Invariance, Geometric Recoverability, and Angular Redundancy

Mathematical Supplement For Beyond Landmark Extraction: A Framework for Robust Geometric Feature Evaluation in Structured Image Classification

This document provides the mathematical definitions and proofs underlying the invariance properties, geometric recoverability, and angular redundancy of the nine landmark representations considered in the main manuscript.

Geometric equivalence and feature representations

Geometric invariance can be understood through the notion of equivalence classes of landmark configurations. Let $\mathcal{X}$ denote the space of admissible configurations $Q = \{q_j\}_{j=0}^{20}$, and let G be a group of geometric transformations acting on $\mathcal{X}$. Two configurations are considered equivalent if they differ only by a transformation in G, that is,

$$Q \sim_G Q' \quad \Longleftrightarrow \quad Q' = gQ \text{ for some } g \in G.$$

The equivalence class

$$[Q]_G = \{gQ : g \in G\}$$

therefore consists of all configurations related to Q through the transformations in G. In the present setting, we take G to be the group generated by translations and positive uniform scalings, since the absolute position and overall size of a hand are treated as nuisance variations. This perspective is motivated in part by classical Procrustes analysis [Gower(1975), Kendall(1984)], which compares landmark configurations after accounting for translation, scaling, and rotation. Unlike Procrustes alignment, however, our construction removes only translation and positive uniform scaling, deliberately preserving orientation relative to the image coordinate system. This distinction allows us to investigate geometric invariance without discarding potentially discriminative orientation information.This provides an abstract geometric description of the information we wish to preserve: configurations belonging to the same equivalence class should be indistinguishable with respect to these transformations. However, an equivalence class is an abstract mathematical object and does not, by itself, specify a numerical representation suitable for computation. To use this geometric description in a classification algorithm, we must construct explicit maps from landmark configurations to finite-dimensional feature spaces. Ideally, such a map $\phi : \mathcal{X} \to \mathbb{R}^p$ should assign the same feature vector to configurations within the same equivalence class, allowing the classifier to operate on representations that are insensitive to the chosen nuisance transformations. Importantly, different constructions may satisfy the same invariance requirements without preserving the same additional geometric information or making that information equally accessible to a learning algorithm. The nine representations introduced below provide different computational realizations of this geometric principle. To establish precisely which transformations each representation removes, and which geometric properties it retains, we first define invariance and the relevant transformation classes, before proving the corresponding properties of the feature maps.

1 Definitions and domain

1.1 Mathematical Definition of Representaion Spaces

For an image I_i of width W_i and height H_i, let $L(I_i) = \{(x_j, y_j, z_j)\}_{j=0}^{20}$ denote the MediaPipe output. Define the aspect-ratio-corrected planar configuration $Q = \{q_j\}_{j=0}^{20}$, where

$$q_j = \frac{(W_i x_j, H_i y_j)}{\max(W_i, H_i)}, \qquad j = 0, \ldots, 20.$$

The image index i is fixed throughout the proof; j indexes landmarks. Set

$$b_i = \begin{pmatrix} \min_j q_{jx} \\ \min_j q_{jy} \end{pmatrix}, \qquad R_k = \max_j q_{jk} - \min_j q_{jk}, \quad k \in \{x, y\},$$

$$\Lambda = \begin{pmatrix} R_x & 0 \\ 0 & R_y, \end{pmatrix}, \qquad \bar{q} = \frac{1}{21} \sum_{j=0}^{20} q_j,$$

$$D_i = \max_{0 \le a < b \le 20} \|q_a - q_b\|_2.$$

Note here that we are assuming the shared usability condition $\min\{R_x, R_y, D_i, \min_{(a,b)\in\mathcal{E}_\theta} \|q_a - q_b\|_2\} > 10^{-6}$ defined in the paper **??**, together with finite inputs and feature maps. Here $\mathcal{E}_\theta$ is the set of landmark pairs forming the arms of the selected angles. These conditions are imposed on all nine maps, even when an individual map could be computed under weaker conditions. Define

$$T_j = q_j - b_i, \qquad A_j = \Lambda_i^{-1}(q_j - b_i), \qquad Z_j = \frac{q_j - \bar{q}}{D_i}, \tag{1}$$

$$d_{ab} = \|q_a - q_b\|_2, \qquad \delta_{ab} = \frac{d_{ab}}{D_i}. \tag{2}$$

Distances are listed for all $0 \le a < b \le 20$, giving $\binom{21}{2} = 210$ entries. For each selected ordered triple (a, b, c), define the unsigned angle at its middle landmark:

$$\theta_{abc} = \arccos \frac{(q_a - q_b)^\top (q_c - q_b)}{\|q_a - q_b\|_2 \, \|q_c - q_b\|_2} \in [0, \pi]. \tag{3}$$

The 15 triples are, for each $j \in \{1, 5, 9, 13, 17\}$, $(0, j, j+1)$, $(j, j+1, j+2)$, and $(j+1, j+2, j+3)$. The nine maps are

$$\phi_q(Q) = \vec{Q} \subset \mathbb{R}^{42}, \qquad \phi_T(Q) = \vec{T} \subset \mathbb{R}^{42}, \qquad \phi_A(Q) = \vec{A} \subset \mathbb{R}^{42}$$

$$\phi_Z(Q) = \vec{Z}, \subset \mathbb{R}^{42}, \qquad \phi_d(Q) = (d_{ab})_{a<b} \subset \mathbb{R}^{210}, \qquad \phi_\delta(Q) = (\delta_{ab})_{a<b} \subset \mathbb{R}^{210}$$

$$\phi_s(Q) = [\phi_A(Q), \phi_\delta(Q)] \subset \mathbb{R}^{252}, \quad \phi_f(Q) = [\phi_A(Q), \phi_\delta(Q), \phi_\theta(Q)] \subset \mathbb{R}^{267}$$

The two ablation-only maps $\phi_{f\setminus A} = (\phi_\delta, \phi_\theta)$ and $\phi_{f\setminus\delta} = (\phi_A, \phi_\theta)$ are not counted among the nine candidate maps.

For clarity in the reconstruction algebra below, we abbreviate $b = b_i$, $\Lambda = \Lambda_i$, and $D = D_i$. In that proof, indices i, j, k denote landmark indices locally, rather than the fixed image index.

1.2 Definition of Invariance

A geometric feature representation is said to be invariant under a transformation if applying that transformation to the underlying landmark configuration does not change the resulting feature vector. Formally, let $Q = \{q_j\}_{j=0}^{20}$ denote an admissible configuration of 21 labeled planar landmarks, and let $\phi : \mathcal{X} \to \mathbb{R}^p$ be a feature map, where $\mathcal{X}$ denotes the set of admissible configurations and p is the dimension of the resulting representation. A transformation g acts on Q by transforming each landmark, producing

$$gQ = \{g(q_j)\}_{j=0}^{20}.$$

We say that ϕ is invariant under g if

$$\phi(gQ) = \phi(Q)$$

for every $Q \in \mathcal{X}$ such that $gQ \in \mathcal{X}$. In other words, two configurations related by the transformation are assigned the same feature vector, even though their landmark coordinates may differ. This is no way a unique definition of invariance, it is simply the standard definition formulated to fit out realization conventions.
We know consider the following geometric transformations:

Translation. A translation displaces every landmark by the same vector $t = (t_x, t_y)^\top \in \mathbb{R}^2$, according to

$$q'_j = q_j + t.$$

The components t_x and t_y specify the horizontal and vertical displacements respectively. Translation changes the absolute location of the configuration in the image without changing the relative arrangement of its landmarks.

Positive uniform scaling. A positive uniform scaling changes the overall size of a configuration by multiplying both coordinates of every landmark by a common factor:

$$q'_j = cq_j, \qquad c \in \mathbb{R}_{>0}.$$

Values $c > 1$ enlarge the configuration, whereas $0 < c < 1$ reduce its size. Because the same factor is applied to both coordinate axes, the relative proportionality of the configuration remains undisturbed.

Rotation. A rotation changes the orientation of a configuration in the image plane. It is represented by

$$q'_j = Uq_j,$$

where U belongs to the special orthogonal group

$$SO(2) = \left\{U \in \mathbb{R}^{2\times 2} : U^\top U = I,\ \det(U) = 1\right\}.$$

U being unimodular here restricts the space to orthogonal transformations that preserve orientation. In two dimensions, every such matrix can be expressed as

$$U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

where α denotes the angle of rotation about the origin. Unlike translation or uniform scaling, rotation changes the orientation of the landmarks relative to the image coordinate axes.

Reflection. A reflection reverses the orientation of a configuration across a line through the origin. It is represented by

$$q'_j = Uq_j,$$

where U is an orthogonal matrix satisfying

$$U^\top U = I, \qquad \det(U) = -1.$$

Such matrices belong to the orthogonal group

$$O(2) = \left\{ U \in \mathbb{R}^{2\times 2} : U^\top U = I \right\},$$

but not to $SO(2)$. The determinant distinguishes reflections from rotations: rotations preserve orientation, whereas reflections reverse it. For example, reflection across the horizontal axis is represented by

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Although rotations and reflections are both orthogonal transformations, they differ in their effect on the orientation of the configuration.

Positive coordinatewise scaling. Coordinatewise scaling allows the horizontal and vertical dimensions of a configuration to be rescaled independently. It is defined by

$$q'_j = Hq_j, \qquad H = \begin{pmatrix} a_x & 0 \\ 0 & a_y \end{pmatrix},$$

where $a_x, a_y > 0$. When $a_x = a_y$, this transformation reduces to positive uniform scaling. Otherwise, the two coordinate axes are rescaled by different factors, potentially changing the proportions and internal geometry of the configuration. This distinction is relevant to the axis-normalized representation ϕ_A, which treats the coordinate ranges independently.
Finally, invariance is a property of a feature map over a class of configurations, rather than a property of an individual example. A representation is described as *not generally invariant* under a transformation when $\exists$ atleast one admissible configuration Q and an admissible transformation g for which

$$\phi(gQ) \neq \phi(Q).$$

This does not imply that every transformed configuration must produce a different feature vector. The following section establishes the invariance properties of the nine representations by examining how their defining quantities behave under the transformations introduced above.

2 Proof of the claimed invariances

Throughout, let $Q = \{q_j\}_{j=0}^{20}$ be an admissible landmark configuration. We write

$$b_i = \begin{pmatrix} \min_j q_{j,x} \\ \min_j q_{j,y} \end{pmatrix}, \qquad \bar{q} = \frac{1}{21}\sum_{j=0}^{20} q_j,$$

and define

$$\Lambda_i = \operatorname{diag}(R_x, R_y), \qquad D_i = \max_{j,k} \|q_j - q_k\|_2.$$

Here R_x and R_y denote the coordinate ranges. All calculations are understood on the domain where the corresponding representations are defined.

2.1 Translation

Let $t \in \mathbb{R}^2$ and consider the transformation

$$q'_j = q_j + t.$$

The coordinatewise minimum transforms as

$$b'_i = \min_j (q_j + t) = \min_j q_j + t = b_i + t,$$

where the minimum is taken coordinatewise. Similarly, the centroid satisfies

$$\bar{q}' = \frac{1}{21} \sum_{j=0}^{20} q'_j = \frac{1}{21} \sum_{j=0}^{20} (q_j + t) = \bar{q} + t. \tag{4}$$

For each coordinate $k \in \{x, y\}$, the range is

$$\begin{aligned} R'_k &= \max_j (q_{j,k} + t_k) - \min_j (q_{j,k} + t_k) \\ &= (\max_j q_{j,k} + t_k) - (\min_j q_{j,k} + t_k) \\ &= R_k. \end{aligned} \tag{5}$$

Consequently,

$$\Lambda'_i = \Lambda_i.$$

The maximum pairwise distance is also unchanged:

$$\begin{aligned} D'_i &= \max_{j,k} \|q'_j - q'_k\|_2 \\ &= \max_{j,k} \|(q_j + t) - (q_k + t)\|_2 \\ &= \max_{j,k} \|q_j - q_k\|_2 \\ &= D_i. \end{aligned} \tag{6}$$

We now examine the coordinate representations. For translated coordinates,

$$\begin{aligned} T'_j &= q'_j - b'_i \\ &= (q_j + t) - (b_i + t) \\ &= q_j - b_i = T_j. \end{aligned} \tag{7}$$

For axis-normalized coordinates,

$$\begin{aligned} A'_j &= (\Lambda'_i)^{-1}(q'_j - b'_i) \\ &= \Lambda_i^{-1}[(q_j + t) - (b_i + t)] \\ &= \Lambda_i^{-1}(q_j - b_i) \\ &= A_j. \end{aligned} \tag{8}$$

For centered isotropic coordinates,

$$\begin{aligned} Z'_j &= \frac{q'_j - \bar{q}'}{D'_i} \\ &= \frac{(q_j + t) - (\bar{q} + t)}{D_i} \\ &= \frac{q_j - \bar{q}}{D_i} \\ &= Z_j. \end{aligned} \tag{9}$$

For pairwise distances,

$$\begin{aligned} d'_{jk} &= \|q'_j - q'_k\|_2 \\ &= \|(q_j + t) - (q_k + t)\|_2 \\ &= \|q_j - q_k\|_2 \\ &= d_{jk}. \end{aligned} \tag{10}$$

It follows that normalized distances satisfy

$$\delta'_{jk} = \frac{d'_{jk}}{D'_i} = \frac{d_{jk}}{D_i} = \delta_{jk}. \tag{11}$$

Finally, consider an angle formed by landmarks q_a, q_b, q_c. Define the angle arms

$$v = q_a - q_b, \qquad w = q_c - q_b.$$

Under translation,

$$v' = (q_a + t) - (q_b + t) = v, \tag{12}$$

$$w' = (q_c + t) - (q_b + t) = w. \tag{13}$$

Therefore,

$$\cos\theta'_{abc} = \frac{(v')^\top w'}{\|v'\|_2 \|w'\|_2} = \frac{v^\top w}{\|v\|_2 \|w\|_2} = \cos\theta_{abc}, \tag{14}$$

and hence

$$\theta'_{abc} = \theta_{abc}.$$

Since the constituent blocks are unchanged,

$$\phi_s(Q + t) = \mathrm{vec}(A', \delta') = \phi_s(Q), \tag{15}$$

$$\phi_f(Q + t) = \mathrm{vec}(A', \delta', \theta') = \phi_f(Q). \tag{16}$$

In contrast, the raw coordinate representation satisfies

$$\begin{aligned} \phi_q(Q + t) &= \mathrm{vec}(q_0 + t, \ldots, q_{20} + t) \\ &= \phi_q(Q) + \mathrm{vec}(t, \ldots, t), \end{aligned} \tag{17}$$

which differs from $\phi_q(Q)$ whenever $t \neq 0$.
Thus $\phi_T, \phi_A, \phi_Z, \phi_d, \phi_\delta, \phi_\theta, \phi_s, \phi_f$ are translation invariant, whereas ϕ_q is not.

2.2 Positive uniform scaling

Let $c > 0$ and consider

$$q'_j = cq_j.$$

Since multiplication by a positive scalar preserves coordinate order,

$$b'_i = \min_j(cq_j) = c \min_j q_j = cb_i. \tag{18}$$

Likewise,

$$\begin{aligned} \bar{q}' &= \frac{1}{21} \sum_{j=0}^{20} cq_j \\ &= c\bar{q}, \end{aligned} \tag{19}$$

and

$$\begin{aligned} R'_k &= \max_j(cq_{j,k}) - \min_j(cq_{j,k}) \\ &= c(\max_j q_{j,k} - \min_j q_{j,k}) \\ &= cR_k. \end{aligned} \tag{20}$$

Therefore,

$$\Lambda'_i = c\Lambda_i.$$

For the maximum pairwise distance,

$$\begin{aligned} D'_i &= \max_{j,k} \|cq_j - cq_k\|_2 \\ &= \max_{j,k} c\|q_j - q_k\|_2 \\ &= cD_i. \end{aligned} \tag{21}$$

The translated coordinates satisfy

$$\begin{aligned} T'_j &= q'_j - b'_i \\ &= cq_j - cb_i \\ &= c(q_j - b_i) \\ &= cT_j. \end{aligned} \tag{22}$$

Thus ϕ_T is not generally invariant.
For axis-normalized coordinates,

$$\begin{aligned} A'_j &= (\Lambda'_i)^{-1}(q'_j - b'_i) \\ &= (c\Lambda_i)^{-1}(cq_j - cb_i) \\ &= \frac{1}{c}\Lambda_i^{-1}c(q_j - b_i) \\ &= A_j. \end{aligned} \tag{23}$$

Similarly, for centered isotropic coordinates,

$$\begin{aligned} Z'_j &= \frac{q'_j - \bar{q}'}{D'_i} \\ &= \frac{cq_j - c\bar{q}}{cD_i} \\ &= \frac{q_j - \bar{q}}{D_i} \\ &= Z_j. \end{aligned} \tag{24}$$

Pairwise distances transform according to

$$\begin{aligned} d'_{jk} &= \|cq_j - cq_k\|_2 \\ &= c\|q_j - q_k\|_2 \\ &= cd_{jk}. \end{aligned} \tag{25}$$

However, normalized distances satisfy

$$\delta'_{jk} = \frac{d'_{jk}}{D'_i} = \frac{cd_{jk}}{cD_i} = \delta_{jk}. \tag{26}$$

For angles, the transformed arms are

$$v' = cv, \qquad w' = cw.$$

Hence,

$$\begin{aligned}\cos\theta'_{abc} &= \frac{(cv)^\top (cw)}{\|cv\|_2 \|cw\|_2} \\ &= \frac{c^2 v^\top w}{c^2 \|v\|_2 \|w\|_2} \\ &= \cos\theta_{abc}. \end{aligned} \tag{27}$$

Consequently,

$$\theta'_{abc} = \theta_{abc}.$$

Since both hybrid representations concatenate scale-invariant blocks,

$$\phi_s(cQ) = \phi_s(Q), \tag{28}$$

$$\phi_f(cQ) = \phi_f(Q). \tag{29}$$

In contrast,

$$\phi_q(cQ) = c\phi_q(Q), \qquad \phi_d(cQ) = c\phi_d(Q).$$

Therefore $\phi_A, \phi_Z, \phi_\delta, \phi_\theta, \phi_s, \phi_f$ are invariant under positive uniform scaling, whereas ϕ_q, ϕ_T, ϕ_d are not generally invariant.

2.3 Rotation and reflection

Let $U \in \mathbb{R}^{2\times 2}$ be orthogonal, so that

$$U^\top U = I.$$

This includes rotations, for which $\det U = 1$, and reflections, for which $\det U = -1$. Consider the transformation

$$q'_j = Uq_j.$$

For any $v \in \mathbb{R}^2$,

$$\begin{aligned}\|Uv\|_2^2 &= (Uv)^\top (Uv) \\ &= v^\top U^\top U v \\ &= v^\top v \\ &= \|v\|_2^2. \end{aligned} \tag{30}$$

Thus

$$\|Uv\|_2 = \|v\|_2.$$

Likewise, for any $v, w \in \mathbb{R}^2$,

$$\begin{aligned}(Uv)^\top (Uw) &= v^\top U^\top U w \\ &= v^\top w. \end{aligned} \tag{31}$$

Consequently, pairwise distances satisfy

$$\begin{aligned}d'_{jk} &= \|Uq_j - Uq_k\|_2 \\ &= \|U(q_j - q_k)\|_2 \\ &= \|q_j - q_k\|_2 \\ &= d_{jk}. \end{aligned} \tag{32}$$

The maximum pairwise distance is therefore

$$\begin{aligned} D_i' &= \max_{j,k} \|U(q_j - q_k)\|_2 \\ &= \max_{j,k} \|q_j - q_k\|_2 \\ &= D_i. \end{aligned} \tag{33}$$

Hence,

$$\delta_{jk}' = \frac{d_{jk}'}{D_i'} = \frac{d_{jk}}{D_i} = \delta_{jk}. \tag{34}$$

For angles, the transformed arms satisfy

$$v' = Uv, \qquad w' = Uw.$$

Therefore,

$$\begin{aligned} \cos\theta_{abc}' &= \frac{(Uv)^\top (Uw)}{\|Uv\|_2 \|Uw\|_2} \\ &= \frac{v^\top U^\top U w}{\|v\|_2 \|w\|_2} \\ &= \frac{v^\top w}{\|v\|_2 \|w\|_2} \\ &= \cos\theta_{abc}. \end{aligned} \tag{35}$$

Thus,

$$\theta_{abc}' = \theta_{abc}.$$

It follows that

$$\phi_d(UQ) = \phi_d(Q), \qquad \phi_\delta(UQ) = \phi_\delta(Q), \qquad \phi_\theta(UQ) = \phi_\theta(Q).$$

The coordinate representations behave differently. For raw coordinates,

$$q_j' = Uq_j,$$

which is not generally equal to q_j.
The centroid transforms as

$$\begin{aligned} \bar{q}' &= \frac{1}{21} \sum_{j=0}^{20} Uq_j \\ &= U\bar{q}. \end{aligned} \tag{36}$$

Consequently, centered isotropic coordinates satisfy

$$\begin{aligned} Z_j' &= \frac{Uq_j - U\bar{q}}{D_i} \\ &= U \frac{q_j - \bar{q}}{D_i} \\ &= UZ_j. \end{aligned} \tag{37}$$

In general,

$$UZ_j \neq Z_j.$$

Thus ϕ_Z is orthogonally equivariant, but not generally invariant.
For translated and axis-normalized coordinates,

$$T'_j = Uq_j - b'_i, \tag{38}$$
$$A'_j = (\Lambda'_i)^{-1}(Uq_j - b'_i). \tag{39}$$

Because coordinatewise minima and ranges do not generally commute with orthogonal transformations,

$$b'_i \neq Ub_i, \qquad \Lambda'_i \neq \Lambda_i$$

in general. Hence neither ϕ_T nor ϕ_A is universally invariant.
Finally, since the hybrid representations contain the axis-normalized coordinate block,

$$\phi_s(UQ) = \mathrm{vec}(A', \delta), \tag{40}$$
$$\phi_f(UQ) = \mathrm{vec}(A', \delta, \theta). \tag{41}$$

Because $A' \neq A$ in general,

$$\phi_s(UQ) \neq \phi_s(Q), \qquad \phi_f(UQ) \neq \phi_f(Q).$$

Therefore only $\phi_d, \phi_\delta, \phi_\theta$ are universally invariant under rotations and reflections among the nine representations.

2.4 Positive axis scaling

Let

$$H = \begin{pmatrix} a_x & 0 \\ 0 & a_y \end{pmatrix}, \qquad a_x, a_y > 0,$$

and consider the transformation

$$q'_j = Hq_j.$$

Since both scaling factors are positive, coordinatewise minima transform as

$$\begin{aligned} b'_i &= \begin{pmatrix} \min_j(a_x q_{j,x}) \\ \min_j(a_y q_{j,y}) \end{pmatrix} \\ &= \begin{pmatrix} a_x \min_j q_{j,x} \\ a_y \min_j q_{j,y} \end{pmatrix} \\ &= Hb_i. \end{aligned} \tag{42}$$

Similarly, the coordinate ranges satisfy

$$\begin{aligned} R'_x &= \max_j(a_x q_{j,x}) - \min_j(a_x q_{j,x}) \\ &= a_x R_x, \end{aligned} \tag{43}$$
$$\begin{aligned} R'_y &= \max_j(a_y q_{j,y}) - \min_j(a_y q_{j,y}) \\ &= a_y R_y. \end{aligned} \tag{44}$$

Consequently,

$$\begin{aligned} \Lambda'_i &= \begin{pmatrix} a_x R_x & 0 \\ 0 & a_y R_y \end{pmatrix} \\ &= H\Lambda_i. \end{aligned} \tag{45}$$

The axis-normalized coordinates therefore satisfy

$$\begin{aligned} A'_j &= (\Lambda'_i)^{-1}(q'_j - b'_i) \\ &= (H\Lambda_i)^{-1}(Hq_j - Hb_i) \\ &= \Lambda_i^{-1}H^{-1}H(q_j - b_i) \\ &= \Lambda_i^{-1}(q_j - b_i) \\ &= A_j. \end{aligned} \tag{46}$$

Thus,

$$\phi_A(HQ) = \phi_A(Q).$$

In contrast, pairwise distances transform as

$$\begin{aligned} (d'_{jk})^2 &= \|H(q_j - q_k)\|_2^2 \\ &= (q_j - q_k)^\top H^\top H(q_j - q_k) \\ &= a_x^2(q_{j,x} - q_{k,x})^2 + a_y^2(q_{j,y} - q_{k,y})^2. \end{aligned} \tag{47}$$

Unless $a_x = a_y$, this is not generally a common scalar multiple of d_{jk}^2. Consequently,

$$\frac{d'_{jk}}{D'_i} \neq \frac{d_{jk}}{D_i}$$

in general, so normalized distances are not invariant to arbitrary anisotropic scaling.
Similarly, for angle arms v, w,

$$\begin{aligned} \cos\theta'_{abc} &= \frac{(Hv)^\top(Hw)}{\|Hv\|_2\|Hw\|_2} \\ &= \frac{v^\top H^\top Hw}{\sqrt{v^\top H^\top Hv}\sqrt{w^\top H^\top Hw}}. \end{aligned} \tag{48}$$

Since

$$H^\top H = \begin{pmatrix} a_x^2 & 0 \\ 0 & a_y^2 \end{pmatrix},$$

this expression is not generally equal to

$$\frac{v^\top w}{\|v\|_2\|w\|_2}.$$

Hence ϕ_θ is not invariant to arbitrary positive separate axis scaling.
The centered isotropic coordinates satisfy

$$\bar{q}' = H\bar{q}, \tag{49}$$

$$Z'_j = \frac{H(q_j - \bar{q})}{D'_i}, \tag{50}$$

which is not generally equal to Z_j.
Likewise,

$$\begin{aligned} T'_j &= Hq_j - Hb_i \\ &= HT_j, \end{aligned} \tag{51}$$

so translated coordinates are not generally invariant. Raw coordinates are also not invariant, since $q'_j = Hq_j$.

Finally, although $A'_j = A_j$, the hybrid maps contain additional noninvariant blocks:

$$\phi_s(HQ) = \mathrm{vec}(A, \delta'), \tag{52}$$
$$\phi_f(HQ) = \mathrm{vec}(A, \delta', \theta'). \tag{53}$$

Since $\delta' \neq \delta$ and $\theta' \neq \theta$ in general,

$$\phi_s(HQ) \neq \phi_s(Q), \qquad \phi_f(HQ) \neq \phi_f(Q).$$

Therefore, among the nine representations, only ϕ_A is universally invariant to arbitrary positive separate axis scaling.
“‘latex

3 Proposition 1: Equivalent geometric encodings

Proposition 1 (Recoverability of the simple hybrid)**.** *Let $Q = \{q_j\}_{j=0}^{20}$ be a noncollinear, usable configuration as defined above. Define A, δ, Z as above, and let*

$$\lambda = \begin{pmatrix} \frac{R_x^2}{D^2} & \frac{R_y^2}{D^2} \end{pmatrix}^{\mathrm{T}}$$

denote the squared coordinate ranges relative to the squared configuration diameter. For each pair $i < j$, define

$$H_{(ij),:} = \big((A_{ix} - A_{jx})^2,\ (A_{iy} - A_{jy})^2\big), \qquad s_{ij} = \delta_{ij}^2. \tag{54}$$

Then $H\lambda = s$, and H has rank two. Consequently, λ is uniquely determined by A and δ, and the centered, diameter-normalized coordinates can be recovered as

$$Z_{ik} = \sqrt{\lambda_k}(A_{ik} - \bar{A}_k), \qquad k \in \{x, y\}. \tag{55}$$

Conversely, Z uniquely determines both A and δ. Thus the simple hybrid representation $\phi_s = (\vec{A}, \delta)$ and the centered, diameter-normalized representation $\phi_Z = \vec{Z}$ uniquely determine one another. Both encode the same labeled configuration up to translation and positive uniform scaling, without quotienting out rotation or reflection.

Proof. We first show that the simple hybrid representation (A, δ) uniquely determines Z.
Recall that A is obtained by normalizing each coordinate independently using its minimum and range. Writing $b_k = \min_j q_{jk}$ and $R_k = \max_j q_{jk} - \min_j q_{jk}$, we have

$$A_{ik} = \frac{q_{ik} - b_k}{R_k}, \qquad k \in \{x, y\}.$$

Since the configuration is usable, both coordinate ranges are positive. We can therefore rearrange this expression to obtain

$$q_{ik} = b_k + R_k A_{ik}.$$

This shows that A preserves the relative positions of the landmarks along each coordinate axis. However, because the horizontal and vertical coordinates are normalized separately, A does not directly retain the relative sizes of the two coordinate ranges. In particular, two configurations may have the same A while differing by independent positive scalings of the horizontal and vertical axes.
To recover Z, we must therefore determine the two coordinate ranges relative to the configuration diameter D. We show that the normalized pairwise distances δ provide precisely this information.

For any pair of landmarks $i < j$, the squared Euclidean distance is

$$\begin{aligned} d_{ij}^2 &= (q_{ix} - q_{jx})^2 + (q_{iy} - q_{jy})^2 \\ &= R_x^2(A_{ix} - A_{jx})^2 + R_y^2(A_{iy} - A_{jy})^2. \end{aligned}$$

Dividing both sides by D^2, and using $\delta_{ij} = \frac{d_{ij}}{D}$, gives

$$\delta_{ij}^2 = \frac{R_x^2}{D^2}(A_{ix} - A_{jx})^2 + \frac{R_y^2}{D^2}(A_{iy} - A_{jy})^2.$$

By the definition of λ, this becomes

$$\delta_{ij}^2 = \lambda_x(A_{ix} - A_{jx})^2 + \lambda_y(A_{iy} - A_{jy})^2. \tag{56}$$

The quantities A_{ik} and δ_{ij} are known from the simple hybrid representation. Thus each landmark pair provides a linear equation in the two unknown squared relative scales λ_x and λ_y. Collecting these equations over all pairs $i < j$ gives

$$H\lambda = s.$$

We now show that these equations uniquely determine λ. Since the original configuration is noncollinear and

$$\operatorname{rank}(H) = 2 \iff Q \text{ is noncollinear}$$

provided $R_x, R_y > 0$;$\exists$ three landmarks i, j, k that are not collinear. Furthermore The coordinate-wise normalization defining A consists of a translation followed by an invertible diagonal scaling, so it preserves noncollinearity. Hence the vectors

$$u = A_j - A_i, \qquad w = A_k - A_i$$

are linearly independent.
Suppose $Hv = 0$ for some $v = (v_x, v_y)^\top$, and define $B = \operatorname{diag}(v_x, v_y)$. By the definition of H, the equation $Hv = 0$ means that

$$v_x(A_{ax} - A_{bx})^2 + v_y(A_{ay} - A_{by})^2 = 0$$

for every landmark pair $a < b$.
In particular, applying this identity to the three pairs formed by i, j, k gives

$$u^\top Bu = 0, \qquad w^\top Bw = 0, \qquad (u - w)^\top B(u - w) = 0.$$

Expanding the third identity yields

$$u^\top Bu - 2u^\top Bw + w^\top Bw = 0,$$

where we have used the symmetry of B. Since the first and last terms are zero, we obtain

$$u^\top Bw = 0.$$

Consequently, the bilinear form associated with B vanishes on every combination of the two basis vectors u and w. Equivalently,

$$\begin{pmatrix} u^\top \\ w^\top \end{pmatrix} B \begin{pmatrix} u & w \end{pmatrix} = 0.$$

Let $P = [u\ w]$. Since u and w are linearly independent, P is invertible. The preceding identity is therefor $P^\top BP = 0$, and multiplication by $P^{-\top}$ and P^{-1} gives $B = 0$. It follows that $v_x = v_y = 0$, so the kernel of H is trivial and

$$\operatorname{rank}(H) = 2.$$

Thus the pairwise distance equations have at most one solution for the two squared relative scales. A solution exists because the original configuration satisfies these equations with

$$\lambda_x = \frac{R_x^2}{D^2}, \qquad \lambda_y = \frac{R_y^2}{D^2}.$$

Therefore this solution is unique. Moreover, since $R_x, R_y, D > 0$, both components of λ are strictly positive, and their positive square roots uniquely recover

$$\frac{R_x}{D} = \sqrt{\lambda_x}, \qquad \frac{R_y}{D} = \sqrt{\lambda_y}.$$

Having recovered these relative coordinate scales, we can reconstruct Z directly from A. Recall that Z is obtained by subtracting the configuration centroid and dividing all coordinates by the common diameter D:

$$Z_{ik} = \frac{q_{ik} - \bar{q}_k}{D}.$$

Averaging the identity $q_{ik} = b_k + R_k A_{ik}$ over all landmarks gives

$$\bar{q}_k = b_k + R_k \bar{A}_k.$$

Subtracting this expression from the original coordinate equation yields

$$q_{ik} - \bar{q}_k = R_k(A_{ik} - \bar{A}_k).$$

Dividing by D, we obtain

$$\begin{aligned} Z_{ik} &= \frac{q_{ik} - \bar{q}_k}{D} \\ &= \frac{R_k}{D}(A_{ik} - \bar{A}_k) \\ &= \sqrt{\lambda_k}(A_{ik} - \bar{A}_k). \end{aligned}$$

This gives an explicit reconstruction of every coordinate of Z from the simple hybrid. First, the normalized pairwise distances determine the relative horizontal and vertical scales through the unique solution of $H\lambda = s$. These scales are then applied to the centered coordinates of A. The result is exactly the configuration centered at its centroid and normalized by its diameter. Hence (A, δ) uniquely determines Z.

We next establish the converse direction, showing that Z uniquely determines the simple hybrid representation (A, δ).

By definition,

$$Z_i = \frac{q_i - \bar{q}}{D}.$$

For any pair of landmarks $i < j$, subtracting their normalized coordinates gives

$$Z_i - Z_j = \frac{q_i - q_j}{D},$$

because the centroid terms cancel. Taking Euclidean norms, we obtain

$$\|Z_i - Z_j\|_2 = \frac{\|q_i - q_j\|_2}{D} = \delta_{ij}.$$

Therefore all normalized pairwise distances are recovered directly from Z as

$$\delta_{ij} = \|Z_i - Z_j\|_2. \tag{57}$$

It remains to recover the coordinate-wise normalized landmarks A. Since

$$Z_{ik} = \frac{q_{ik} - \bar{q}_k}{D}$$

and $D > 0$, the transformation from q_{ik} to Z_{ik} consists of subtracting the same constant and dividing by the same positive factor for every landmark. It therefore preserves the ordering of the coordinates along each axis.
In particular,

$$\min_j Z_{jk} = \frac{b_k - \bar{q}_k}{D},$$

while

$$\max_j Z_{jk} = \frac{b_k + R_k - \bar{q}_k}{D}.$$

The coordinate range of Z is consequently

$$\max_j Z_{jk} - \min_j Z_{jk} = \frac{R_k}{D}.$$

Applying coordinate-wise min–max normalization to Z, we obtain

$$\begin{aligned} \frac{Z_{ik} - \min_j Z_{jk}}{\max_j Z_{jk} - \min_j Z_{jk}} &= \frac{\frac{q_{ik} - b_k}{D}}{\frac{R_k}{D}} \\ &= \frac{q_{ik} - b_k}{R_k} \\ &= A_{ik}. \end{aligned}$$

Thus

$$A_{ik} = \frac{Z_{ik} - \min_j Z_{jk}}{\max_j Z_{jk} - \min_j Z_{jk}}. \tag{58}$$

Equations (57) and (58) recover every component of the simple hybrid representation from Z. Together with the preceding reconstruction of Z from (A, δ), this establishes that ϕ_s and ϕ_Z uniquely determine one another. Finally, we characterize the geometric ambiguity shared by the two representations. Suppose that two labeled configurations Q and Q' produce the same Z. Writing D and D' for their respective diameters, we have

$$\frac{q'_j - \bar{q}'}{D'} = \frac{q_j - \bar{q}}{D}$$

for every landmark j. Rearranging gives

$$q'_j = \frac{D'}{D} q_j + \left(\bar{q}' - \frac{D'}{D} \bar{q} \right).$$

Hence

$$q'_j = c q_j + t, \qquad c = \frac{D'}{D} > 0, \qquad t = \bar{q}' - c\bar{q}.$$

Thus two configurations with the same Z differ only by a translation and a positive uniform scaling. Conversely, if $q'_j = c q_j + t$ for every landmark, with $c > 0$, then

$$\bar{q}' = c\bar{q} + t, \qquad D' = cD,$$

and therefore

$$Z'_j = \frac{cq_j + t - (c\bar{q} + t)}{cD} = Z_j.$$

Such transformations consequently leave Z unchanged and, by the established equivalence, also leave (A, δ) unchanged.

Because Z retains the coordinate values of the landmarks in the fixed image frame, neither representation additionally quotients out rotations or reflections. Therefore the simple hybrid and centered, diameter-normalized representations encode precisely the same labeled configuration up to translation and positive uniform scaling. □

3.1 Why noncollinearity matters

Noncollinearity is a sufficient condition for the rank-two property used in the reconstruction, not a requirement of the implemented evaluation mask. Without this condition, the normalized pairwise distances need not uniquely determine the relative horizontal and vertical scales.

To see this, consider distinct $t_i \in [0, 1]$, including 0 and 1, and define two collinear configurations by

$$Q_i = (t_i, t_i), \qquad Q'_i = (t_i, 2t_i).$$

Both configurations have the same coordinate-wise normalized landmarks,

$$A_i(Q) = A_i(Q') = (t_i, t_i).$$

Their pairwise Euclidean distances differ by a constant factor, but this factor is canceled by diameter normalization. Consequently,

$$\delta_{ij}(Q) = \delta_{ij}(Q') = |t_i - t_j|.$$

Thus the two configurations have identical simple hybrid representations. However, their centered, diameter-normalized coordinates are

$$Z_i(Q) = \frac{t_i - \bar{t}}{\sqrt{2}}(1, 1), \qquad Z_i(Q') = \frac{t_i - \bar{t}}{\sqrt{5}}(1, 2),$$

which are different. The ambiguity arises because all landmark differences have proportional horizontal and vertical components. The pairwise distance equations therefore constrain only a combination of the two squared relative scales, rather than determining them separately. This example shows that the simple hybrid need not recover Z for collinear configurations. The experiment can nevertheless retain collinear hands whenever the common usability criteria are satisfied.

4 The angular block is determined by normalized distances

Lemma 1 (Unsigned angles from three pairwise distances)**.** *For any triple (a, b, c) with nonzero arms at b, the unsigned angle θ_{abc} satisfies*

$$\cos\theta_{abc} = \frac{d_{ab}^2 + d_{bc}^2 - d_{ac}^2}{2d_{ab}d_{bc}} = \frac{\delta_{ab}^2 + \delta_{bc}^2 - \delta_{ac}^2}{2\delta_{ab}\delta_{bc}}. \tag{59}$$

Therefore the full vector ϕ_θ is a deterministic function of ϕ_δ, and ϕ_f is a deterministic function of ϕ_s.

Proof. Put $u = q_a - q_b$ and $v = q_c - q_b$. Since $q_a - q_c = u - v$,

$$d_{ac}^2 = \|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2u^\top v = d_{ab}^2 + d_{bc}^2 - 2d_{ab}d_{bc}\cos\theta_{abc}.$$

Rearranging yields the first identity. Substitute $d_{ij} = D\delta_{ij}$; all terms have degree two in D, so the common factor D^2 cancels. Since $\arccos : [-1, 1] \to [0, \pi]$ is single-valued, each unsigned angle is uniquely recovered. Applying this identity to the 15 selected triples recovers the entire angle block. □

References

[Kendall(1984)] D. G. Kendall. Shape manifolds, Procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984. https://doi.org/10.1112/blms/16.2.81.

[Dryden and Mardia(2016)] I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis, with Applications in R*. 2nd edition, Wiley, 2016. https://doi.org/10.1002/9781119072492.

[Gower(1975)] J. C. Gower. Generalized Procrustes analysis. *Psychometrika*, 40:33–51, 1975. https://doi.org/10.1007/BF02291478.