Collapsing Rips complexes*

Dominique Attali† André Lieutier‡ David Salinas†

Abstract

Given a finite set of points that samples a shape in Euclidean space, the Rips complex of the points provides an approximation of the shape which can be used in manifold learning. Indeed, it suffices to compute the proximity graph of the points to encode the whole Rips complex as the latter is an example of flag completion. Recently, it has been proved that the Rips complex reflects the homotopy type of the shape when sufficiently densely sampled by the points. Unfortunately, the Rips complex is generally high-dimensional. In this paper, we focus on the simplification of Rips complexes that approximate manifolds with the goal of reducing the dimension of the complex to the one of the manifold. We first propose an algorithm that iteratively applies elementary operations that preserve both the homotopy type and the property of the complex to be a flag completion and then show how our algorithm performs on real datasets.

1 Introduction

Manifold learning aims at recovering low-dimensional structures hidden in high-dimensional data [7]. An example of data might be a collection of m by m images of a rigid body taken under different poses. The collection of images can be thought of as a point cloud in $\mathbb{R}^{m\times m}$. Assuming the space of images is equipped with a reasonable metric (possibly the Euclidean metric), we expect the points to be distributed over a 6-dimensional manifold corresponding to the group of rigid displacements (e.g. rotations and translations). A manifold learning algorithm should be able, given as input the points, to output a topologically consistent representation of that manifold. Typically, the representation can encode a simplicial complex which, in the ideal case, is homeomorphic to the underlying manifold.

Given a finite set of points that samples a shape in Euclidean space, a classical approach for building an approximation of the shape consists in returning the Rips complex of the points [4]. Formally, the *Rips complex* of a set of points P at scale α is the simplicial complex whose simplices are subsets of points in P with diameter at most 2α. Recently, it has been proved that the Rips complex reflects the homotopy type of the shape, assuming the shape has a positive reach and is sufficiently densely sampled by the points [2]. The Rips complex has the computational advantage to be a flag completion: it suffices to compute its 1-skeleton to encode the whole complex. Unfortunately, the Rips complex is generally high-dimensional so that the true dimension of the shape remains elusive in the representation. To retrieve the intrinsic dimension of the shape, we propose to simplify Rips complexes by repeatedly applying generalized vertex and edge collapses (see definition bellow). We propose and compare several heuristics for finding such a sequence of collapses.

2 Collapses

Given a simplicial complex K, an *elementary collapse* is the operation that removes a pair of simplices $(\sigma_{\min}, \sigma_{\max})$ assuming $\sigma_{\max}$ is the unique proper coface of $\sigma_{\min}$. The result is the simplicial complex $K \setminus \{\sigma_{\min}, \sigma_{\max}\}$ to which K deformation retracts. The reverse operation, which adds back the two simplices $\sigma_{\min}$ and $\sigma_{\max}$ is called an *elementary anti-collapse* and is clearly also a homotopy-preserving operation. A simplicial complex is said to be *collapsible* if it can be reduced to a single vertex by a finite sequence of elementary collapses. For instance, the closure of a simplex σ, $\mathrm{Cl}(\sigma) = \bigcup_{\emptyset\neq\tau\subset\sigma}\{\tau\}$, is collapsible. $\mathrm{Cl}(\sigma)$ is an example of cone. A *cone* is a simplicial complex K which contains a vertex o such that the following implication holds: $\sigma \in K \implies \sigma \cup \{o\} \in K$. Cones are also collapsible. Another elementary operation that we shall use is the *edge contraction*. The edge contraction $ab \mapsto c$ is the operation that identifies the two vertices $a \in K$ and $b \in K$ to the vertex c. It preserves the homotopy-type whenever $\mathrm{Lk}_K(ab) = \mathrm{Lk}_K(a) \cap \mathrm{Lk}_K(b)$ where $\mathrm{Lk}_K(\sigma) = \{\tau \in K, \tau \cap \sigma = \emptyset \text{ and } \tau \cup \sigma \in K\}$ designates the link of σ in K [1, 5].

We now list several possible generalizations of elementary collapses. To do so, we call the collection of simplices of K having σ as a face the *star* of σ and denote it as $\mathrm{St}_K(\sigma)$. Finally, we call the operation that removes $\mathrm{St}_K(\sigma_{\min})$ from K:

*Research partially supported by the French “Agence nationale pour la Recherche” under grant ANR-09-BLAN-0331-01 “Giga”.

†Gipsa-lab – CNRS UMR 5216, Grenoble, France. Firstname.Lastname@gipsa-lab.grenoble-inp.fr

‡Dassault systèmes, Aix-en-Provence, France. andre.lieutier@3ds.com

• a *(classical) collapse*: if the star of $\sigma_{\min}$ has a unique inclusion-maximal element $\sigma_{\max} \neq \sigma_{\min}$; the reverse operation is called a *(classical) anti-collapse.*

• an *(extended) collapse*: if the link of $\sigma_{\min}$ is a cone [1]; the reverse operation is called an *(extended) anti-collapse.*

• a *(generalized) collapse*: if the link of $\sigma_{\min}$ can be reduced to a point by a sequence of collapses, anti-collapses and homotopy-preserving edge contractions; the reverse operation is called a *(generalized) anti-collapse.*

All three collapses (classical, extended, generalized) preserve the homotopy-type. Deciding whether the operation that removes $\mathrm{St}_K(\sigma_{\min})$ from K is a classical or extended collapse can be done efficient (*i.e. in polynomial time*) using the data structure described in [1]. We will see shortly that deciding whether the operation that removes $\mathrm{St}_K(\sigma_{\min})$ from K is a generalized collapse is computationally more involved. Here, we will focus on Rips complexes which will allow us to design specific reduction sequence.

3 Simplifying Rips complexes whoses vertices approximate a manifold

The goal of this section is to present and compare strategies for simplifying a Rips complex whose vertex set samples a manifold. In the ideal case, we would like to get a complex homeomorphic to the manifold or at least whose dimension is as close as possible to that of the manifold. Throughout the section, we will assume that P is a point cloud that samples a d-dimensional manifold A embedded in D-dimensional Euclidean space and suppose that we can find a value of α such that $\mathrm{Rips}(P, \alpha)$ and A has the same homotopy type.

3.1 Simplification algorithm

The Rips complex is simplified in two stages: the first stage iteratively collapses vertices and the second stage iteratively collapses edges; see Algorithm 1. During the simplification, the complex remains a flag completion, since this property is not altered by collapsing vertices and edges. For $k \in \{0, 1\}$, stage k proceeds as follows. Initially, all k-dimensional simplices are stored in a priority queue Q. Each k-simplex receives as priority its diameter. During stage k, we iteratively take the k-simplex σ with highest priority and remove it from the current complex K together with all its cofaces whenever REDUCIBLE($\mathrm{Lk}_K(\sigma)$) returns true; see Algorithm 2. Ideally, we would like the function REDUCIBLE($\mathrm{Lk}_K(\sigma)$) to be true if and only if the operation that removes σ and all its cofaces is a generalized collapse. This means that ideally, we would like the function REDUCIBLE to take as input a simplicial complex L and returns true whenever there exists a sequence of homotopy-preserving elementary operations (collapses, anti-collapses and edge-contractions) that goes from L to a point and false otherwise. Unfortunately, the problem of deciding whether a complex L is reducible to a point by a sequence of elementary operation is NP-complete, even when we limit ourselves to elementary collapses and 3-dimensional complexes [6]. Instead, we will propose four more or less sophisticated heuristics to find such a sequence, drawing inspiration from the constructive proofs of [3] and sometimes taking advantage of the fact that L is a flag complex.

Algorithm 1 SIMPLIFY(Simplicial complex K)

```
SIMPLIFY(K,0,true) {Collapse the vertices of K}
SIMPLIFY(K,1,false) {Collapse the edges of K }
```

Algorithm 2 SIMPLIFY(Simplicial complex K, integer k, boolean reinsert)

```
Q = K^(k) \ K^(k-1)
while Q ≠ ∅ do
  Remove the simplex σ from Q with highest priority
  if REDUCIBLE (Lk_K(σ)) then
    K = K \ St_K(σ)
    if reinsert then
      Insert in Q the k-dimensional simplices
      whose link have changed
    end if
  end if
end while
```

Observe that the only difference between the two stages is that at stage 0, when we collapse a vertex, we reintroduce the vertices of its link in the priority queue. We do not do the same at stage 1 because, in our experiments, it slows the computation and does not improve the number of times we get a complex either homeomorphic to or with the same dimension as the manifold A.

3.2 Finding reduction sequences

We present four possible procedures that can be used in place of REDUCIBLE in Algorithm 2. Each procedure takes as input a simplicial complex L. The third strategy is the only one that requires its input to be a flag completion. The pseudo-code for each procedure can be found either in [1] or [3]. For later reference, strategies are numbered from (S1) to (S4). The intuition behind strategies (S2) and (S3) is that if the vertices of L sample a convex set densely enough then results in [3] ensure that those two strategies will succeed.

(S1) ISCONE: returns true if and only if L is a cone.

(S2) REDUCIBLE_BY_SWEEP: starts by picking a vertex o of L and tries to apply vertex and edge extended collapses to reduce L to o. To do so, the strategy computes a priority φ which is evaluated to $\varphi(a) = d(o, a) - \alpha$ for each vertex a of L and $\varphi(ab) = d(o, B(a, \alpha) \cap B(b, \alpha))$ for each edge ab of L. Then, we put all vertices of L except o in a priority queue Q and put all edges ab of L in Q iff the priority of ab is greater than the priority of its vertices a and b that is $\varphi(ab) > \min\{\varphi(a), \varphi(b)\}$. Finally the strategy tries to reduce L to o by performing extended collapses of all simplices in Q in the order of decreasing priority φ. The strategy returns true iff it manages to do so.

(S3) REDUCIBLE_BY_COMPLETION: applies a sequence of edge extended anti-collapses in the order of increasing length. If at some point, the result is a cone, returns true. If at some point, an edge could not be inserted, returns false.

(S4) REDUCIBLE_BY_EDGE_CONTRACTIONS: simplify L by applying a sequence of edge contractions $ab \mapsto \frac{a+b}{2}$ in the order of increasing length assuming $\mathrm{Lk}_L(ab) = \mathrm{Lk}_L(a) \cap \mathrm{Lk}_L(b)$ as explained in [5, 1]. Returns true iff the simplex L after simplification consists of a single vertex.

If we assume that P is initially a dense sampling of A, the vertices in the link of a simplex are likely to be close to a convex (at least at the beginning of the simplification). In this situation, it is proved that the complex is always reduced to a point with the two strategies (S2) and (S3) [3]. We now describe various computational experiments and the results we obtained.

4 Experiments

4.1 Data-sets

We present the four data-sets used in our experiments.

`Cat`. A collection of 72 images of a toy cat placed on a turntable and observed by a fixed camera. Each image has size $128^2 = 16384$ and can be thought of as a point-cloud that samples a 1-manifold in $\mathbb{R}^{16384}$. For this data-set, $d = 1$ and $D = 16384$.

`Sphere`. A sampling of a 2-sphere with 2646 points corrupted by noise. In this case, $d = 2$ and $D = 3$.

`Ramses`. A 3D scan data consisting of 193252 points measured on the surface of a statue representing Ramses II. The surface of the statue is homeomorphic to $\mathbb{S}^2$. For this data-set, $d = 2$ and $D = 3$.

`SO3`. A point set `SO3` $\subset \mathbb{R}^9$ with size 10000 that samples the special orthogonal group. Recall that this group is diffeomorphic to $\mathbb{RP}^3$ which is a 3-dimensional manifold that can be embedded in $\mathbb{R}^9$ by representing each rotation in 3D by a 3×3 matrix. We have $d = 3$ and $D = 9$.

P	$\sharp P$	size of $\mathrm{Rips}(P, \alpha)$	dimension of $\mathrm{Rips}(P, \alpha)$
`Cat`	72	$> 10^6$	19
`Sphere`	2646	$> 2 \times 10^9$	> 12
`Ramses`	$193, 252$	$> 6 \times 10^6$	14
`SO3`	$10, 000$	$> 2.8 \times 10^8$	16

Table 1: For each data-set P, we indicate the number of points in P, the number of simplices in $\mathrm{Rips}(P, \alpha)$ and the dimension of $\mathrm{Rips}(P, \alpha)$.

Table 1 gives for each data-set the number of points, the number of simplices in $\mathrm{Rips}(P, \alpha)$ and the dimension of $\mathrm{Rips}(P, \alpha)$.

4.2 Results

To compare our four strategies, we first perform the following experiment. We apply our simplification algorithm (Algorithm 2), using for REDUCIBLE the function which returns true iff one of the four strategies returns true. Let σ_j be the k-simplex removed at step j during the simplification. Let $s_k^x(i)$ be the number of times (Sx) returns false when applied to σ_j for j ranging over $\{1, \ldots, i\}$. In other words, $s_k^x(i)$ counts the number of times the strategy (Sx) has failed to find a sequence of reduction while another strategy had succeeded during the i first steps of the simplification process. In Figure 1, we plot s_0^x for $x \in [1, 4]$ (that is for all strategies), and for all data-sets. When collapsing edges, all strategies give the same answer for all data-sets except for `Ramses` and `SO3` thus we plot s_1^x only for these two data-sets, see Figure 1. We observe that (S4) seems to be the most efficient strategy : it finds a sequence of reduction whenever another strategy finds one when simplifying our four data-sets.

We now use our simplification algorithm with a fixed strategy for finding reduction sequences. In Table 2, we describe the complex K_{out} obtained after simplification using each of our 4 strategies in turn to find reduction sequences. As suggested by the previous experiment, the best results are obtained when using (S4). Indeed, using (S4), the result of the simplification K_{out} is a flag complex homeomorphic to the sampled manifold A, except for the data-set `SO3`. Still, in that case, we get a complex with the correct dimension. Future work will include a better understanding of the performances of strategy $(S4)$ together with the search of a condition ensuring that our algorithm outputs a complex homeomorphic to A.

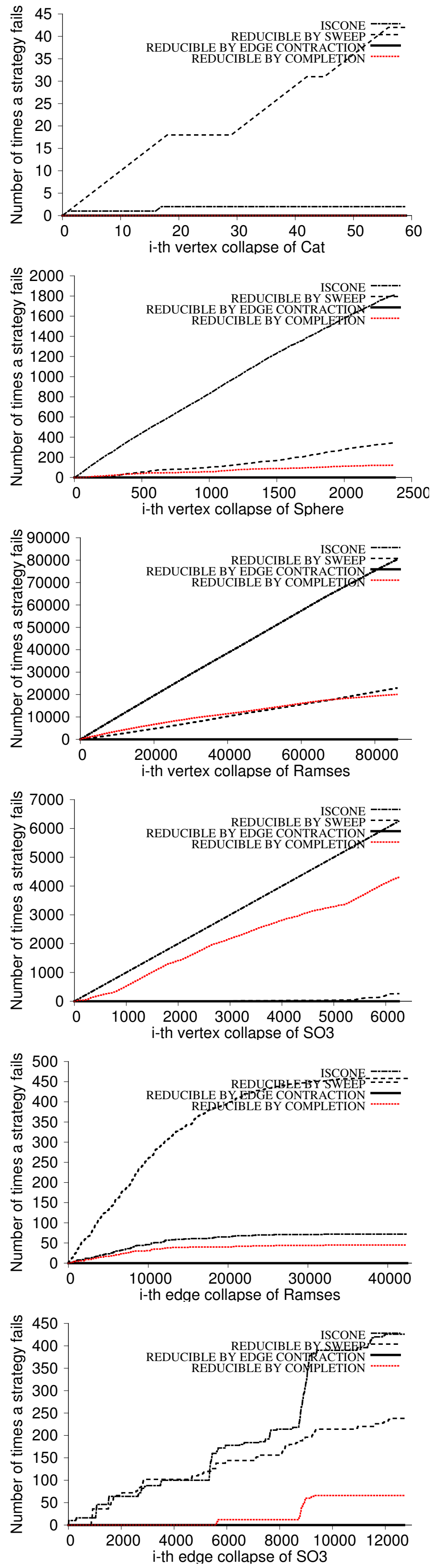

Figure 1: Number of failures for each strategy while performing vertex collapse (first four figures) and edge collapse (last two figures).

(S1)	$\dim(K_{\text{out}})$	$K_{\text{out}} \approx A$	running time
`Cat`	1	YES	1 s
`Sphere`	5	NO	1 min
`Ramses`	3	NO	25 min
`SO3`	6	NO	94 s
(S2)	$\dim(K_{\text{out}})$	$K_{\text{out}} \approx A$	running time
`Cat`	2	NO	3 min
`Sphere`	3	NO	6 min
`Ramses`	3	NO	180 min
`SO3`	4	NO	10 min
(S3)	$\dim(K_{\text{out}})$	$K_{\text{out}} \approx A$	running time
`Cat`	1	YES	1 s
`Sphere`	2	YES	2 min
`Ramses`	2	NO	160 min
`SO3`	4	NO	33 min
(S4)	$\dim(K_{\text{out}})$	$K_{\text{out}} \approx A$	running time
`Cat`	1	YES	2 s
`Sphere`	2	YES	2 min
`Ramses`	2	YES	150 min
`SO3`	3	NO	7 min

Table 2: Description of the simplicial complex output K_{out} when using each of the four strategies for finding sequences of reductions. We indicate the dimension of K_{out} together with the fact that it is homeomorphic or not to A and the computation time. All computation are done with a 2.8 GHz processor and 8 GB RAM.

References

[1] D. Attali, A. Lieutier, and D. Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. *International Journal of Computational Geometry and Applications (IJCGA)*, 22(4):279–303, 2012.

[2] D. Attali, A. Lieutier, and D. Salinas. Vietoris-Rips complexes also provide topologically correct reconstructions of sampled shapes. *Computational Geometry: Theory and Applications (CGTA)*, 2012.

[3] D. Attali, A. Lieutier, and D. Salinas. Collapsing rips complexes. In *29th Ann. Sympos. Comput. Geom.*, Rio de Janeiro, Brazil, 2013. Submitted.

[4] V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007.

[5] T. K. Dey, H. Edelsbrunner, and S. Guha. Computational topology. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*. AMS, Providence, 1999.

[6] M. Tancer. Recognition of collapsible complexes is np-complete. *CoRR*, abs/1211.6254, 2012.

[7] J. Tenenbaum, V. De Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.