Discrete Object Recognition and Symmetry Detection

Prof. Dr. Mireille Boutin

Lead, Discrete Algebra and Geometry group
Department of Mathematics and Computer Science
Einhoven Institute of Technology
and
Purdue University, West Lafayette, Indiana, USA

4TU.NAMI Meeting
June 7, 2024.

Abstract

How do you determine if two objects have the same shape? How do you reconstruct a room from echoes? How do you find HAZMAT signs in a picture? How do you reconstruct 3D objects from a movie? These questions boil down to the problem of characterizing the orbits of the action of a Lie group on a manifold. In this talk, I will discuss how to use invariants to solve such problems. In particular, I will discuss how to recognize configurations of points up to a rigid motion and relabeling using the "bag of distances," the related problem of reconstructing the shape of a room consisting of planar walls from the echoes heard by four microphones held in a rigid configuration on a drone, and how to detect symmetries in an object on an image using a pyramid of moment invariants.

A Bit of Invariant Theory

Definition

A **group action** or **transformation group** is given by a group G, a smooth manifold M, and a smooth map $\Phi : G \times M \to M$, denoted by $\Phi(g, x) = g \cdot x$, which satisfies

$$\begin{aligned} e \cdot x &= x, \\ g \cdot (h \cdot x) &= (g * h) \cdot x, \end{aligned}$$

for all $x \in M$ and all $g \in G$.

A Bit of Invariant Theory

Example: Rotations in $\mathbb{R}^2$

- $G = SO(2)$, group of 2×2 orthogonal matrices with determinant 1
- $M = \mathbb{R}^2$
- $$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

A Bit of Invariant Theory

Given is the action of a group G on a manifold M.

Definition:
Let $x \in M$. The **orbit through** x is

$$\{g \cdot x | g \in G\}.$$

A Bit of Invariant Theory

Example:

When $SO(2)$ acts on the plane by rotation, then the orbits are either

1. circles around the origin (submanifolds of dimension one)
2. a single point at the origin (submanifold of dimension zero)

A Bit of Invariant Theory

Definition:
Given the action of a group G on a manifold M, an **invariant** is a real-valued function

$$I : M \to \mathbb{R}$$

such that

$$I(g \cdot x) = I(x), \text{ for all } x \in M \text{ and all } g \in G.$$

Example:

The Euclidean norm $|x|$ of a point $x \in \mathbb{R}^2$ is invariant under the action of $SO(2)$ on the plane by rotation.

A Bit of Invariant Theory

Theorem

Given is action of group G on a manifold M with dimension m. Assume all the orbits of action of G on M have same dimension s. Assume also that each point of M has an arbitrarily small neighborhood whose intersection with each orbit is connected. Then

- (*Invariant Generation*)There exist $m - s$ functionally independent invariants $I_1, \ldots, I_{m-s}$, called **fundamental invariants**, such that any other invariant can be expressed, locally, as a function of those fundamental invariants.
- (*Orbit Separation*) Two points $x_1, x_2 \in M$ are in the same orbit if and only if

$$I_1(x_1) = I_1(x_2), \ldots, I_{m-2}(x_1) = I_{m-2}(x_2).$$

The Moving Frame Method

- originally formulated by Elie Cartan in 1937[1]
- yields a fundamental set of separating invariants.
- builds on idea of Frenet-Serret frame for planar curves from 1851 and 1852[2][3]
- Reformulated as a systematic method by Fels and Olver in 1998[4].

[1] Cartan, "La théorie des groupes finis et continus et la géométrie différentielle traitée par la méthode du repère mobile".

[2] Frenet, "Sur les courbes a double courbure."

[3] Serret, "Sur quelques formules relatives à la théorie des courbes à double courbure."

[4] Fels and Olver, "Moving coframes: I. A practical algorithm".

Recognizing the shape of a smooth planar curve

Given are two smooth curves in the plane.

We would like to know if one can rotate and translate the first curve so to obtain the second curve.

Group is $SE(2)$

- $SE(2) = SO(2) \ltimes \mathbb{R}^2$, called *special Euclidean group*

Recognizing the shape of a smooth planar curve

No invariant for "point transformations"

- When $SE(2)$ acts on $\mathbb{R}^2$, every orbit has dimension 2.
- dim (space)-dim(orbits)=2-2=0
- So 0 invariants

Recognizing the shape of a smooth planar curve

Trick:

increase dimension of M by **prolonging** action on "jet space" (bundle)

Example

- Suppose curve is parameterized as $y = u(x)$.
- coordinatize curve as a point $(x, u(x), u'(x), u''(x)) \in \mathbb{R}^4$
- Orbit dimension of prolonged action is 3
- dim (space)-dim(orbits)=4-3=1
- So one fundamental invariant: curvature

$$\kappa(x_1) = \frac{u''(x)}{(1 + u'(x)^2)^{\frac{3}{2}}}$$

differential invariant = invariant of the prolonged action of a group on the jet space

Recognizing the shape of a smooth planar curve

Theorem (Cartan)

Two submanifolds are (locally) equivalent if and only if they have the same functional relationships among all their differential invariants.

Definition

The **special Euclidean signature curve** associated with a plane curve is the planar curve parametrized by the curvature and its first derivative with respect to arc length.

Theorem[5][6]

Two analytic curves can be mapped to each other by a special Euclidean transformation $g \in SE(2)$ (rotation and translation) if and only if their special Euclidean signature curves are identical.

[5] Calabi et al., "Differential and numerically invariant signature curves applied to object recognition".

[6] Musso and Nicolodi, "Invariant signatures of closed planar curves".

Recognizing the shape of a discretized planar curve

What if we are only given a finite number of points on/near the curve?

Recognizing the shape of a discretized planar curve

Trick:

increase dimension of space by **prolonging** action on cartesian product $M \times M \times \ldots \times M$.

Example

- Suppose $G = SE(2)$ acts on $\mathbb{R}^2$ by rotation/translation:

$$g \cdot x = Rx + T, \quad R \in SO(2), T \in \mathbb{R}^2.$$

- We can prolong the action of G onto $\mathbb{R}^2 \times \mathbb{R}^2$ by acting on two points simultaneously

$$g \cdot (x_1, x_2) = (Rx_1 + T, Rx_2 + T), \quad R \in SO(2), T \in \mathbb{R}^2.$$

- One fundamental invariant: distance between two points

$$I_1(x_1, x_2) = \|x_2 - x_1\|$$

joint invariant=Invariant of the prolonged action of a group on cartesian product

Idea:

Replace differential invariant signature by joint invariant signature

Two Ways

1. Approximate differential invariants by joint invariants[7]
2. Parameterize entirely new signature curve using distances and areas[8]

[7] Calabi et al., "Differential and numerically invariant signature curves applied to object recognition"; Boutin, "Numerically invariant signature curves".

[8] Boutin, "Joint invariant signatures for curve recognition".

Recognizing the shape of a point configuration

What if the points given are not ordered?

Two approaches

- Method 1: moment invariants (Pascal Triangle)
- Method 2: distribution of distances

Recognizing the shape of a point configuration

Method 1: moment invariants (Pascal Triangle)

- Given are K points in $\mathbb{R}^2$ at positions $(x_k, y_k), k = 1, \ldots, N$.
- Can assign a weight $\omega_k \in \mathbb{R}_{\geq 0}$ to each point
- View the points are a (greyscale) digital image of an object

Recognizing the shape of a point configuration

Method 1: moment invariants (Pascal Triangle)

- Let $z_k = x_k + iy_k$
- Compute the moments $\mu_{m,r} = \sum_k \omega_k z_k^m (z_k^*)^r$
- The pascal triangle is this pyramid of moments:

$$
\begin{array}{c}
\mu_{0,0} \\
\mu_{0,1} \quad \mu_{1,0} \\
\mu_{0,2} \quad 2\mu_{1,1} \quad \mu_{2,0} \\
\mu_{0,3} \quad 3\mu_{1,2} \quad 3\mu_{2,1} \quad \mu_{3,0} \\
\mu_{0,4} \quad 4\mu_{1,3} \quad 6\mu_{2,2} \quad 4\mu_{3,1} \quad \mu_{4,0} \\
\vdots \\
\mu_{0,r} \quad \binom{r}{1}\mu_{1,r-1} \quad \cdots \quad \binom{r}{l}\mu_{l,r-l} \quad \cdots \quad \binom{r}{r-1}\mu_{r-1,1} \quad \mu_{r,0} \\
\vdots
\end{array}
$$

Method 1: moment invariants (Pascal Triangle)

Moving Frame method yields coordinates invariants under rotations/translations[9]

(assuming center of mass is at origin, $\mu_{0,1} = \mu_{1,0} = 0$.)

$$\mu_{0,0}$$

$$\frac{\mu_{0,1}}{e^{i\theta_0}} \qquad \frac{\mu_{1,0}}{e^{-i\theta_0}}$$

$$|\mu_{0,2}| \qquad 2\mu_{1,1} \qquad |\mu_{2,0}|$$

$$\frac{\mu_{0,3}}{e^{i3\theta_0}} \qquad 3\,\frac{\mu_{1,2}}{e^{i\theta_0}} \qquad 3\,\frac{\mu_{2,1}}{e^{-i\theta_0}} \qquad \frac{\mu_{3,0}}{e^{-i3\theta_0}}$$

$$\frac{\mu_{0,4}}{e^{i4\theta_0}} \qquad 4\,\frac{\mu_{1,3}}{e^{i2\theta_0}} \qquad \mu_{2,2} \qquad 4\,\frac{\mu_{3,1}}{e^{-i2\theta_0}} \qquad \frac{\mu_{4,0}}{e^{-i4\theta_0}}$$

$$\vdots$$

$$\frac{\mu_{0,r}}{e^{ir\theta_0}} \qquad \cdots \qquad \binom{r}{l}\frac{\mu_{l,r-l}}{e^{i(r-2l)\theta_0}} \qquad \cdots \qquad \frac{\mu_{r,0}}{e^{-ir\theta_0}}$$

where

$$\theta_0 = \begin{cases} \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) & \text{if } \sphericalangle(\vec{\mu}_{1,2}, \vec{e}_1) - \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) \in [-\frac{\pi}{2}, \frac{\pi}{2}], \\ \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) + \pi & \text{if } \sphericalangle(\vec{\mu}_{1,2}, \vec{e}_1) - \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) \in (\frac{\pi}{2}, \frac{3\pi}{2}], \\ \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) - \pi & \text{if } \sphericalangle(\vec{\mu}_{1,2}, \vec{e}_1) - \frac{1}{2}\sphericalangle(\vec{\mu}_{0,2}, \vec{e}_1) \in [-\frac{3\pi}{2}, -\frac{\pi}{2}). \end{cases}$$

[9] Boutin, Huang, et al., "The Pascal triangle of a discrete image: definition, properties and application to shape analysis".

Recognizing the shape of a point configuration

Method 1: moment invariants (Pascal Triangle)

The moments represent the shape of the object

w.l.o.g. assume center of mass is at origin ($\mu_{0,1} = \mu_{1,0} = 0$.)

- the quantity $0 \leq \frac{|\mu_{0,2}|}{\mu_{1,1}} \leq 1$ measures the elongation of the object; the object lies within a straight line if and only if $\frac{|\mu_{0,2}|}{\mu_{1,1}} = 0$
- The object is symmetric with respect to the x-axis if and only all the $\mu_{i,j}$'s are real.
- The object has a K-fold rotational symmetry if and only if $\mu_{i,j} = 0$ when $j - i$ is not a multiple of K.

Method 2: Bag of distances

- Compute the pairwise distances
- Forget the labeling

Can we reconstruct the points from the bag of distances?

Recognizing the shape of a point configuration

Reconstructing a point configuration from unlabeled distances is also a problem encountered in

- x-ray crystallography (Patterson 1935, Patterson 1944)
- mapping of restriction sites of DNA- *partial digest problem*- (Stefik 1978, Dix and Kieronska 1988, Gwangsoo 1988,...)
- material science (Jiao-Stillinger-Torquato 2010)
- ...

"Turnpike Problem" or "Partial Digest Problem": points lie in $\mathbb{R}$.
"Beltway Problem": the points lie on a circle.

Recognizing the shape of a point configuration

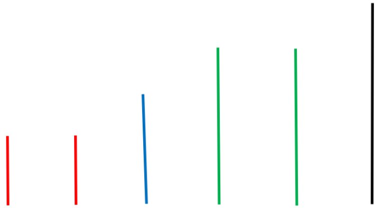
Question: Is the problem well-posed?

i.e., is the shape of a point-set *uniquely* determined by its unlabeled pairwise distances?

Example[10]

Is there a unique configuration of 4 points in the plane (up to a rigid motion) whose pairwise distances are

$$\{\sqrt{2}, \sqrt{2}, 2, \sqrt{10}, \sqrt{10}, 4\}?$$



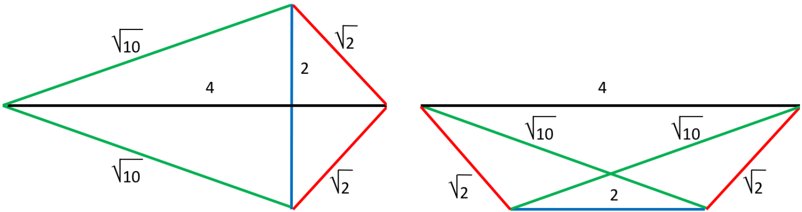
[10] Boutin and Kemper, "On reconstructing n-point configurations from the distribution of distances or areas".

Recognizing the shape of a point configuration

Question: Is the problem well-posed?

No.

Counterexample[11]



Two point-sets with the same pairwise distances

[11] Boutin and Kemper, "On reconstructing n-point configurations from the distribution of distances or areas".

Recognizing the shape of a point configuration

Question: Is the problem well-posed?

For Turnpike Problem ($D = 1$):

- Picard (1939): Proof of uniqueness when no repeated distances.
- Bloom (1977): 6-point counterexample.

Theorem (B.-Kemper)

Let $n \in \mathbb{N}$ with $0 < n \leq 3$ or $n \geq m + 2$
There exists a non-zero polynomial in mn variables such that every n-point configuration $p_1, \ldots, p_n \in \mathbb{R}^m$ with $f(p_1, \ldots, p_n) \neq 0$ is uniquely determined, up to a rigid motion, by the multiset of its unlabeled pairwise distances.

Corollary

- The set of exceptional point configurations has measure zero.
- Fast comparison algorithm that is accurate with probability 1.

Extension to other cases

- Noisy point sets: use probability density function of distance between two points drawn at random (i.i.d.) as a signature[12].
- Weighted graphs: use multiset of weights (unlabeled) and multiset of sum of pairs of adjacent weights (unlabeled) as a signature[13].

[12] Santos-Villalobos and Boutin, "Computationally efficient method to compare the shape of planar Gaussian mixtures from point samples".

[13] Boutin and Kemper, "Lossless representation of graphs using distributions".

Reconstructing a wall arrangement from echoes

Given is a *room*:

- arrangement of planar *"walls"* (ceilings, floors,. . .);
- not necessarily convex;
- not necessarily closed;
- position and number of walls is unknown.

Reconstructing a wall arrangement from echoes

We have 4 microphones:

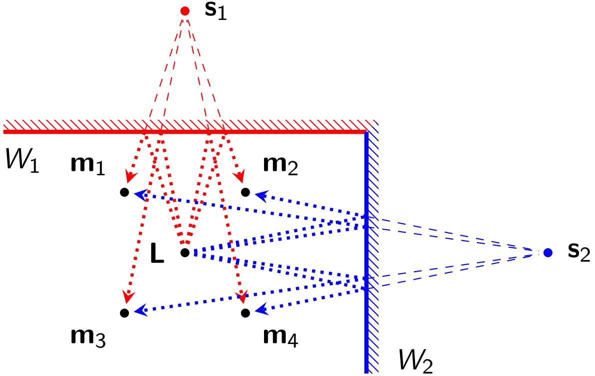
- known positions $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4 \in \mathbb{R}^3$;
- placed on a drone (allows rigid transformation).

An omnidirectional speaker emits a short pulse

- high-frequency so ray acoustics approximation holds;
- 1^{st} order echoes: pulses heard after they bounce off the walls;
- other echoes discarded
- First order echoes give (unlabeled) distances to walls.

Reconstructing a wall arrangement from echoes

Represent each wall W by a mirror point $s \in \mathbb{R}^3$



Virtually, sound comes from mirror points $\mathbf{s}_1$ and $\mathbf{s}_2$.

Reconstructing a wall arrangement from echoes

If room has only one wall $\mathbf{s}$

- Time between pulse emitted and first order echoes gives microphone-wall distance

$$d_i = \|m_i - \mathbf{s}\|^2.$$

- Reconstruct $\mathbf{s}$ from $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$ and d_1, d_2, d_3, d_4.
- Reconstruct wall plane from $\mathbf{s}$ and $\mathbf{L}$.
- Reconstruct 4 points on wall by intersecting wall plane with line from $\mathbf{s}$ to $\mathbf{m}_i$, $i = 1, \ldots, 4$.

If room has several walls: must sort (i.e., label) the echoes

- Determining whether four distances $(d_1, d_2, d_3, d_4) \in \mathcal{D}$ correspond to one wall,

i.e. these exists $\mathbf{s}$ s.t. $\|\mathbf{m}_i - \mathbf{s}\|^2 = d_i$, for $i = 1, 2, 3, 4$.

Echo Sorting

Five Point Echo Sorting Criterion[14]:

Let $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4, \mathbf{m}_5 \in \mathbb{R}^3$.
Let $D_{i,j} = \|\mathbf{m}_i - \mathbf{m}_j\|^2$, $i, j = 1, 2, 3, 4, 5$ and $u_1, \ldots, u_5 \in \mathbb{R}$. Let

$$E := \begin{pmatrix} 0 & u_1 & \cdots & u_5 \\ u_1 & D_{1,1} & \cdots & D_{1,5} \\ \vdots & \vdots & & \vdots \\ u_4 & D_{4,1} & \cdots & D_{4,5} \\ u_5 & D_{5,1} & \cdots & D_{5,5} \end{pmatrix} \quad \text{and} \quad g_E(u_1, u_2, \ldots, u_5) := \det(E).$$

Then $g_E(d_1, d_2, d_3, d_4, d_5) = 0$ when d_1, d_2, d_3, d_4, d_5 correspond to same wall $\mathbf{s}$.

[14] Dokmanić et al., "Acoustic echoes reveal room shape".

Echo Sorting

Why?

Because if $x_1, \ldots, x_k \in \mathbb{R}^n$, then the Euclidean Distance Matrix

$$\begin{pmatrix} \|x_1 - x_1\|^2 & \|x_1 - x_2\|^2 & \cdots & \|x_1 - x_k\|^2 \\ \|x_2 - x_1\|^2 & \|x_2 - x_2\|^2 & \cdots & \|x_2 - x_k\|^2 \\ \|x_3 - x_1\|^2 & \|x_3 - x_2\|^2 & \cdots & \|x_3 - x_k\|^2 \\ \vdots & \vdots & & \vdots \\ \|x_k - x_1\|^2 & \|x_k - x_2\|^2 & \cdots & \|x_k - x_k\|^2 \end{pmatrix} \in \mathbb{R}^{k \times k}$$

has rank at most $n + 2$.

Echo Sorting

Four Microphone Echo Sorting Criterion[15]:

Let $D_{i,j} := \|\mathbf{m}_i - \mathbf{m}_j\|^2$ and let $u_1 \dots u_4 \in \mathbb{R}$. Let

$$D := \begin{pmatrix} 0 & u_1 & \cdots & u_4 & 1 \\ u_1 & D_{1,1} & \cdots & D_{1,4} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ u_4 & D_{4,1} & \cdots & D_{4,4} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \quad \text{and} \quad f_M(u_1 \dots u_4) := \det(D). \tag{1}$$

Then $f_M(d_1 \dots d_4) = 0$ when d_1, d_2, d_3, d_4 correspond to the same wall $\mathbf{s}$.

[15] Boutin and Kemper, "A drone can hear the shape of a room".

Echo Sorting

Why?

- Set $\mathbf{m}_0 = \mathbf{s}$.
- Then

$$\det D = \det \begin{pmatrix} D_{0,0} & D_{0,1} & \cdots & D_{0,4} & 1 \\ D_{1,0} & D_{1,1} & \cdots & D_{1,4} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ D_{4,0} & D_{4,1} & \cdots & D_{4,4} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix}$$

  is the Cayley-Menger determinant of the 5-simplex $\mathbf{m}_0, \mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$.

The Algorithm

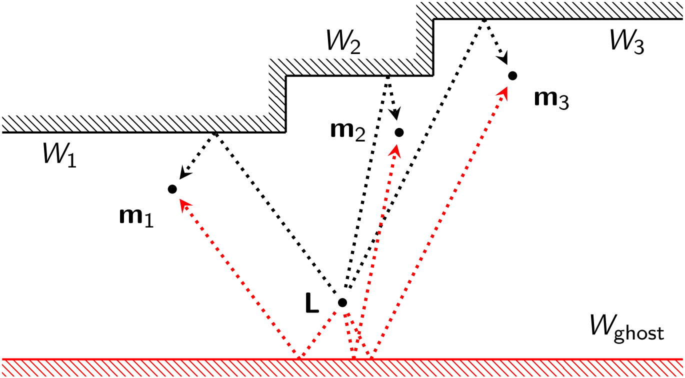
1. For $i = 1, \dots, 4$, collect the times of the first-order echoes recorded by the ith microphone in the set $\mathcal{T}_i$.
2. Set $\mathcal{D}_i := \{c^2(t - t_0)^2 \mid t \in \mathcal{T}_i\}$ $(i = 1, \dots, 4)$, where c is the speed of sound and t_0 is the time of sound emission.
3. FOR $(d_1, d_2, d_3, d_4) \in \mathcal{D}_1 \times \mathcal{D}_2 \times \mathcal{D}_3 \times \mathcal{D}_4$ DO
   - 3.4 IF $f_M(d_1, \dots, d_4) = 0$ THEN
     - 3.4.5 Compute the mirror point $\mathbf{s}$ from $(d_1, \dots, d_4)$.
     - 3.4.6 Compute four non-collinear points on the wall with mirror point $\mathbf{s}$ and, if desired, a normal vector.
     - 3.4.7 OUTPUT the data of this wall.

Wall Reconstruction Algorithm

Observe:

- Algorithm reconstructs all walls heard by 4 microphones.
- Algorithm could reconstruct walls that are not there (ghost walls).

Example of Ghost wall



Wall Reconstruction Algorithm

Theorem (B.-Kemper)[16]

The set of bad drone positions lies in a subspace of dimension ≤ 5 within the 6-dimensional space of possible drone positions.

$\Rightarrow$ If drone in generic position, our algorithm only reconstructs walls that are there.

So

- A drone in generic position can hear the shape of a room from echoes.

[16] Boutin and Kemper, "A drone can hear the shape of a room".

Concluding Remarks

- Invariant theory is well developed; has a rich history.
- New objects (discrete, noise, complex structure) still pose challenge.

Where to apply?

- Look for problem involving equations with extraneous parameters;
- See if extraneous parameters can be viewed as parameters of a group action ;
- Use invariants to rephrase the problem without the extraneous parameters.

Merci!