A metric approach to shape comparison via multidimensional persistence

Patrizio Frosini[1,2]

[1]Department of Mathematics, University of Bologna, Italy
[2]ARCES - Vision Mathematics Group, University of Bologna, Italy
`frosini@dm.unibo.it`

Colloquium on Computer Graphics
PRIP - Vienna University of Technology
17 December 2010

Outline

1. **A Metric Approach to Shape Comparison**
2. **Size functions and persistent homology groups**
3. **A new lower bound for the Natural Pseudo-distance**

1. **A Metric Approach to Shape Comparison**
2. Size functions and persistent homology groups
3. A new lower bound for the Natural Pseudo-distance

Informal position of the problem

Let us start

from three examples of questions

about the concept of comparison...

Informal position of the problem

How similar are the colorings of these leaves?

Informal position of the problem

How similar are the Riemannian structures of these manifolds?

Informal position of the problem

How similar are the spatial positions of these wires?

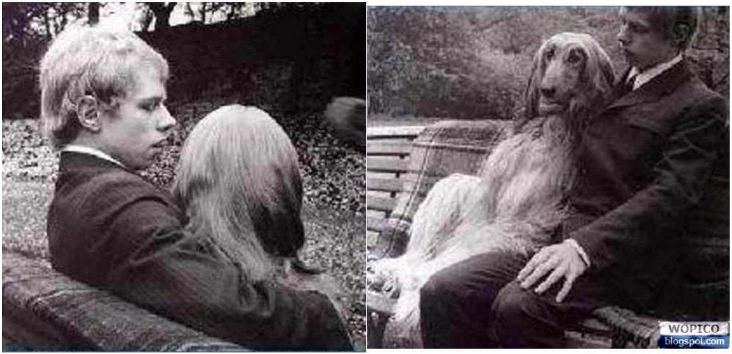
Informal position of the problem

Every comparison of properties involves the presence of

- an observer perceiving the properties
- a methodology to compare the properties

Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:



Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:

Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:

Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:

Julian Beever

Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:

Julian Beever

Informal position of the problem

The perception properties depend on the subjective interpretation of an observer:

Julian Beever

Informal position of the problem

The concept of shape is subjective and relative. It is based on the act of perceiving, depending on the chosen observer. Persistent perceptions are fundamental in order to approach this concept.

- "Science is nothing but perception." *Plato*
- "Reality is merely an illusion, albeit a very persistent one." *Albert Einstein*

Our formal setting

Our formal setting:

- Each perception is formalized by a pair $(X, \vec{\varphi})$, where X is a topological space and $\vec{\varphi}$ is a continuous function.
- X represents the set of observations made by the observer, while $\vec{\varphi}$ describes how each observation is interpreted by the observer.

Our formal setting

Example a Let us consider Computerized Axial Tomography, where for each unit vector v in the real plane a real number is obtained, representing the total amount of mass $\varphi(v)$ encountered by an X-ray beam directed like v. In this case the topological space X equals the set of all unit vectors in $\mathbb{R}^2$, i.e. S^1. The filtering function is $\varphi : S^1 \to \mathbb{R}$.

Example b Let us consider a rectangle R containing an image, represented by a function $\vec{\varphi} = (\varphi_1, \varphi_2, \varphi_3) : R \to \mathbb{R}^3$ that describes the RGB components of the colour for each point in the image. The filtering function is $\vec{\varphi} : R \to \mathbb{R}^3$.

Our formal setting

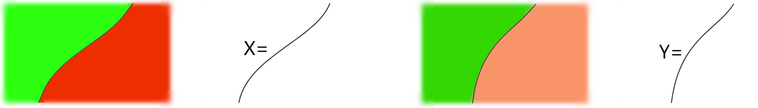
- Persistence is quite important. Without persistence (in space, time, with respect to the analysis level...) perception could have little sense. This remark compels us to require that
    - X is a topological space and $\vec{\varphi}$ is a continuous function; this function $\vec{\varphi}$ describes X from the point of view of the observer. It is called a measuring function (or filtering function).
    - The stable properties of the pair $(X, \vec{\varphi})$ are studied.

Our formal setting

- A possible objection: sometimes we have to manage discontinuous functions (e.g., colour).

Our formal setting

- A possible objection: sometimes we have to manage discontinuous functions (e.g., colour).
- An answer: in that case the topological space X can describe the discontinuity set, and persistence can concern the properties of this topological space with respect to a suitable measuring function.



As measuring functions we can take $\vec{\varphi} : X \to \mathbb{R}^2$ and $\vec{\psi} : Y \to \mathbb{R}^2$, where the components φ_1, φ_2 and ψ_1, ψ_2 represent the colors on each side of the considered discontinuity set.

Our formal setting

A categorical way to formalize our approach

Let us consider a category $\mathcal{C}$ such that

- The objects of $\mathcal{C}$ are the pairs $(X, \vec{\varphi})$ where X is a compact topological space and $\vec{\varphi} : X \to \mathbb{R}^k$ is a continuous function.
- The set $Hom\left((X, \vec{\varphi}), (Y, \vec{\psi})\right)$ of all morphisms between the objects $(X, \vec{\varphi})$, $(Y, \vec{\psi})$ is a subset of the set of all homeomorphisms between X and Y (possibly empty).

If $Hom\left((X, \vec{\varphi}), (Y, \vec{\psi})\right)$ is not empty we say that the objects $(X, \vec{\varphi})$, $(Y, \vec{\psi})$ are comparable.

Our formal setting

Do not compare apples and oranges...

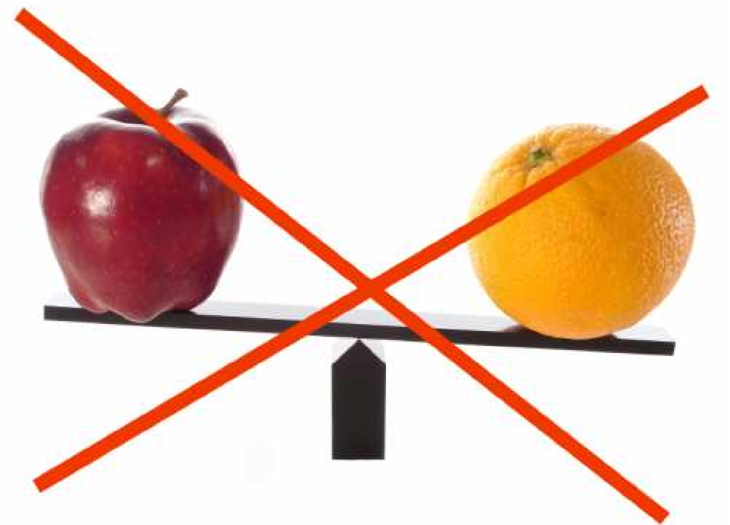
Remark: $Hom\left((X,\vec{\varphi}),(Y,\vec{\psi})\right)$ can be empty also in case X and Y are homeomorphic.

Example:

- Consider a segment $X = Y$ embedded into $\mathbb{R}^3$ and consider the set of observations given by measuring the colour $\vec{\varphi}(x)$ and the triple of coordinates $\vec{\psi}(x)$ of each point x of the segment.
- It does not make sense to compare the perceptions $\vec{\varphi}$ and $\vec{\psi}$. In other words the pairs $(X,\vec{\varphi})$ and $(Y,\vec{\psi})$ are not comparable, even if $X = Y$.
- We express this fact by setting $Hom\left((X,\vec{\varphi}),(Y,\vec{\psi})\right) = \emptyset$.

Our formal setting

Do not compare apples and oranges...



Our formal setting

We can now define the following (extended) pseudo-metric:

$$\delta\left((X,\vec{\varphi}),(Y,\vec{\psi})\right)=\inf_{h\in Hom\left((X,\vec{\varphi}),(Y,\vec{\psi})\right)}\max_i \max_{x\in X}|\varphi_i(x)-\psi_i\circ h(x)|$$

if $Hom\left((X,\vec{\varphi}),(Y,\vec{\psi})\right)\neq\emptyset$, and $+\infty$ otherwise.
We shall call $\delta\left((X,\vec{\varphi}),(Y,\vec{\psi})\right)$ the natural pseudo-distance between $(X,\vec{\varphi})$ and $(Y,\vec{\psi})$.

The functional $\Theta(h)=\max_i \max_{x\in X}|\varphi_i(x)-\psi_i\circ h(x)|$ represents the "cost" of the matching between observations induced by h. The lower this cost, the better the matching between the two observations.

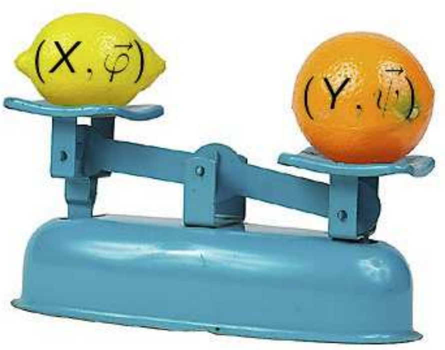
Our formal setting

- The natural pseudo-distance δ measures the dissimilarity between the perceptions expressed by the pairs $(X, \vec{\varphi})$, $(Y, \vec{\psi})$.
- The value δ is small if and only if we can find a homeomorphism between X and Y that induces a small change of the measuring function (i.e., of the shape property we are interested to study).

- For more information:
- P. Donatini, P. Frosini, *Natural pseudodistances between closed manifolds*, Forum Mathematicum, 16 (2004), n. 5, 695-715.
- P. Donatini, P. Frosini, *Natural pseudodistances between closed surfaces*, Journal of the European Mathematical Society, 9 (2007), 331-353.

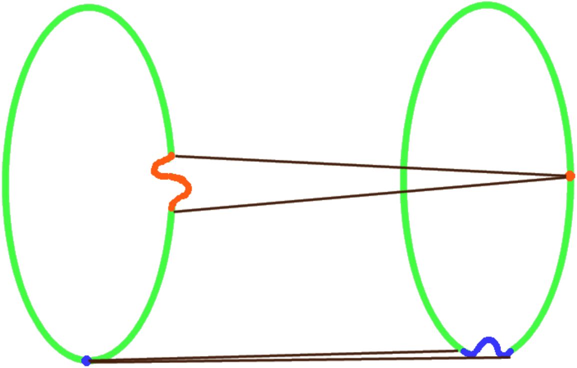
Our formal setting

In plain words, the natural pseudo-distance δ is obtained by trying to match the observations (taken in the topological spaces X and Y), in a way that minimizes the change of properties that the observer judges relevant (the filtering functions $\vec{\varphi}$ and $\vec{\psi}$).



Our formal setting

Why do we just consider homeomorphisms between X and Y?
Why couldn't we use, e.g., relations between X and Y?



Our formal setting

The following result suggests not to do that:

Non-existence Theorem

Let $\mathcal{M}$ be a closed Riemannian manifold. Call H the set of all homeomorphisms from $\mathcal{M}$ to $\mathcal{M}$. Let us endow H with the uniform convergence metric d_{UC}: $d_{UC}(f,g) = \max_{x\in\mathcal{M}} d_{\mathcal{M}}(f(x), g(x))$ for every $f, g \in H$, where $d_{\mathcal{M}}$ is the geodesic distance on $\mathcal{M}$.
Then (H, d_{UC}) cannot be embedded in any compact metric space (K, d) endowed with an internal binary operation $\bullet$ that extends the usual composition $\circ$ between homeomorphisms in H and commutes with the passage to the limit in K.

In particular, we cannot embed H into the set of binary relations on $\mathcal{M}$.

Our formal setting

What is shape, in our approach?

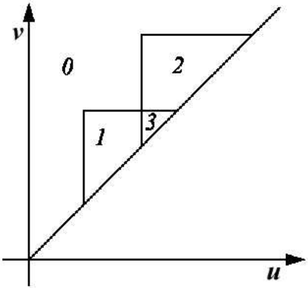
Shape is seen as an (unknown) pseudo-metric d expressed by an observer. We just try to approximate it:

- An observer communicates the values $d(\bar{O}, \bar{O}')$ for some pairs $(\bar{O}, \bar{O}')$ of "objects" (in the generic sense);
- We choose a functional F, associating each object O to a set of observations $\{(X_i, \varphi_i)\}$. The functional F is chosen in such a way that the distance between the values $d(\bar{O}, \bar{O}')$ and $\max_i \delta\left((\bar{X}_i, \bar{\varphi}_i), (\bar{X}_i', \bar{\varphi}_i')\right)$ is minimized for the pairs $(\bar{O}, \bar{O}')$ at which the observer has expressed her opinion.
- We *hope* that the distance between the values $d(O, O')$ and $\max_i \delta\left((X_i, \varphi_i), (X_i', \varphi_i')\right)$ is "small" for *every* pair (O, O').

Obviously, this is just a program for mathematical research, since there is no general rule to choose the functional F, at this time.

Natural pseudo-distance and size functions

- The natural pseudo-distance is usually difficult to compute.
- Lower bounds for the natural pseudo-distance δ can be obtained by computing the size functions.



Main definitions:

Given a topological space X and a continuous function $\vec{\varphi}: X \to \mathbb{R}^k$,

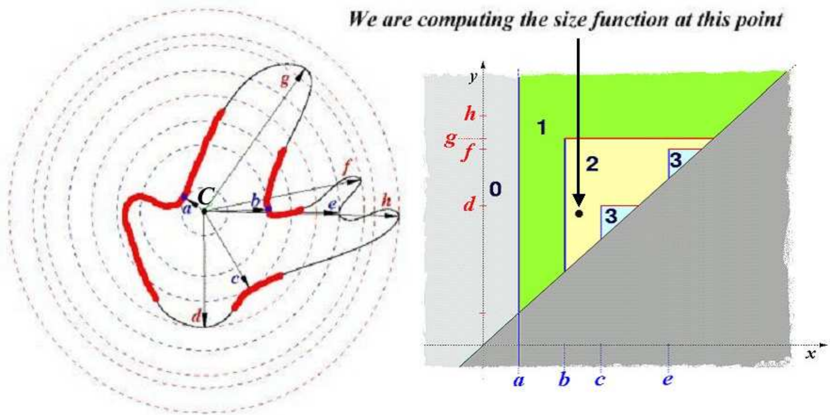
Lower level sets

For every $\vec{u} \in \mathbb{R}^k$, $X\langle\vec{\varphi} \preceq \vec{u}\rangle = \{x \in X : \vec{\varphi}(x) \preceq \vec{u}\}$.
($(u_1, \ldots, u_k) \preceq (v_1, \ldots, v_k)$ means $u_j \leq v_j$ for every index j.)

Definition (F. 1991)

The Size Function of $(X, \vec{\varphi})$ is the function ℓ that takes each pair $(\vec{u}, \vec{v})$ with $\vec{u} \prec \vec{v}$ to the number $\ell(\vec{u}, \vec{v})$ of connected components of the set $X\langle\vec{\varphi} \preceq \vec{v}\rangle$ that contain at least one point of the set $X\langle\vec{\varphi} \preceq \vec{u}\rangle$.

Example of a size function



We observe that each size function can be described by giving a set of points (vertices of triangles in figure). sizeshow.jar+cerchio.avi

Persistent homology groups and size homotopy groups

Size functions have been generalized by Edelsbrunner and al. to homology in higher degree (i.e., counting the number of holes instead of the number of connected components). This theory is called Persistent Homology:

H. Edelsbrunner, D. Letscher, A. Zomorodian, *Topological persistence and simplification*, Discrete & Computational Geometry, vol. 28, no. 4, 511–533 (2002).

Size functions have been also generalized to size homotopy groups:

P. Frosini, M. Mulazzani, *Size homotopy groups for computation of natural size distances*, Bulletin of the Belgian Mathematical Society, vol. 6, no. 3, 455–464 (1999).

Some important theoretical facts:

- The theory of size functions for filtering functions taking values in $\mathbb{R}^k$ can be reduced to the case of size functions taking values in $\mathbb{R}$, by a suitable foliation of their domain;
- On each leaf of the foliation, size functions are described by a collection of points (the vertices of the triangles seen previously);
- Size functions can be compared by measuring the difference between these collections of points, by a matching distance;
- Size functions are stable with respect to perturbations of the filtering functions (measured via the max-norm).

The same statements hold for persistent homology groups.

The distance d_T

Definition

Let X, Y be two topological spaces, and $\vec{\varphi}: X \to \mathbb{R}^k$, $\vec{\psi}: Y \to \mathbb{R}^k$ two continuous filtering functions. Let $\ell_{\vec{\varphi}}$ and $\ell_{\vec{\psi}}$ the size functions associated with the pairs $(X, \vec{\varphi})$ and $(Y, \vec{\psi})$, respectively. Let us consider the set E of all $\epsilon \geq 0$ such that, setting $\vec{\epsilon} = (\epsilon, \ldots, \epsilon) \in \mathbb{R}^k$, $\ell_{\vec{\psi}}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon}) \leq \ell_{\vec{\varphi}}(\vec{u}, \vec{v})$ and $\ell_{\vec{\varphi}}(\vec{u} - \vec{\epsilon}, \vec{v} + \vec{\epsilon}) \leq \ell_{\vec{\psi}}(\vec{u}, \vec{v})$ for every $\vec{u} \prec \vec{v}$. We define $d_T\left(\ell_{\vec{\varphi}}, \ell_{\vec{\psi}}\right)$ equal to $\inf E$ if E is not empty, and equal to ∞ otherwise.

This definition can be extended to persistent homology groups (possibly with torsion), substituting the previous inequalities with the existence of suitable surjective homomorphisms between groups.

d_T is a stable distance

Theorem

The function d_T is a distance. Moreover, if X, Y are two compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^k$, $\vec{\psi} : Y \to \mathbb{R}^k$, then

$$d_T\left(\ell_{\vec{\varphi}}, \ell_{\vec{\psi}}\right) \leq \delta\left((X, \vec{\varphi}), (Y, \vec{\psi})\right).$$

This theorem allows us to get lower bounds for the natural pseudo-distance, which is intrinsically difficult to compute.

d_T is a stable distance

From the previous theorem, two useful corollaries follow:

Corollary

Let X, Y be two compact topological spaces endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^k$, $\vec{\psi} : Y \to \mathbb{R}^k$. If two pairs $(\vec{u}, \vec{v})$, $(\vec{u}', \vec{v}')$ exist such that $\vec{u} \prec \vec{v}$, $\vec{u}' \prec \vec{v}'$ and $\ell_{\vec{\psi}}(\vec{u}', \vec{v}') > \ell_{\vec{\varphi}}(\vec{u}, \vec{v})$, then

$$\delta\left((X, \vec{\varphi}), (Y, \vec{\psi})\right) \geq \min_i \min\{u_i - u_i', v_i' - v_i\}.$$

Corollary

Let X be a compact topological space endowed with two continuous functions $\vec{\varphi} : X \to \mathbb{R}^k$, $\vec{\varphi}' : X \to \mathbb{R}^k$. Then $d_T\left(\ell_{\vec{\varphi}}, \ell_{\vec{\varphi}'}\right) \leq \|\vec{\varphi} - \vec{\varphi}'\|_\infty$.

Conclusions

- We have illustrated the concept of natural pseudo-distance δ, seen as a mathematical tool to compare shape properties;
- Some theoretical results about δ have been recalled;
- A new lower bound for δ has been illustrated.