Influence of the Non-Uniqueness of Logarithm

A note on "Symmetry and Orbit Detection via Lie-Algebra Voting" (SGP 2016)

Zeyun Shi, Juncheng Zhong, Pierre Alliez, Mathieu Desbrun, Hujun Bao, Jin Huang

Abstract

In this report, we show that the non-uniqueness of the matrix logarithm has only a small influence on the results of [Shi et al., 2016]. More specifically, it has no influence on transformations in SO(3) (when considering both transformation and inverse transformation as mentioned in the paper); for more general transformations, the ordering of the distances with or without considering the non-uniqueness of logarithm remain mostly identical.

1 Non-Unique Logarithm of a Transformation

In this section, we briefly introduce the non-uniqueness issue of the logarithm of a transformation.

1.1 Periodicity in logarithm of a SO(3) transformation

The issue comes from the 2π periodicity of rotation. For a rotation $R \in \mathrm{SO}(3)$, its logarithm is a skew-symmetric (i.e., anti-symmetric) matrix S assembled from a length-3 vector $\boldsymbol{\omega} = \theta U \in \mathbb{R}^3$, where U is a unit vector in $\mathbb{R}^3$ representing the rotation axis while the scalar $\theta = \|\boldsymbol{\omega}\| \geq 0$ is the angle of the rotation. The matrix S is then assembled as:

$$S(\boldsymbol{\omega}) = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \tag{1}$$

The logarithm of R is not unique because the angle of a rotation is defined up to a multiple of 2π, which leads to the fact that

$$R = \exp(S(\theta U)) = \exp(S((\theta + 2k\pi)U)), \quad k \in \mathbb{Z}. \tag{2}$$

Usually, one picks θ to be the smallest rotation angle between 0 and π, leading to a modified definition of logarithm:

$$\begin{aligned} &\log(R) = S(\phi\boldsymbol{\omega}), \quad \phi \in [0,\pi], \|\boldsymbol{\omega}\| = 1, \\ &\phi = \min_k |\theta + 2k\pi|, \quad \boldsymbol{\omega} = \begin{cases} U & \theta + 2k\pi \geq 0, \\ -U & \theta + 2k\pi < 0. \end{cases} \end{aligned} \tag{3}$$

However, this arbitrary choice renders the construction of the logarithm not continuous for rotations of angles around π.

In the remainder of this note, we will often (abusively) refer to ω as the logarithm of a rotation R as it simplifies the exposition.

1.2 Non-unique logarithm of a Sim(3) transform

As a consequence of the non-uniqueness of logarithm for a rotation, the logarithm of similarity transform $T \in \mathrm{SIM}(3)$ is also not unique. A similarity transform has the form

$$T = \begin{pmatrix} R & \mathbf{t} \\ 0 & w^{-1} \end{pmatrix}, \tag{4}$$

and its logarithm can be represented by a skew-symmetric matrix assembled from a $\mathbb{R}^7$ vector (see Eq. (2) in our original paper). For consiveness, we will use

$$\log(T) = (\boldsymbol{\omega}^t, \mathbf{u}^t, s)^t, \tag{5}$$

where $\boldsymbol{\omega} = \log(R)$, $\mathbf{u} \in \mathbb{R}^3$ (which is related to the value of $\boldsymbol{\omega}$, $\mathbf{t}$ and w), $s = \log(w)$ (details about the calculation can be found in [Eade, 2013]).

When $\mathbf{t} = 0$ and $w = 1$, $T \in \mathrm{SO}(3)$, the logarithm of T equals to $(\boldsymbol{\omega}^t, \mathbf{0}^t, 0)^t$, and we are back to the pure rotation case as expected.

2 Distance Computation in [Shi et al., 2016]

The variational (squared) distance used in our paper (Eq. (9) and Eq. (14)) is defined as

$$\mathcal{D}^\dagger(T_A, T_B) \triangleq \min\{\mathcal{D}(T_A, T_B), \mathcal{D}(T_A, T_B^{-1})\}, \tag{6}$$

and

$$\mathcal{D}(T_A, T_B) = \begin{cases} \beta\|\mathbf{u}\|^2 + \gamma\, s^2 & \boldsymbol{\omega} = 0 \\ \alpha\|\boldsymbol{\omega}\|_2^2 + \beta\langle \boldsymbol{\omega}/\|\boldsymbol{\omega}\|, \mathbf{u}\rangle^2 + \gamma\, s^2 & \text{otherwise} \end{cases}, \tag{7}$$

where $\boldsymbol{\omega} = \boldsymbol{\omega}_A - \boldsymbol{\omega}_B$, $\mathbf{u} = \mathbf{u}_A - \mathbf{u}_B$, $s = s_A - s_B$.

In [Shi et al., 2016], we used log in the computation, but did not properly analyze the effect of considering all possible logarithms.

In the following, we will show that the non-uniqueness has no influence for transformations in SO(3). For more general Sim(3) transformations, we will prove that ignoring the non-uniqueness of logarithm leads to different distances, but has only little influence on the results.

3 The case of SO(3)

When $R_A, R_B \in \mathrm{SO}(3)$, we have

$$\begin{aligned} \boldsymbol{\omega}_A = \log(R_A) = \theta_a A, \quad \boldsymbol{\omega}_B = \log(R_B) = \theta_b B, \\ \theta_a, \theta_b \in [0,\pi], \|A\| = \|B\| = 1. \end{aligned} \tag{8}$$

The distance can be computed as

$$\mathcal{D}(R_A, R_B) = \|\theta_a A - \theta_b B\|_2^2 = \theta_a^2 + \theta_b^2 - 2\theta_a\theta_b\langle A, B\rangle. \tag{9}$$

If we consider the non-uniqueness of the logarithm, we have a different definition of distance $\widetilde{\mathcal{D}}$:

$$\widetilde{\mathcal{D}}_{a,b}(R_A, R_B) = \|(\theta_a + 2a\pi)A - (\theta_b + 2b\pi)B\|_2^2, a, b \in \mathbb{Z}. \tag{10}$$

In many cases, $\mathcal{D}(R_A, R_B) = \widetilde{\mathcal{D}}_{0,0}(R_A, R_B)$ may not be equal to $\min_{a,b} \widetilde{\mathcal{D}}_{a,b}$. However, if we consider the inverse of the transformation instead, they are identical:

$$\mathcal{D}^\dagger(R_A, R_B) \equiv \widetilde{\mathcal{D}}^\dagger(R_A, R_B). \tag{11}$$

The proof will be given in this section.

3.1 Without the inverse transformation

To compare the value of $\mathcal{D}(R_A, R_B)$ and $\widetilde{\mathcal{D}}_{a,b}(R_A, R_B)$, we take a simple case with $A = -B = W$ below.

First we have

$$\mathcal{D}(R_A, R_B) = \theta_a^2 + \theta_b^2 + 2\theta_a\theta_b = (\theta_a + \theta_b)^2. \tag{12}$$

Then we find

$$\begin{aligned}\widetilde{\mathcal{D}}_{a,b}(R_A, R_B) &= \|(\theta_a + 2a\pi)W + (\theta_b + 2b\pi)W\| \\ &= \|(\theta_a + \theta_b + 2a\pi + 2b\pi)W\| \\ &= (\theta_a + \theta_b + 2(a+b)\pi)^2.\end{aligned} \tag{13}$$

When $\theta_a + \theta_b > \pi$, we deduce

$$\widetilde{\mathcal{D}}_{0,-1}(R_A, R_B) = (2\pi - (\theta_a + \theta_b))^2 < \mathcal{D}(R_A, R_B), \tag{14}$$

which means that in general cases, $\mathcal{D}(R_A, R_B)$ is *not* the smallest distance considering periodic angle changes, namely $\mathcal{D}(R_A, R_B) \neq \min_{a,b} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B)$.

3.2 Considering the inverse transformation

However, in our application, the inverse of R_B is also involved in computing the distance, which makes a significant difference. An intuitive illustration can be found in Fig. 1.

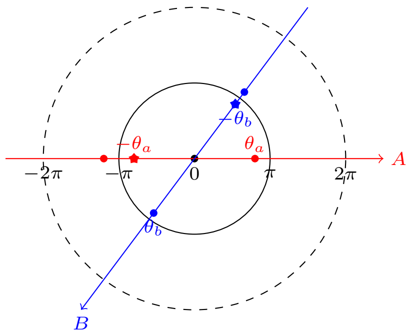

Figure 1: Illustration: the red circular dots indicate the logarithms of rotation R_A, and the blue ones are for rotation R_B. The red and blue star shapes are the logarithms of R_A^{-1} and R_B^{-1} respectively.

In the following, we will prove that considering the non-uniqueness or not has no influence to the result. We first introduce two lemmas which help to simplify the proof.

Lemma 1. $\min_{a,b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B) = \min_{b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B)$.

Proof. Note that $\theta_a, \theta_b \in [0, \pi]$. If $ab \geq 0$, $\widetilde{\mathcal{D}}_{a,b}(R_A, R_B) - \widetilde{\mathcal{D}}_{0,b-a}(R_A, R_B) = 4a\pi(\theta_a + \theta_b + 2b\pi)(1 - \langle A, B\rangle) \geq 0$. If $ab < 0$, $\widetilde{\mathcal{D}}_{a,b}(R_A, R_B) - \widetilde{\mathcal{D}}_{0,b+a}(R_A, R_B) = 4a\pi(\theta_a - \theta_b - 2b\pi)(1 + \langle A, B\rangle) \geq 0$. Therefore, $\min_{a,b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B) = \min_{b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B)$. □

Figure 2 gives an intuitive illustration of the above lemma.

Lemma 2. $\min_{a,b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B) = \min_{b\in\{0,-1\}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B)$.

Proof. Using the previous lemma, we only need to prove $\min_{b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B) = \min_{b\in\{0,-1\}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B)$. Note that $\theta_a, \theta_b \in [0, \pi]$. When $b \neq -1$, $\widetilde{\mathcal{D}}_{0,b}(R_A, R_B) - \widetilde{\mathcal{D}}_{0,0}(R_A, R_B) = 4b\pi(b\pi + \theta_b - \theta_a\langle A, B\rangle) \geq 0$. Thus, $\min_{b\in\mathbb{Z}} \widetilde{\mathcal{D}}_{0,b}(R_A, R_B)$ has minimum value when $b = 0$ or $b = -1$. □

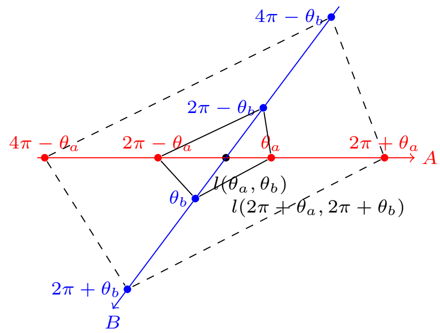

Figure 2: Illustration. We use line $l(\theta_a, \theta_b)$ to represent the logarithm distance between two rotations, it is clear that moving θ_a, θ_b towards the origin simultaneously by 2π will make the length of the line shorter. For example, $l(\theta_a, \theta_b) < l(2\pi+\theta_a, 2\pi+\theta_b)$, $l(2\pi-\theta_a, \theta_b) < l(4\pi-\theta_a, 2\pi+\theta_b)$.

Now, we prove the following theorem:

Theorem 1.

$$\mathcal{D}^\dagger(R_A, R_B) \equiv \widetilde{\mathcal{D}}^\dagger(R_A, R_B), \quad R_A, R_B \in \mathrm{SO}(3). \tag{15}$$

Proof. Let $\langle A, B\rangle = \gamma \in [-1, 1]$. We have

$$\begin{aligned}\widetilde{\mathcal{D}}^\dagger &\triangleq \min\{\min_{a,b} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B), \min_{a,b} \widetilde{\mathcal{D}}_{a,b}(R_A, R_B^{-1})\} \\ &= \min_{b\in\{0,-1\}} \|\theta_a A \pm (\theta_b + 2b\pi)B\|_2^2 \\ &= \min_{b\in\{0,-1\}} \theta_a^2 + (\theta_b + 2b\pi)^2 \pm 2\theta_a(\theta_b + 2b\pi)\gamma.\end{aligned} \tag{16}$$

When $b = 0$,

$$\widetilde{\mathcal{D}}^\dagger_{b=0} = \min\{\theta_a^2 + \theta_b^2 \pm 2\theta_a\theta_b\gamma\}. \tag{17}$$

When $b = -1$,

$$\widetilde{\mathcal{D}}^\dagger_{b\neq 0} = \min\{\theta_a^2 + (\theta_b - 2\pi)^2 \pm 2\theta_a(\theta_b - 2\pi)\gamma\}. \tag{18}$$

If $\gamma > 0$,

$$\begin{aligned}\widetilde{\mathcal{D}}^\dagger_{b=0} &= \theta_a^2 + \theta_b^2 - 2\theta_a\theta_b\gamma \\ \widetilde{\mathcal{D}}^\dagger_{b=-1} &= \theta_a^2 + (\theta_b - 2\pi)^2 + 2\theta_a(\theta_b - 2\pi)\gamma \\ \widetilde{\mathcal{D}}^\dagger_{b=0} - \widetilde{\mathcal{D}}^\dagger_{b=-1} &= -4(\pi - \theta_b)(\pi - \theta_a\gamma) < 0.\end{aligned} \tag{19}$$

If $\gamma < 0$,

$$\begin{aligned}\widetilde{\mathcal{D}}^\dagger_{b=0} &= \theta_a^2 + \theta_b^2 + 2\theta_a\theta_b\gamma \\ \widetilde{\mathcal{D}}^\dagger_{b=-1} &= \theta_a^2 + (\theta_b - 2\pi)^2 - 2\theta_a(\theta_b - 2\pi)\gamma \\ \widetilde{\mathcal{D}}^\dagger_{b=0} - \widetilde{\mathcal{D}}^\dagger_{b=-1} &= -4(\pi - \theta_b)(\pi + \theta_a\gamma) < 0.\end{aligned} \tag{20}$$

Thus, $\widetilde{\mathcal{D}}^\dagger = \widetilde{\mathcal{D}}^\dagger_{b=0} = \mathcal{D}^\dagger$. □

4 The case of $\mathrm{S{\small IM}}(3)$

In the previous section, we showed that the log non-uniqueness has no influence in our symmetry-finding application for $\mathrm{SO}(3)$ transformation. However, for similarity transformations ($\mathrm{S{\small IM}}(3)$), non-uniqueness does have consequences. A simple example is shown first, before studying the influence of these consequences on distance evaluations.

4.1 An example

Suppose there are two similarity transformations T_A and T_B, with

$$T_A = \begin{pmatrix} R_A & \mathbf{t}_A \\ 0 & 1 \end{pmatrix}, T_B = \begin{pmatrix} R_B & \mathbf{t}_B \\ 0 & 1 \end{pmatrix}, \tag{21}$$

such that

$$\begin{gathered} \log(R_A) = \theta_a A, \quad \log(R_B) = \theta_b B, \\ \theta_a, \theta_b \in (0, \pi], \quad A = (1,0,0)^t, \quad B = (0,1,0)^t, \\ \mathbf{t}_A = (t_a, 0, 0)^t, \quad \mathbf{t}_B = (0, 0, -2t_b)^t. \end{gathered} \tag{22}$$

Let $\widetilde{\theta}_a$ and $\widetilde{\theta}_b$ denote $\theta_a + 2a\pi$ and $\theta_b + 2b\pi$ for some $a, b \in \mathbb{Z}$ respectively. The logarithms of the two transformations then involve:

$$\mathbf{u}_A = (t_a, 0, 0)^t, \mathbf{u}_B = (\widetilde{\theta}_b t_b, 0, \widetilde{c}_b)^t, \tag{23}$$

where $\widetilde{c}_b$ is a value computed by Equs. (70)-(78) in [Eade, 2013]. According to the definition of variational distance, we have:

$$\begin{aligned} \widetilde{\mathcal{D}}(T_A, T_B) &= \alpha(\widetilde{\theta}_a^2 + \widetilde{\theta}_b^2) + \beta \frac{(t_a - \widetilde{\theta}_b t_b)^2 \widetilde{\theta}_a^2}{\widetilde{\theta}_a^2 + \widetilde{\theta}_b^2}, \\ \widetilde{\mathcal{D}}(T_A, T_B^{-1}) &= \alpha(\widetilde{\theta}_a^2 + \widetilde{\theta}_b^2) + \beta \frac{(t_a + \widetilde{\theta}_b t_b)^2 \widetilde{\theta}_a^2}{\widetilde{\theta}_a^2 + \widetilde{\theta}_b^2}. \end{aligned} \tag{24}$$

If we choose $t_b > 0$, $t_a = (\theta_b + 2\pi)t_b$, then we have

$$\begin{aligned} \widetilde{\mathcal{D}}_{0,1}(T_A, T_B) =& \alpha(\theta_a^2 + (\theta_b + 2\pi)^2), \\ \widetilde{\mathcal{D}}_{0,0}(T_A, T_B) =& \alpha(\theta_a^2 + \theta_b^2) + \beta \frac{(2\pi t_b)^2 \theta_a^2}{\theta_a^2 + \theta_b^2}, \\ \widetilde{\mathcal{D}}_{0,0}(T_A, T_B^{-1}) =& \alpha(\theta_a^2 + \theta_b^2) + \beta \frac{(2\theta_b t_b + 2\pi t_b)^2 \theta_a^2}{\theta_a^2 + \theta_b^2}. \end{aligned} \tag{25}$$

It is easy to see that $\widetilde{\mathcal{D}}_{0,0}(T_A, T_B) < \widetilde{\mathcal{D}}_{0,0}(T_A, T_B^{-1})$. By choosing a large enough t_b, we can make $\widetilde{\mathcal{D}}_{0,1}(T_A, T_B) < \widetilde{\mathcal{D}}_{0,0}(T_A, T_B) = \min\{\mathcal{D}(T_A, T_B), \mathcal{D}(T_A, T_B^{-1})\}$, which provides an example to demonstrate that $\mathcal{D}^\dagger > \widetilde{\mathcal{D}}^\dagger$.

4.2 Practical influence

However, when used in a voting framework, the non-uniqueness issue could be acceptable if the induced difference is small enough. In this section we will analyze the influence in more detail.

Criteria for influence In our application, the distance between two transformations is used in the mean shift and RANSAC clustering methods. In these voting methods, the *relative* distance value is more important than the absolute distance value. As an intuitive example, if T_C is closer to T_A than T_B (for a distance measurement D_x), a distance measurement D_y which gives the same order $D_y(T_A, T_C) < D_y(T_B, T_C)$ will lead to the same result when merging T_C into cluster centers T_A or T_B. In such a situation, we will say that D_y is consistent with D_x. If $\widetilde{\mathcal{D}}^\dagger$ and $\mathcal{D}^\dagger$ give the same orders for any transformation pairs, then even if these two distance definitions give different values in some cases, the clustering results computed from them may still be the same, depending on the specific clustering method being used. In order to measure the consistency between them, we first generate large number of random pairs of transformations, and then evaluate the change of the orders that the non-uniqueness creates.

Setting of experiments We first randomly sample $N = 1000$ *pairs* of transformations:

- for the rotation part, the three Euler angles are in $\mathcal{U}(0, 2\pi)$;
- for the translation part, each component is in $\mathcal{U}(-1, 1)$;
- for the scale component, it is in $\mathcal{U}(0.9, 1.1)$.

where $\mathcal{U}(x, y)$ indicates the uniform probabilistic distribution in $[x, y]$.

For each pair of transformations, we compute the distance using $\widetilde{\mathcal{D}}^\dagger$ (we evaluate $\widetilde{\mathcal{D}}_{a,b}$ with a, b in $[-10, 10]$), then sort pairs into $P_i = (T_{A,i}, T_{B,i})$ in ascending order so that $\widetilde{\mathcal{D}}^\dagger(T_{A,k}, T_{B,k}) \leq \widetilde{\mathcal{D}}^\dagger(T_{A,k+1}, T_{B,k+1})$. This order is viewed as "ground truth".

Then we compute $\mathcal{D}^\dagger(T_{A,i}, T_{B,i})$, and sort them in ascending order by a sequence of indices $o : [0, \cdots, N-1] \to [0, \cdots, N-1]$ so that $d_{o(k)} \leq d_{o(k+1)}$. Finally we check the difference between the sequences o and $[0, 1, \cdots, N-1]$, which order the pairs of transformations according to the distances from $\mathcal{D}^\dagger$ and $\widetilde{\mathcal{D}}^\dagger$ respectively.

Results In Figure 3, the coordinates of the i-th point are $(i, , o(i))$. We test for transformations in SO(3) and SIM(3) separately.

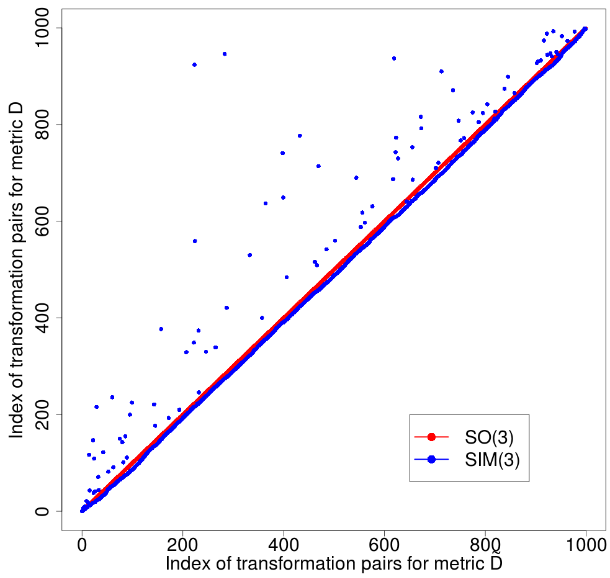

Figure 3: Comparison of the indices in ascending order of distances. The red points represent the result of taking all the transformation pairs from SO(3), while the blue points represent the result of taking transformation pairs from SIM(3). These red and blue points are almost distributed along lines of slope 1, indicating that the orders of transformation pairs in either SO(3) or SIM(3) are approximately the same regardless of considering non-uniqueness or not.

As evidence of the theoretical analysis in the previous sections, the points of the SO(3) transformation pairs form a perfect line of slope 1, indicating that the ascending order of the transformation pairs for these two distance definitions are totally unchanged. Although the distance values from $\mathcal{D}^\dagger$ and $\widetilde{\mathcal{D}}^\dagger$ are not always the same for SIM(3) transformation pairs, it still shows a clear line of slope 1, with only a limited number of points far from the line.

To further compare these two distance methods for $SIM(3)$ transformations, we first define the "discrete gradients" of o (difference of adjacent values), then measure the error by the difference of these gradients to 1, the expected slope of the perfect line.

$$\nabla o(k) \triangleq o(k+1) - o(k), \quad e(k) = |\nabla o(k) - 1|. \tag{26}$$

In Figure 4 we show the histogram distribution of error e. Clearly, most of the error are zero. This shows that the order of transformation pairs only changes a little.

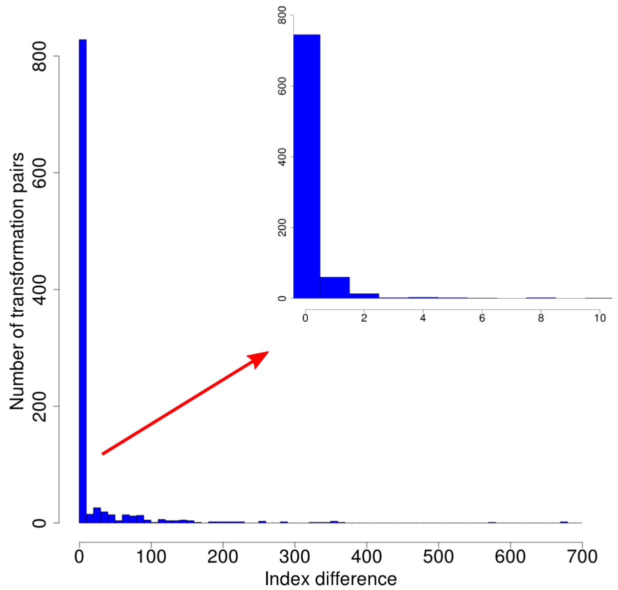

Figure 4: Histogram of e. The computation of index difference is explained in Section 4.2. The chart on the upper right side shows a zoomed-in view of distribution in $[0, 10]$.

In Figure 5, we visualize the error distribution under different transformation pair distances. For each interval $[x, y]$, the height of the bar is $h_{x,y} = \sum_{k=x}^{y} e(k)$. It is easy to see that the error is distributed randomly, meaning that the error does not affect pairs close to each other more than remote pairs.

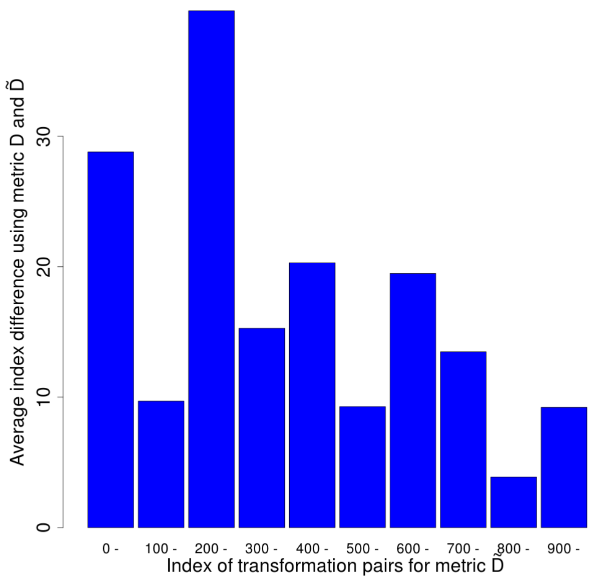

Figure 5: The distribution of error when ignoring the non-uniqueness. Each interval contains 100 successive transformation pairs. The index difference is random and small relative to the range (1000).

5 Conclusion

To sum up, considering the statistical character of the clustering method, the influence of the non-uniqueness of matrix logarithm is negligible. A more specific conclusion would require studying the effect of these errors on a particular clustering method.

Acknowledgement

We sincerely thank the anonymous reviewers and the audience of our presentation at SGP for pointing out this issue that we did not carefully examine before.

References

[Eade, 2013] Eade, E. (2013). Lie groups for 2D and 3D transformations. http://www.ethaneade.org/lie.pdf.

[Shi et al., 2016] Shi, Z., Alliez, P., Desbrun, M., Bao, H., and Huang, J. (2016). Symmetry and orbit detection via Lie-algebra voting. *Comput. Graph. Forum*, 35(5):217–227.