 (1998)
20. Kropatsch, W.: Building irregular pyramids by dual graph contraction. In: IEE-Proc. Vision, Image and Signal Processing. pp. 366–374 (1995)
21. Matas, J., Chum, O., Martin, U., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proceedings of British Machine Vision Conference. vol. 1, pp. 384–393. London (2002)
22. Miettinen, K.: Nonlinear multiobjective optimization. Kluwer Academic Publishers, Boston (1999)
23. Mrozek, M., Batko, B.: Coreduction homology algorithm. Discrete Comput. Geom. 41, 96–118 (2009)
24. Peltier, S., Ion, A., Haxhimusa, Y., Kropatsch, W., Damiand, G.: Computing Homology Group Generators of Images Using Irregular Graph Pyramids. In: Graph-Based Representations in Pattern Recognition. LNCS, vol. 4538, pp. 283–294. Springer Berlin / Heidelberg, Alicante Spain (2007)
25. Uras, C., Verri, A.: Computing size functions from edge maps. International Journal of Computer Vision 23(2), 169–183 (1997)