The Metric Approach for Shape Matching

Alfonso Ros

June 5, 2014
# Table of contents

1. Introduction
2. Gromov-Hausdorff distance
3. Correspondences and distortion
4. $L^p$ Gromov-Hausdorff relaxation
5. Discretization
6. Sampling
7. Properties
8. Applications
The similarity between two objects can be expressed as a distance function on some abstract shape space.

This distance needs the isometry invariance property to be insensitive to isometric deformations.
A **Metric space** is a ordered pair \((M, d)\) where \(M\) is a set and \(d\) is a metric on \(M\), i.e. a function \(d : M \times M \to \mathbb{R}\) such that for every \(x, y, z \in M\), the following holds:

1. \(d(x, y) \geq 0\) (non-negative)
2. \(d(x, y) = 0 \iff x = y\) (identity of indistinguishables)
3. \(d(x, y) = d(y, x)\) (symmetry)
4. \(d(x, z) \leq d(x, y) + d(y, z)\) (triangle inequality)
Isometry and $\varepsilon$-Isometry

**Definition**

Let $(X, d_X)$ and $(Y, d_Y)$ be two metric spaces. A map $f : X \to Y$ is called an **isometry** if it is

- **distance preserving**: $d_X(x, x') = d_Y(f(x), f(x'))$
- **bijective**

**Definition**

Let $(X, d_X)$ and $(Y, d_Y)$ be two metric spaces. A map $f : X \to Y$ is said to be an **almost isometry** or **$\varepsilon$-isometry** if there exist some $\varepsilon \geq 0$ such that $|d_Y(f(x), f(x')) - d_X(x, x')| \leq \varepsilon$ for all $x, x' \in X$, and for every $y \in Y$ there exist $x \in X$ such that $d_Y(y, f(x)) \leq \varepsilon$.
Let $\mathbb{M}$ be the space of all shapes.

Isometries form an equivalence relation of shapes in the space $\mathbb{M}$.

The set of all equivalent shapes (isometric) to $X \in \mathbb{M}$ is called equivalent class of $X$.

Considering the equivalent classes as a single unit, $\mathbb{M}$ can be transformed into a new space, referred to as the quotient space $\mathbb{M}^*$.

**Definition**

An Equivalence relation is a relation between pairs of objects, which is reflexive, symmetric and transitive.
Quotient Space
Hausdorff distance

**Definition**

Let \((Z, d_Z)\) be a metric space and \(X, Y \subset Z\). The **Hausdorff distance** between \(X\) and \(Y\) is defined as

\[
d^Z_H(X, Y) := \max \left( \sup_{x \in X} \inf_{y \in Y} d_Z(x, y), \sup_{y \in Y} \inf_{x \in X} d_Z(x, y) \right)
\]  

(1)
It *embeds* the two metric spaces $X$ and $Y$ into a common ambient metric space $Z$ and then compare them with the **Hausdorff** distance.

\[
\begin{align*}
(X, d_X) & \overset{d_{GH}}{\iff} (Y, d_Y) \\
\downarrow f & \quad \downarrow g \\
(f(X), d_Z|_{X'}) & \overset{d_{\mathcal{H}}}{\iff} (g(Y), d_Z|_{Y'})
\end{align*}
\]
We require $d_{GH}(X, Y) < r$ for some $r > 0$ if and only if there exist a metric space $Z$ and subspaces $X', Y' \subset Z$ which are isometric to $X$ and $Y$, and such that $d_H(X', Y') < r$. 
The **Gromov-Hausdorff** distance between two metric spaces \((X, d_X)\) and \((Y, d_Y)\) is defined by

\[
d_{GH} := \inf_{f: X \rightarrow Z, \ g: Y \rightarrow Z} d_H^Z(f(X), g(Y))
\]  

Where \(f : X \rightarrow Z\) and \(g : Y \rightarrow Z\) are isometric embeddings into the metric space \((Z, d)\)

- \(d_{GH}(X, Y) = 0\) if and only if \(X\) and \(Y\) are isometric.
- \(d_{GH}\) is a metric on the space of isometry classes of compact metric spaces.
The original formulation of the GH distance is very abstract and not friendly in terms of computation. For this reason, we define the concept of correspondence to derive a more practical form of the GH distance.

**Definition**

For sets $A$ and $B$, a subset $R \subseteq A \times B$ is a **correspondence** (between $A$ and $B$) if and only if

1. $\forall a \in A$, there exist $b \in B$ s.t. $(a, b) \in R$
2. $\forall b \in B$, there exist $a \in A$ s.t. $(a, b) \in R$

Let $\mathcal{R}(A, B)$ denote the set of all possible correspondences between sets $A$ and $B$. 
Correspondences and distortion

Let \((X, d_X)\) and \((Y, d_Y)\) two metric spaces. For any \((x, y)\) and \((x', y')\) correspondences in \(R\), their distortion is given by a function 
\(\Gamma : X \times X \times Y \times Y \to \mathbb{R}^+\)

\[
\Gamma(x, x'; y, y') := |d_X(x, x') - d_Y(y, y')|
\]
Correspondences and distortion

Correspondence Distortion

In general, the **distortion** of the entire correspondence $R$, is given by

$$
\|\Gamma\|_{L^\infty(R \times R)} = \sup_{(x_1,y_1),(x_2,y_2) \in R} \Gamma(x_1, x_2; y_1, y_2)
$$

(4)

If $\|\Gamma\|_{L^\infty(R \times R)} = 0$ then the correspondence $R$ is associated to an **isometry**.
With these concepts we can obtain an equivalent definition for the **Gromov-Hausdorff distance** as follows

\[
    d_{GH}(X, Y) = \inf_{R \in \mathcal{R}} \| \Gamma \|_{L^\infty(R \times R)}
\]

Restricting to the case of finite metric spaces \(X = \{x_1, \ldots, x_n\}\) and \(Y = \{y_1, \ldots, y_n\}\), we have

\[
    d_{GH}(X, Y) = \frac{1}{2} \min_{R \in \mathcal{R}} \max_{(x_1, y_1), (x_2, y_2) \in R} \Gamma(x_1, x_2; y_1, y_2)
\]
Let the **binary function** $\mu^R : X \times Y \to \{0, 1\}$ represent a correspondence $R$ subject to the condition

$$\max_{x \in X} \mu^R(x, y) = \max_{y \in Y} \mu^R(x, y) = 1$$  \hspace{1cm} (7)

This basically means $\mu(x, y) = 1$ if $(x, y) \in R$ and $\mu(x, y) = 0$ if $(x, y) \notin R$. Introducing this binary function $\mu^R$ to equation in the previous slide, together with the condition 7 we have

$$d_{\mathcal{GH}}(X, Y) = \frac{1}{2} \min_{R \in \mathcal{R}} \max_{i,j,k,l} \Gamma(x_i, x_k; y_j, y_l) \mu^R(x_i, y_j) \mu^R(x_k, y_l)$$  \hspace{1cm} (8)
Equivalent Gromov-Hausdorff distance

\[ d_{\mathcal{H}}(X, Y) = \frac{1}{2} \min_{R \in \mathcal{R}} \max_{i,j,k,l} \Gamma(x_i, x_k; y_j, y_l) \mu^R(x_i, y_j) \mu^R(x_k, y_l) \]

- Problem: The \( L^\infty \) norm makes this distance sensitive to noise and outliers.
Equivalent Gromov-Hausdorff distance

\[ d_{\mathcal{GH}}(X, Y) = \frac{1}{2} \min_{R \in \mathcal{R}} \max_{i,j,k,l} \Gamma(x_i, x_k; y_j, y_l) \mu^R(x_i, y_j) \mu^R(x_k, y_l) \]

- Problem: The \( L^\infty \) norm makes this distance sensitive to noise and outliers.
- Possible solution: using \( L^p \) norm would make the distance more robust to this problem.
Equivalent Gromov-Hausdorff distance

\[ d_{GH}(X, Y) = \frac{1}{2} \min_{\mathbb{R}} \max_{i,j,k,l} \Gamma(x_i, x_k; y_j, y_l) \mu^R(x_i, y_j) \mu^R(x_k, y_l) \]

- Problem: The \( L^\infty \) norm makes this distance sensitive to noise and outliers.
- Possible solution: using \( L^p \) norm would make the distance more robust to this problem.
- Problem: The \( L^p \) norm makes the distance inconsistent to non-uniform sampling.
Switching to $L^p$ norm

Just to give an intuition why the $L^p$ norm is not consistent to non-uniform sampling. Let $p = 1$

\[
\int_{R \times R} \Gamma dr \approx \frac{1}{N} \sum_{X'} \Gamma = \| \Gamma \|_{L^1(R \times R)}
\]

(a) uniform sampling

\[
\int_{R \times R} \Gamma dr \neq \left( \frac{1}{N} \right) \sum_{X'} \Gamma \neq \| \Gamma \|_{L^1(R \times R)}
\]

(b) non-uniform sampling

The $\Gamma$ function needs to be *weighted* by the factor corresponding to the area between the points.
Relaxation of the binary constraint

First step is to relax the function $\mu^R$ to be continuous in the interval $[0, 1]$.

$$\max_{x \in X} \mu^R(x, y) = \max_{y \in Y} \mu^R(x, y) = 1$$

$$\int_A \int_Y \mu(x, y) \, dy \, dx = \int_A \, dx; \quad \int_B \int_X \mu(x, y) \, dx \, dy = \int_B \, dy$$
Relaxation of the binary constraint

To understand a little bit more this relaxed constraint, consider the area of $A$ or $B$ to be a **single element**. The previous equation would be equivalent to

\[
\int_Y \mu(x, y) dy = 1; \quad \int_X \mu(x, y) dx = 1
\]

This means, that the function $\mu$ defines a **probability distribution** over the possible correspondences for one point into the other set. This property is called **marginalization constraint**.
Now we have in the setting not only the metric spaces \((X, d_X)\) and \((Y, d_Y)\), but we get into a more restrictive setting where we consider the probability measures \(\mu_X\) and \(\mu_Y\).

Joined together, they form a triplet \((X, d_X, \mu_X)\) and \((Y, d_Y, \mu_Y)\). These are known as **metric measure spaces** (**mm-space** for short).

In this more restrictive setting, only the isometries that preserve the area on the surfaces are allowed (isomorphisms). The two mm-spaces are called **isomorphic**.
Relaxed Gromov-Hausdorff

Using this relaxed notation of correspondence a new family of distances can be described as

$$d_{GH}(X, Y) = \frac{1}{2} \inf_{\Gamma} \| \Gamma \|_{L^p(R \times R)}$$

Where $1 \leq p \leq \infty$, and

$$\| \Gamma \|_{L^p(R \times R)} = \left( \int_{(R \times R)^2} \Gamma^p(x, x'; y, y') \mu^R(x, y) \mu^R(x', y') dxdydx'dy' \right)^{1/p}$$

This are the so called Mass Transportation distances or in the Math community as Wasserstein-Kantorovich-Rubinstein distances or Earth Mover’s distance.
In the discrete setting we have two sums instead of the integrals.

\[ d_{G^H}(X, Y) = \frac{1}{2} \min_R \left( \sum_{i,j} \sum_{k,l} \Gamma_{ij kl}^p \mu_{ij} R_{kl} \mu_{kl} \right)^{1/p} \]

This representation has a direct association to the **Quadratic Assignment Problem** described in the next slide.
This relaxation of the $L^\infty$ norm, relates to a family of problems known as the **Quadratic Assignment Problem**. Which its formulation is as follows

$$L_p(\delta) := \min_{\delta \in \Delta} \sum_{ij} \sum_{kl} \Gamma^{(p)}_{ij,kl}\delta_{ij}\delta_{kl}$$

Where $\Delta$ is the set of matrices constrained to

1. $\delta_{ij} \in \{0, 1\}$ for all $i, j$
2. $\sum_i \delta_{ij} = 1$ for all $j$
3. $\sum_j \delta_{ij} = 1$ for all $i$

These conditions force $\delta$ to be a permutation matrix. The resulting problem is known to be **NP-hard**.
Discretization

Back to the discrete setting

$$d_{G\mathcal{H}}(X, Y) = \frac{1}{2} \min_R \sum_{i,j}^{n,m} \sum_{k,l}^{n,m} \Gamma_{ij, kl} \mu_{ij}^R \mu_{kl}^R$$

Then we can write this formulation of the GH distance in matrix form

$$\frac{1}{2} \min_{R \in [0,1]^{n \times n}} \text{vec}\{R\}^T C \text{vec}\{R\}$$

s.t. $R^T 1 = 1, R1 = 1$

Where $R \in \mathbb{R}^{n \times n}$ is a **doubly stochastic** matrix and $C \in \mathbb{R}^{n^2 \times n^2}$ is the cost matrix.
If $R$ is not square, meaning the two surfaces have different number of points, it is not possible to fulfill the *doubly stochastic* property.

**Proof.**

Let $R \in \mathbb{R}^{n \times m}$ with $n \neq m$ s.t. $R^T \mathbf{1} = \mathbf{1}$ and $R \mathbf{1} = \mathbf{1}$

\[
\sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} = \sum_{j=1}^{m} \sum_{i=1}^{n} r_{ij}
\]

\[
\sum_{i=1}^{n} 1 = \sum_{j=1}^{m} 1
\]

$n = m$
As $R$ can no longer be *doubly stochastic* one of the constraints should be dropped. Let's say w.l.o.g. $n > m$ and we drop the constraint $R^T 1 = 1$, the resulting matrix is **row-stochastic** and **column-substochastic**.

The way to tackle this problem is adding an extra point $m^*$ represented by a new column on the matrix $R$. This point will “adopt” those without a correspondence.
Determining the cost to create a correspondence with the point $m^*$ is a very difficult problem.
Sampling

Coverings and sampling

Two criteria for defining a good sampling

**Definition**

$X' = \{x_1, \cdots, x_N\}$ is an **r-covering** of $X$ if

$$\bigcup_{x \in X'} B_r(x) = X$$

Where $B_r(x)$ is a metric ball in $X$ with radius $r$.

**Definition**

$X'$ is **r-separated** if $d(x, x') > r$ for all $x, x' \in X'$. 

Alfonso Ros ()
The Metric Approach for Shape Matching
June 5, 2014 30 / 41
Farthest Point Sampling

The **Farthest Point Sampling** produces a sampling from the metric space \((X, d_x)\) given an initial point \(x_1\).

**FPS algorithm**

- \(X' \leftarrow \{x_1\}\)
- \(d(x) \leftarrow d_X(x, x')\)
- **while** \(r > r_0\) **do**
  1. Find the farthest point from \(X'\), \(x' = \arg \max_{x \in X} d(x)\)
  2. Update the set of selected samples: \(X' \leftarrow X' \cup \{x'\}\)
  3. Update the distance function \(d(x)\),
     \[
     d(x) \leftarrow \min_{x \in X} \{d(x), d_X(x, x')\}
     \]
     \[
     r \leftarrow \max_{x \in X} d(x)
     \]

The resulting \(X' = \{x_1, \cdots, x_N\} \subset X\) is \(r\)-covering and \(r\)-separated.
Farthest point sampling

\[ X \]

\[ x_1 \]
Farthest point sampling
Farthest point sampling
Farthest point sampling
Farthest point sampling
Farthest point sampling
Let \( \{x_1, \cdots, x_N\} \) be an \( r \)-covering of the compact metric space \( (X, d_X) \), then

\[
d_{\mathcal{H}}(X, \{x_1, \cdots, x_N\}) \leq r
\]

The sampling is close to the underlying shape.
Some important properties of the GH distance

- Let $X = \{x_i\}_{i=1}^m$ be an $r$-covering of $X$ and $Y = \{y_j\}_{j=1}^{m'}$ be an $r'$-covering of $Y$. Then

$$|d_{GH}(X, Y) - d_{GH}(X, Y)| \leq d_{GH}(X, X) + d_{GH}(Y, Y) \leq r + r'$$

$d_{GH}$ is consistent to sampling

- For compact metric spaces $(X, d_X)$ and $(Y, d_Y)$:

$$\frac{1}{2}|\text{diam}(X) - \text{diam}(Y)| \leq d_{GH}(X, Y) \leq \frac{1}{2} \max(\text{diam}(X), \text{diam}(Y))$$

Where $\text{diam}(X) := \max_{x, x' \in X} d_X(x, x')$ stands for the diameter of the metric space $(X, d_X)$. 

\[\text{diam}(X) := \max_{x, x' \in X} d_X(x, x') \text{ stands for the diameter of the metric space } (X, d_X).\]
Pros and cons

- **Cons:**
  - Slow convergence
  - Local optimum
  - Gradient descent parameters
  - Binarize the final solution

- **Pros:**
  - Easy to implement
  - Local optima are usually good enough in practice
Applications

- Database with 7 type of shapes (camel, cat, elephant, faces, heads, horse and lion).
- Each class contains several different poses.
- First sampling of 4000 points using Euclidean FPS. Second sampling of 50 points using FPS with an intrinsic distance between the 4000 initial points.
- Use the density inside Voronoi partitions as the measure to build \((X, d_X, \mu_X)\).