An observer-oriented approach to topological data analysis

Part 2: The algebra of group invariant non-expansive operators and its application in the project GIPHOD - Group Invariant Persistent Homology Online Demonstrator

Patrizio Frosini

Department of Mathematics and ARCES, University of Bologna
patrizio.frosini@unibo.it

Second School/Conference in TDA,
Stochastic Topology and related topics
Querétaro, 7-11 December 2015

Outline

Group invariant non-expansive operators from Φ to $\Psi \neq \Phi$

How can we produce new group invariant non-expansive operators?

GIPHOD

Group invariant non-expansive operators from Φ to $\Psi \neq \Phi$

How can we produce new group invariant non-expansive operators?

GIPHOD

Could we consider operators from Φ to Ψ?

In the previous talk we have considered group invariant non-expansive operators (GINOs) from Φ to Φ.

In several applications it can be useful to consider operators that take the space Φ to a different topological space Ψ.

As an example, we can consider the operator F taking each function $\varphi \in \Phi = C^0([0,1] \times [0,1], \mathbb{R})$ to the function $F(\varphi) \in \Psi = C^0([0,1], \mathbb{R})$ defined by setting $F(\varphi)(x) := \max\{\varphi(x,y) : y \in [0,1]\}$ for every $x \in [0,1]$.

In the next slides we will show how our approach based on GINOs from Φ to Φ can be extended to GINOs from Φ to $\Psi \neq \Phi$.

Perception spaces

We are interested in the pairs (Φ, G) where

- Φ is a set of continuous functions from a compact topological space X to $\mathbb{R}$, containing at least the constant functions.
- G is a subgroup of the topological group $\mathrm{Homeo}(X)$.
- Φ is closed under the right-action of G on Φ (i.e., $\varphi \in \Phi \implies \varphi \circ g \in \Phi$.

We will say that (Φ, G) is a perception space.

Natural pseudo-metric d_G

For each perception space (Φ, G) we can consider the natural pseudo-metric d_G associated with the group G.

We recall its definition:

$$d_G(\varphi, \psi) = \inf_{g \in G} \max_{x \in X} |\varphi(x) - \psi(g(x))|$$

Group-invariant non-expansive operators for $\Phi \neq \Psi$

Now we have to extend the concept of GINO to the case $\Phi \neq \Psi$. Let us assume that (Φ, G) and (Ψ, H) are two perception spaces and that a homomorphism $T : G \to H$ is given.

Each function $F : \Phi \to \Psi$ is called a group-invariant non-expansive operator (GINO) from (Φ, G) to (Ψ, H) with respect to T if

(a) $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for every $\varphi \in \Phi$ and every $g \in G$;

(b) $\|F(\varphi_1) - F(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty$ for every $\varphi_1, \varphi_2 \in \Phi$ (i.e., F is a 1-Lipschitzian operator from Φ to Ψ).

An example

As an example, think of the following case. Let us consider the perception spaces (Φ, G), (Ψ, H), where

- Φ is the set of all continuous functions from the disk $D^2 := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ to $\mathbb{R}$
- Ψ is the set of all continuous functions from $S^1 := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ to $\mathbb{R}$
- G is the group of all rotations of D^2
- H is the group of all rotations of S^1.

Let us set $F(\varphi) := \psi$ where $\psi(s) := \int_0^1 \varphi(ts)\ dt$, and $T(g) := g_{|S^1}$.
F is a group-invariant non-expansive operator from (Φ, G) to (Ψ, H) with respect to T.

A distance between GINOs

The symbol $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$ will be used to denote the collection of all GINOs from (Φ, G) to (Ψ, H) with respect to T.

If $\mathscr{F} \neq \emptyset$ is a subset of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$ and Ψ is bounded, then we can consider the function

$$d_{\mathscr{F}}(F_1, F_2) := \max_{\varphi \in \Phi} \|F_1(\varphi) - F_2(\varphi)\|_\infty$$

from $\mathscr{F} \times \mathscr{F}$ to $\mathbb{R}$.

Proposition

$d_{\mathscr{F}}$ is a metric on $\mathscr{F}$.

The pseudo-metric $D^{\mathscr{F}}_{\mathrm{match}}$

Let $\mathscr{F} \neq \emptyset$ be a subset of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$.
For every $\varphi_1, \varphi_2 \in \Phi$ we still set

$$D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2) := \sup_{F \in \mathscr{F}} d_{match}(\rho_k(F(\varphi_1)), \rho_k(F(\varphi_2)))$$

where $\rho_k(\psi)$ denotes the persistent Betti number function (i.e. the rank invariant) of ψ in degree k, while d_{match} denotes the usual bottleneck distance that is used to compare the persistence diagrams associated with $\rho_k(F(\varphi_1))$ and $\rho_k(F(\varphi_2))$.

Proposition

$D^{\mathscr{F}}_{match}$ is a G-invariant and stable pseudo-metric on Φ.

The G-invariance of $D^{\mathscr{F}}_{match}$ means that for every $\varphi_1, \varphi_2 \in \Phi$ and every $g \in G$ the equality $D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2 \circ g) = D^{\mathscr{F}}_{\mathrm{match}}(\varphi_1, \varphi_2)$ holds.

Relation between $D_{match}^{\mathscr{F}((\Phi,G),(\Psi,H),T:G\to H)}$ and d_G

The main link between GINOs and persistent homology is given by the following result.

Theorem

$D_{match}^{\mathscr{F}((\Phi,G),(\Psi,H),T:G\to H)} = d_G$.

This result suggests that G-invariant persistent homology can be used to approximate the natural pseudo-distance also in the case $\Phi \neq \Psi$.

Compactness of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$

Also in the case $\Phi \neq \Psi$ the compactness of our spaces of functions implies the compactness of the space of all GINOs:

Theorem

If the metric spaces Φ, Ψ are compact, then also the metric space $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$ is compact.

Compactness of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$

Similarly to what happened in the case $\Phi = \Psi$, the compactness of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$ suggests a method to approximate $D_{match}^{\mathscr{F}((\Phi,G),(\Psi,H),T:G\to H)}$ (and hence d_G).

Corollary

Assume that the metric spaces Φ, Ψ are compact with respect to their sup-norms. Let $\mathscr{F}$ be a non-empty subset of $\mathscr{F}((\Phi, G), (\Psi, H), T : G \to H)$. For every $\varepsilon > 0$, a finite subset $\mathscr{F}^*$ of $\mathscr{F}$ exists, such that

$$\left| D_{match}^{\mathscr{F}^*}(\varphi_1, \varphi_2) - D_{match}^{\mathscr{F}}(\varphi_1, \varphi_2) \right| \leq \varepsilon$$

for every $\varphi_1, \varphi_2 \in \Phi$.

Group invariant non-expansive operators from Φ to $\Psi \neq \Phi$

How can we produce new group invariant non-expansive operators?

GIPHOD

Composition of GINOs

Our approach to G-invariant TDA is based on the availability of GINOs.

How could we build new GINOs from other GINOs?

A simple method consists in producing new GINOs by composition of other GINOs:

Proposition

If F_1 is a GINO from (Φ, G) to (Ψ, H) with respect to T and F_2 is a GINO from (Ψ, H) to (χ, K) with respect to S, then $F := F_2 \circ F_1$ is a GINO from (Φ, G) to (χ, K) with respect to $S \circ T$.

Building GINOs via 1-Lipschitzian functions

We can also produce new GINOs by means of a 1-Lipschitzian function applied to other GINOs:

Proposition

Assume that two perception spaces (Φ, G), (Ψ, H) *are given. Let* $\mathscr{L}$ *be a* 1*-Lipschitzian map from* $\mathbb{R}^n$ *to* $\mathbb{R}$, *where* $\mathbb{R}^n$ *is endowed with the usual norm* $\|(x_1, \ldots, x_n)\|_\infty = \max_{1 \leq i \leq n} |x_i|$. *Assume also that* $F_1, \ldots, F_n$ *are GINOs from* (Φ, G) *to* (Ψ, H) *with respect to* T.
Let us define $\mathscr{L}^*(F_1, \ldots, F_n) : \Phi \to \Psi$ *by setting*
$\mathscr{L}^*(F_1, \ldots, F_n)(\varphi)(x) := \mathscr{L}(F_1(\varphi)(x), \ldots, F_n(\varphi)(x))$. *Then the operator* $\mathscr{L}^*(F_1, \ldots, F_n)$ *is a GINO from* (Φ, G) *to* (Ψ, H) *with respect to* T.

From this proposition the following three results follow.

Building new GINOs via translations, weighted averages and the maximum operator

Proposition (Translation)

Assume F is a GINO from (Φ, G) to (Ψ, H) with respect to T, and $b \in \mathbb{R}$. Then the operator $F_b := F - b$ is also a GINO from (Φ, G) to (Ψ, H) with respect to T.

Proposition (Weighted average)

Assume that $F_1, \ldots, F_n$ are GINOs from (Φ, G) to (Ψ, H) with respect to T, and $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. Then $F = \sum_{i=1}^n a_i F_i$ is also a GINO from (Φ, G) to (Ψ, H) with respect to T.

Building new GINOs via translations, weighted averages and the maximum operator

Proposition (Maximum)

Assume $F_1, \ldots, F_n$ *are GINOs from* (Φ, G) *to* (Ψ, H) *with respect to* T *taking values in* $\mathbb{R}$. *Then the operator* $F = \max_i F_i$ *is also a GINO from* (Φ, G) *to* (Ψ, H) *with respect to* T.

An interesting GINO in kD persistent homology

Previous propositions imply the following statement.

Proposition

Assume $F_1, \ldots, F_n$ are GINOs from (Φ, G) to (Ψ, H) with respect to T. Assume also that $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_n) \in \mathbb{R}^n$, with $a_1, \ldots, a_n > 0$, $\sum_{i=1}^n a_i = 1$ and $\sum_{i=1}^n b_i = 0$. Then the operator

$$F = \min_j a_j \cdot \max \left\{ \frac{F_1 - b_1}{a_1}, \ldots, \frac{F_n - b_n}{a_n} \right\}$$

is also a GINO from (Φ, G) to (Ψ, H) with respect to T.

An interesting GINO in kD persistent homology

The previous proposition shows an interesting link between GINOs and the operator used to reduce multidimensional persistent Betti numbers to 1D persistent Betti numbers.

We recall that this operator takes the $\mathbb{R}^n$-valued function $\varphi = (\varphi_1, \ldots, \varphi_n)$ to the $\mathbb{R}$-valued functions

$$\varphi_{a,b} := \min_j a_j \cdot \max\left\{\frac{\varphi_1 - b_1}{a_1}, \ldots, \frac{\varphi_n - b_n}{a_n}\right\}$$

with $a_1, \ldots, a_n > 0$, $\sum_{i=1}^n a_i = 1$ and $\sum_{i=1}^n b_i = 0$. The key point is that each sublevel set of φ can be represented as a sublevel set of a filtering function $\varphi_{a,b}$, so that our operator can be used to reduce a multidimensional filtration to a collection of 1D-filtrations.

An interesting GINO in kD persistent homology

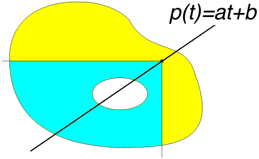


Figure: The collection of the 1D-filtrations associated with the lines $p(t) = at + b$ such that $a_1, a_2 > 0$, $a_1 + a_2 = 1$ and $b_1 + b_2 = 0$ is equivalent to the 2D-filtration associated with the filtering function $p \mapsto (x(p), y(p))$.

[A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, C. Landi, *Betti numbers in multidimensional persistent homology are stable functions*, Mathematical Methods in the Applied Sciences, vol. 36 (2013), 1543–1557.]

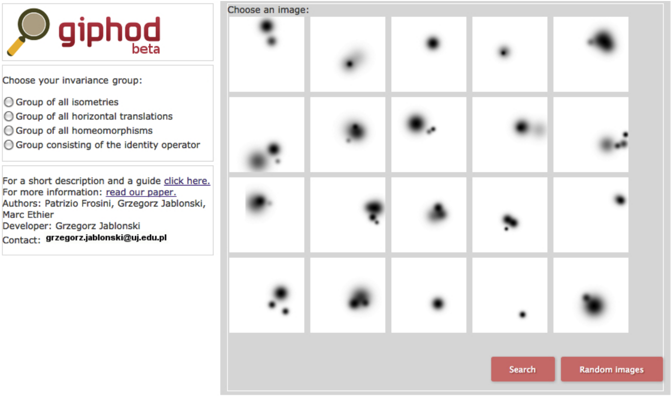
Group invariant non-expansive operators from Φ to $\Psi \neq \Phi$

How can we produce new group invariant non-expansive operators?

GIPHOD

GIPHOD

Joint project with Grzegorz Jabłoński (Jagiellonian University - Kraków) and Marc Ethier (Université de Saint-Boniface - Canada)



GIPHOD (Group Invariant Persistent Homology On-line Demonstrator)

GIPHOD is an on-line demonstrator, allowing the user to choose an image and an invariance group. GIPHOD searches for the most similar images in the dataset, with respect to the chosen invariance group.

Purpose: to show the use of G-invariant persistent homology for image comparison.

Dataset: 10.000 grey-level synthetic images obtained by adding randomly chosen bell-shaped functions. The images are coded as functions from $\mathbb{R}^2 \to [0,1]$.

GIPHOD WILL BE AVAILABLE IN JANUARY 2016.

The beta version of GIPHOD can be tested at `http://giphod.capdnet.ii.uj.edu.pl`.

Thanks to everyone that will give suggestions for improvement (please send them to `grzegorz.jablonski@uj.edu.pl`)

GIPHOD: Examples for the group of isometries

We will now show some results obtained by GIPHOD when the invariance group G is the group of isometries:

Some data about the pseudo-metric $D^{\mathscr{F}}_{match}$ in this case:

- Mean distance between images: 0.2984;
- Standard deviation of distance between images: 0.1377;
- Number of GINOs that have been used: 5.

GIPHOD: Examples for the group of isometries

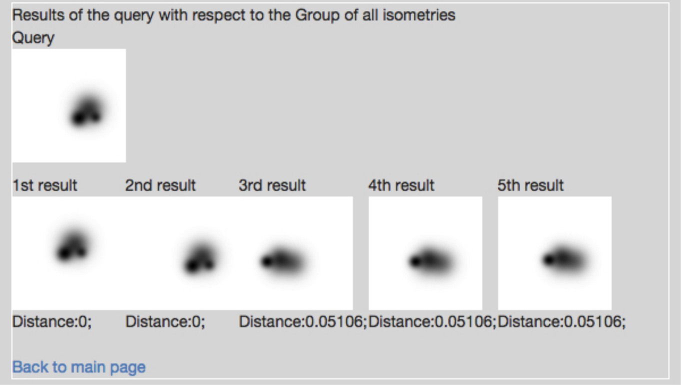
Here are the five GINOs that we used for the invariance group of isometries:

- $F(\varphi) = \varphi$.
- $F(\varphi) :=$ constant function taking each point to the value $\int_{\mathbb{R}^2} \varphi(x)\ dx$.
- $F(\varphi)$ defined by setting

$$F(\varphi)(x) := \int_{\mathbb{R}^2} \varphi(x-y) \cdot \beta\left(\|y\|_2\right)\ dy$$

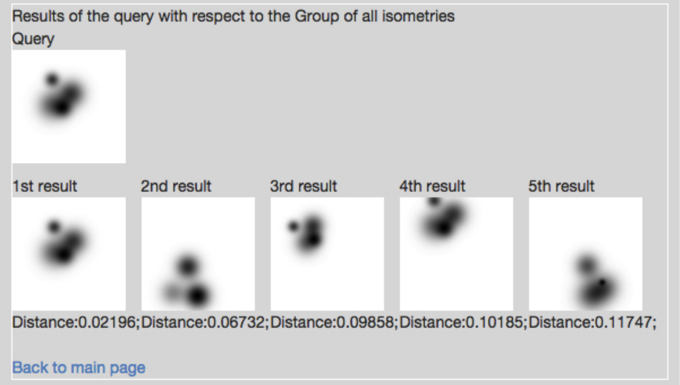
  where $\beta : \mathbb{R} \to \mathbb{R}$ is an integrable function with $\int_{\mathbb{R}^2} \left|\beta\left(\|y\|_2\right)\right|\ dy \leq 1$
  (we have used three operators of this kind).

GIPHOD: Examples for the group of isometries



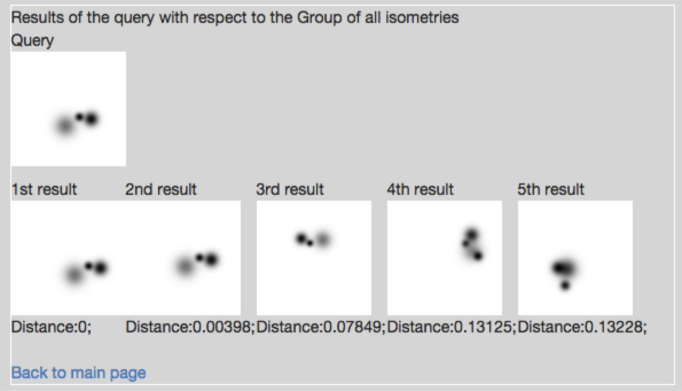
Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

GIPHOD: Examples for the group of isometries



Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

GIPHOD: Examples for the group of isometries



Mean distance: 0.2984; Standard deviation of distance: 0.1377; Number of GINOs: 5.

GIPHOD: Examples for the group $\mathrm{Homeo}(\mathbb{R}^2)$

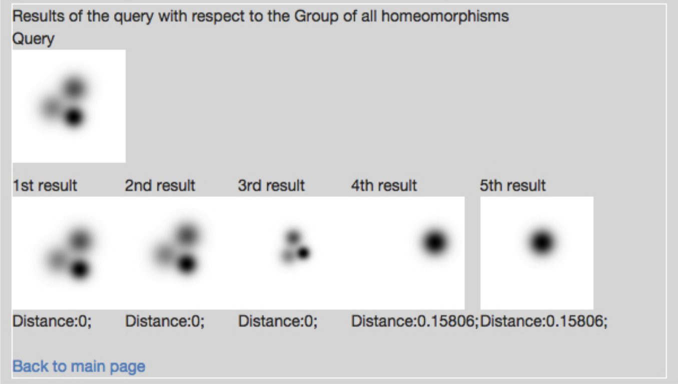
We will now show some results obtained by GIPHOD when the invariance group G is the group $\mathrm{Homeo}(\mathbb{R}^2)$:

Some data about the pseudo-metric $D^{\mathscr{F}}_{match}$ in this case:

- Mean distance between images: 0.2357;
- Standard deviation of distance between images: 0.1721;
- Number of GINOs that have been used: 1.

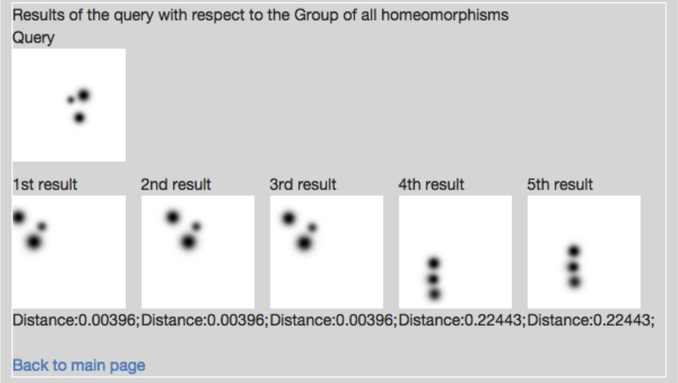
For the invariance group $\mathrm{Homeo}(\mathbb{R}^2)$ we have used just the identity operator.

GIPHOD: Examples for the group $\mathrm{Homeo}(\mathbb{R}^2)$



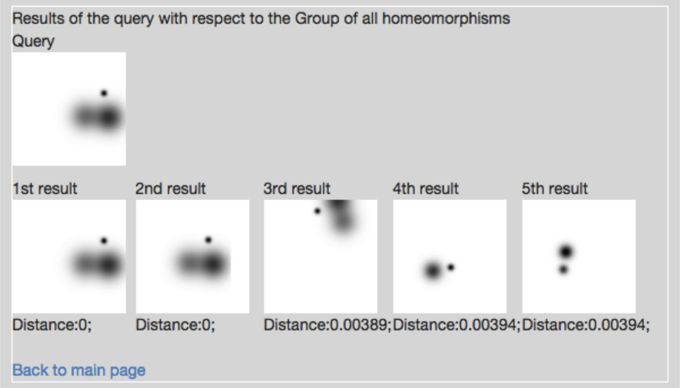
Mean distance: 0.2357; Standard deviation of distance: 0.1721; Number of GINOs: 1.

GIPHOD: Examples for the group $\mathrm{Homeo}(\mathbb{R}^2)$



Mean distance: 0.2357; Standard deviation of distance: 0.1721; Number of GINOs: 1.

GIPHOD: Examples for the group $\mathrm{Homeo}(\mathbb{R}^2)$



Mean distance: 0.2357; Standard deviation of distance: 0.1721; Number of GINOs: 1.

GIPHOD: Group of all horizontal translations

We will now show some results obtained by GIPHOD when the invariance group G is the group of all horizontal translations:

Some data about the pseudo-metric $D^{\mathscr{F}}_{match}$ in this case:

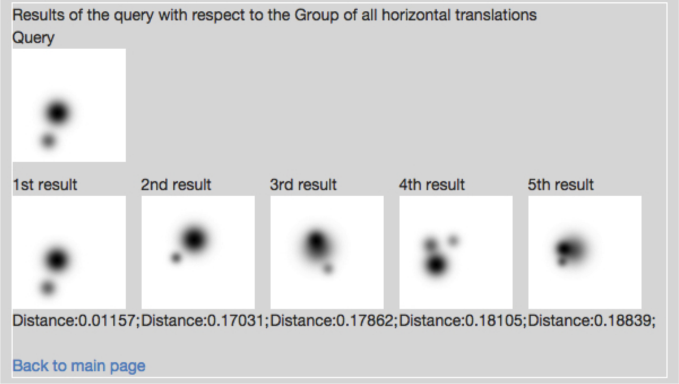
- Mean distance between images: 0.4837;
- Standard deviation of distance between images: 0.1475;
- Number of GINOs that have been used: 10.

GIPHOD: Examples for the group of all horizontal translations

Here are the ten GINOs that we have used for the invariance group of all horizontal translations:

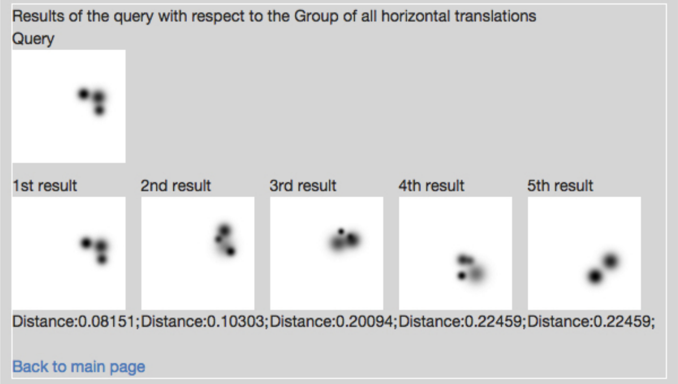
- $F(\varphi) = \varphi$.
- $F(\varphi) :=$ constant function taking each point to $\int_{\mathbb{R}^2} \varphi(x)\ dx$.
- $F(\varphi)$ defined by setting $F(\varphi)(x) := \int_{\mathbb{R}^2} \varphi(x-y) \cdot \beta\left(\|y\|_2\right)\ dy$ with $\beta : \mathbb{R} \to \mathbb{R}$ and $\int_{\mathbb{R}^2} \left|\beta\left(\|y\|_2\right)\right|\ dy \leq 1$
  (we have used two operators of this kind).
- Let us choose a continuous real-valued function $\chi(y_1, y_2)$ that is 1-Lipschitz in the variable y_1, for every y_2. Let us define $F(\varphi)$ by setting $F(\varphi)(x_1, x_2) := \chi(\varphi(x_1, x_2), x_2)$
  (we have used six operators of this kind).

GIPHOD: Group of all horizontal translations



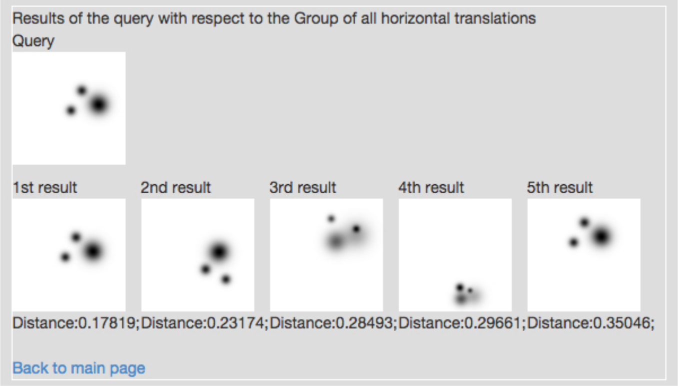
Mean distance: 0.4837; Standard deviation of distance: 0.1475; Number of GINOs: 10.

GIPHOD: Group of all horizontal translations



Mean distance: 0.4837; Standard deviation of distance: 0.1475; Number of GINOs: 10.

GIPHOD: Group of all horizontal translations



Mean distance: 0.4837; Standard deviation of distance: 0.1475; Number of GINOs: 10.

GIPHOD: Examples for the identical group

Finally, we will show some results obtained by GIPHOD when the invariance group G is the identical group:

Some data about the pseudo-metric $D^{\mathscr{F}}_{match}$ in this case:

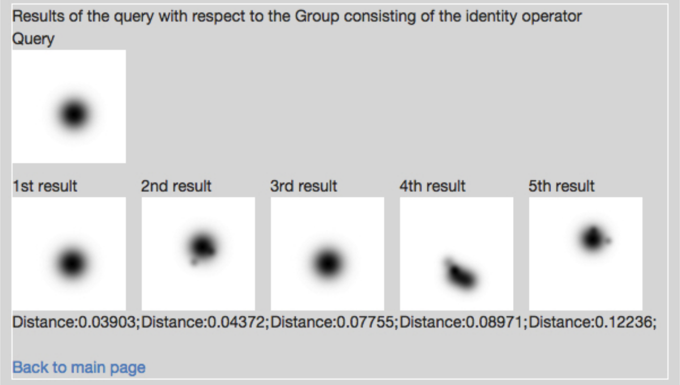
- Mean distance between images: 0.6412;
- Standard deviation of distance between images: 0.1257;
- Number of GINOs that have been used: 15.

GIPHOD: Examples for the identical group

Here are the fifteen GINOs used for the identical group:

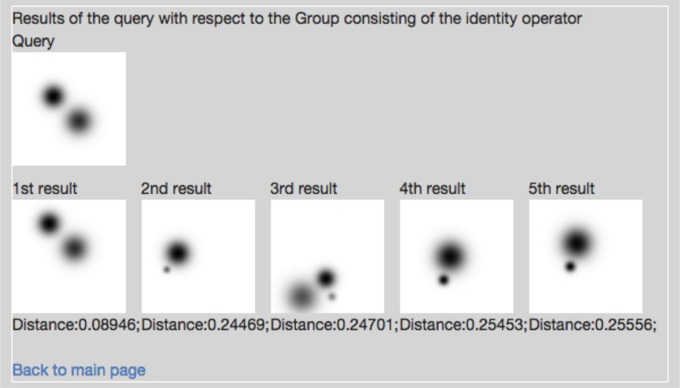
- $F(\varphi) = \varphi$.
- $F(\varphi) :=$ constant function taking each point to $\int_{\mathbb{R}^2} \varphi(x)\, dx$.
- $F(\varphi)$ defined by setting $F(\varphi)(x) := \int_{\mathbb{R}^2} \varphi(x-y) \cdot \beta(\|y\|_2)\, dy$ with $\beta : \mathbb{R} \to \mathbb{R}$ and $\int_{\mathbb{R}^2} |\beta(\|y\|_2)|\, dy \leq 1$
  (we have used six operators of this kind).
- Let us choose a continuous function $\chi(y_1, y_2)$ that is 1-Lipschitz in the variable y_1, for every y_2. Let us define $F(\varphi)$ by setting
  $F(\varphi)(x_1, x_2) := \chi(\varphi(x_1, x_2), x_2)$
  (we have used six operators of this kind).
- Let us choose a continuous function $\omega(y_1, y_2)$ that is 1-Lipschitz in the variable y_2, for every y_1. Let us define $F(\varphi)$ by setting
  $F(\varphi)(x_1, x_2) := \omega(x_1, \varphi(x_1, x_2))$
  (we have used one operator of this kind).

GIPHOD: Examples for the identical group



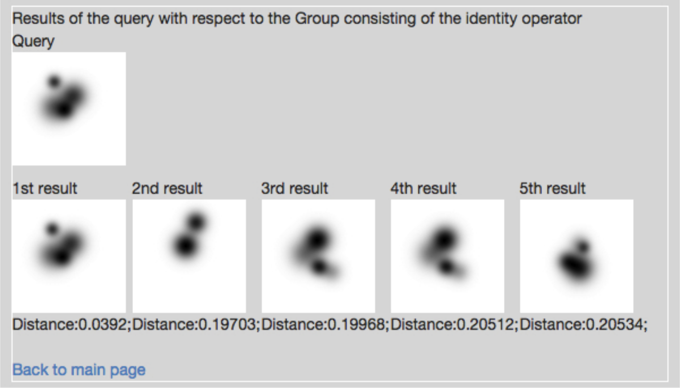
Mean distance: 0.6413; Standard deviation of distance: 0.1314; Number of GINOs: 15.

GIPHOD: Examples for the identical group



Mean distance: 0.6413; Standard deviation of distance: 0.1314; Number of GINOs: 15.

GIPHOD: Examples for the identical group



Mean distance: 0.6413; Standard deviation of distance: 0.1314; Number of GINOs: 15.

Conclusions

- The theory developed for GINOs from Φ to Φ can be easily extended to GINOs from Φ to $\Psi \neq \Phi$.
- An algebra of group invariant non-expansive operators can be introduced to build new GINOs.
- The demonstrator GIPHOD can be used to test the approach to TDA that we have described.

GIPHOD WILL BE AVAILABLE IN JANUARY 2016.
The beta version of GIPHOD can be tested at
`http://giphod.capdnet.ii.uj.edu.pl`.
Thanks to everyone that will give suggestions for improvement (please send them to `grzegorz.jablonski@uj.edu.pl`)


THANKS
FOR YOUR ATTENTION!