

ROTATION MATRIX SAMPLING SCHEME FOR MULTIDIMENSIONAL PROBABILITY DISTRIBUTION TRANSFER

P. Srestasathien *, S. Lawawirojwong, R. Suwantong, and P. Phuthong

Geo-Informatics and Space Technology Development Agency (GISTDA), Laksi, Bangkok, Thailand
(panu.siam.rata.patrawuth)@gistda.or.th

Commission VII, WG VII/4

KEY WORDS: Histogram specification, Distribution transfer, Color transfer, Data Gaussianization

ABSTRACT:

This paper address the problem of rotation matrix sampling used for multidimensional probability distribution transfer. The distribution transfer has many applications in remote sensing and image processing such as color adjustment for image mosaicing, image classification, and change detection. The sampling begins with generating a set of random orthogonal matrix samples by Householder transformation technique. The advantage of using the Householder transformation for generating the set of orthogonal matrices is the uniform distribution of the orthogonal matrix samples. The obtained orthogonal matrices are then converted to proper rotation matrices. The performance of using the proposed rotation matrix sampling scheme was tested against the uniform rotation angle sampling. The applications of the proposed method were also demonstrated using two applications i.e., image to image probability distribution transfer and data Gaussianization.

1. INTRODUCTION

In remote sensing, the analysis of multi-temporal data is widely used in many applications such as urban expansion monitoring, deforestation, change detection, and agriculture monitoring. One of the difficulties when dealing with multi-temporal images is the difference in image statistics because the images were acquired at different times.

In image classification context, it is possible that the image used for training the system can have statistics different from the image we want to classify. Moreover, images from different areas or times are combined to develop efficient classification system. To obtain a reliable classification result, the probability distributions of the training data and the data we want to classify should be similar.

In (Inamdar et al., 2008), it was pointed out that there are three approaches for relieving the problem of different statistics between images. The first approach is to use the atmospheric model to reduce the interference from atmosphere. The second is to use ground data to calibrate the images. The last one is to transform the image data such that its probability distribution of spectral information is similar to that of the reference or target image. The first two approaches are not possible, especially when the physical information of the atmosphere or ground data is not available.

In the last approach, digital numbers of pixels in a multispectral image are treated as multidimensional random variable. Its goal is to transfer statistical properties by reshaping the probability distribution of the source image such that its shape is matched with that of a reference one. In literature, The probability distribution transfer is also known as the histogram specification or histogram matching.

Pitié and Kokaram (Pitié and Kokaram, 2007) presented a linear approach for probability distribution transfer. The method uses Monge-Kantorovich theory of mass transportation to transform the shape of the probability distribution. Namely, the principle

is based on the Monge's transport problem which is to find minimal displacement for mass transportation. The solution from this linear approach can be used as initial solution for non-linear probability distribution transfer methods.

Similarly, in (Rabin et al., 2014), Rabin et al. also used the optimal transport technique to address the problem of statistical property transfer. Especially, the relaxed and regularized discrete optimal transport method was utilized. Moreover, the spatial distribution is also taken into account in the convex minimization process in order to deal with the spatial distortion introduced by the mapping.

In (Reinhard et al., 2001), a method for transfer color between image was proposed. By rotating the color from RGB color space to Lab color space in order to decorrelate the information between color channels, the statistical properties transfer between each channel of two images can be performed independently. Particularly, the color in the rotated space is translated and scaled to transfer the statistical properties. The major drawback of this approach is that it is limited to only RGB color space.

In stead of rotating the original data to a specific color space as presented in (Reinhard et al., 2001), Xiao and Ma (Xiao and Ma, 2006) computed the rotation using the axes defined by the principal components of each image. The transfer function is then formulated as a series of affine matrix transformation.

In (Pitié et al., 2007), an iterative method for transforming the probability distribution of an RGB image was proposed. The concept of this approach is to rotate data space using a random rotation matrix and perform 1D distribution transfer on each axis of the new coordinate system. The process iterates with different random rotations and stops when there is no change in the probability distribution of the modified image.

Inamdar et al. (Inamdar et al., 2008) applied the method proposed in (Pitié et al., 2007) to use with multidimensional data i.e., multispectral satellite images. Particularly, the N-D rotation matrix was used for transfer the multidimensional density function. The

*Corresponding author.

applicability of the proposed method was tested in three scenarios including 1.) supervised classification 2.) semi-supervised classification and 3.) change detection.

The probability distribution transfer is also used transform data such that the shape of its probability distribution is matched with a parametric probability distribution such as uniform, or normal distribution. For example, Laparra et al. (Laparra et al., 2011) proposed a method for data Gaussianization of which the aim is to make the transformed data distributed with a unit-covariance normal distribution.

In this paper, we adopted the iterative approach presented in (Pitié et al., 2007). Its disadvantage is that it was originally proposed for 3D probability distribution transfer using 3D rotation matrix. In order to deal with multidimensional rotation matrix, we proposed to use the sampling technique proposed by Stewart (Stewart, 1980) to generate a set of random orthogonal matrices. This sampling scheme uses Householder transformation converting normally distributed random variables to a random orthogonal matrix. This result in the uniform sampling of the orthogonal matrix. The obtained orthogonal matrices are then converted to proper rotation matrices. The iterative approach presented in (Pitié et al., 2007) can then be generalized to multidimensional case.

The rest of this paper is organized as follow. The concept of probability distribution transfer is briefly presented in Section 2.. In Section 3., the proposed random rotation matrix sampling scheme will be presented. The performance of the proposed method will be demonstrated in Section 4.. The conclusion will be discussed in Section 5..

2. PROBABILITY DISTRIBUTION TRANSFER

2.1 The one-dimensional case

In order to explain the distribution transfer which is also known as *histogram specification*, lets use the one dimensional data as an example because its formulation is very simple. Let c_s and c_t are the probability distribution of the source and target data, respectively. The goal of the distribution transfer is to transform the source data such that the transformed data has probability distribution similar to the target's one.

According to the change of variable formula, one obtains the following constraint:

$$c_s(u_s)du_s = c_t(u_t)du_t, \quad (1)$$

where u_s and u_t are respectively 1D random variables i.e., the digital number of a gray value image. Let the source data is transformed by a transformation T i.e., $T(u_s) = u_t$. By integrating both side of (1), we come up with:

$$\int^{u_s} c_s(u_s)du_s = \int^{T(u_s)} c_t(u_t)du_t \quad (2)$$

The transformation T can then be expressed in terms of the cumulative distributions of the source and target data:

$$C_s(u_s) = C_t(T(u_s)) \rightarrow T(u_s) = C_t^{-1} \circ C_s(u_s) \quad (3)$$

where C_s and C_t are the cumulative distributions of the source and target data, respectively. Since the 1-D cumulative distribution function is non-decreasing function, the mapping can then be easily implemented using look-up table.

Algorithm 1: Multi-dimensional probability distribution transfer

Input : \mathbf{X}_s : source data and \mathbf{X}_t : target data

Output: Transformed source data

```

1 while Not converge do
2   Choose an  $N \times N$  rotation matrix R;
3   At each iteration  $k$ , rotate the data using the rotation matrix
   R i.e.,  $\mathbf{X}_s^r(k) \leftarrow R\mathbf{X}_s(k-1)$  and  $\mathbf{X}_t^r \leftarrow R\mathbf{X}_t$ ;
4   Perform 1-D histogram matching on each of the axis in the
   rotated space;
5   Transform the adjusted source data back to the original
   space:  $\mathbf{X}_s(k) \leftarrow R^{-1}\mathbf{X}_s^r(k)$ ;
6 end
```

2.2 N-dimensional case

To transform the multidimensional probability distribution, this paper adopted the method proposed in (Pitié et al., 2007). The concept is similar to the Radon transformation such that the data is projected onto the principal axes of rotated space. The one-dimensional distribution transform is then performed independently on each principal axis. Therefore, the probability distribution is iteratively transformed to the target one using rotation matrix sequences.

The work flow of the multidimensional probability distribution transfer is illustrated in Algorithm 1. The proposed rotation matrix sampling scheme is presented in the next Section.

3. RANDOM SAMPLING OF ROTATION MATRIX

The proposed random rotation matrix sampling begins with the random sampling of orthogonal matrix. The obtained orthogonal matrices are then converted to proper rotation matrices. A matrix \mathbf{Q} is orthogonal if and only if $\det(\mathbf{Q}) = \pm 1$ and $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$. The set of $N \times N$ orthogonal matrices is denoted by $O(N)$ which is so called *orthogonal group*. The rotation matrix is a special case of the orthogonal group because the determinant of the rotation matrix is 1. Therefore, the set of rotation matrices is so called *special orthogonal group* which is denoted by $SO(N)$ where N is the dimension of the rotation matrix.

3.1 Uniformity of orthogonal matrix distribution

Unlike uniform sampling of real number, the sampling of orthogonal, or rotation matrices needs special care in order to avoid unintentionally bias. For example, uniform Euler angle sampling does not yield uniformly distributed rotation matrices. Particularly, it results in a set of rotation matrix that is heavily distributed in polar region (Kuffner, 2004).

In (Mitchell, 2008), it is pointed out that a scheme can uniformly sample orthogonal matrix $\mathbf{Q} \in O(N)$, if these conditions satisfy:

$$\begin{aligned} \mu(\Omega_1) &= \mu(\Omega_2) \\ P(\mathbf{Q} \in \Omega_1) &= P(\mathbf{Q} \in \Omega_2) \end{aligned}, \quad (4)$$

where μ is the Haar measure assigning orthogonal transformation invariant volume to subset¹ $\Omega_1, \Omega_2 \in O(N)$. This means that, when the set Ω_1 and Ω_2 have the same "volume", the sampling is uniform, if the chances of choosing a orthogonal matrix \mathbf{Q} from Ω_1 , or Ω_2 are equal. This constraint is also used for the uniform sampling of rotation matrices.

¹A measure can be intuitively interpreted as size, or volume of a set

3.2 Uniform orthogonal matrix sampling

To sample orthogonal matrices, the method proposed by Stewart (Stewart, 1980) is adopted because of uniform sampling. The obtained orthogonal matrices are then converted to proper rotation matrices. The orthogonal matrix sampling method is based on the use of Householder transformation to decompose a square matrix into orthogonal and upper triangular matrices i.e., QR decomposition.

Given a vector \mathbf{a} , there exists a Householder transformation H such that:

$$H\mathbf{a} = r\mathbf{e}, \quad (5)$$

where $\mathbf{e} = [1, 0, \dots, 0]^T$ and $r = \pm\|\mathbf{a}\|_2$. The Householder transformation is defined as:

$$H = I - 2\frac{\mathbf{u}\mathbf{u}^T}{\|\mathbf{u}\|_2^2}, \quad (6)$$

where \mathbf{u} is a non-zero vector. It is trivial to prove that the Householder transformation H is symmetric orthogonal. The geometrical meaning of the Householder transformation is the reflection of a vector \mathbf{a} across the hyperplane orthogonal to \mathbf{u} which is so called *reflector*. The reflector \mathbf{u} can be computed by:

$$\mathbf{u} = \text{sign}(a_1)\|\mathbf{a}\|_2\mathbf{e} + \mathbf{a}, \quad (7)$$

where a_1 is the first element of the vector \mathbf{a} .

A series of Householder transformation can then be utilized to perform orthogonal triangularization on an $N \times N$ matrix \mathbf{A} . That is, there exists a sequence of Householder transformation H_2, H_3, \dots, H_N satisfying:

$$H_2H_3 \dots H_N\mathbf{A} = \mathbf{U}, \quad (8)$$

where \mathbf{U} is upper triangular. By setting

$$\mathbf{Q} = H_N \dots H_3H_2, \quad (9)$$

we may write

$$\mathbf{A} = \mathbf{Q}\mathbf{U}, \quad (10)$$

which is known as QR decomposition².

To sampling orthogonal matrices, the method proposed in (Stewart, 1980) is based on the decomposition of random $N \times N$ matrix \mathbf{A} whose elements are independently sampled from normal distribution with zero mean and a common variance. It was proved that the QR decomposition of the random matrix such that the diagonal elements of the upper triangular matrix are positive results in the uniform distribution of orthogonal matrix \mathbf{Q} according to Haar measure.

3.3 Optimal rotation matrix

The obtained orthogonal matrix can be an improper rotation matrix i.e., $|Q| \neq 1$. In this paper, the method proposed by Kanatani (Kanatani, 1994) is adopted to project the obtained orthogonal matrix onto the $SO(N)$ manifold. Therefore, the projected orthogonal matrix is a rotation matrix.

Given an orthogonal matrix \mathbf{Q} , the goal is then to find the rotation matrix maximizing $\text{trace}(\mathbf{R}^T\mathbf{Q})$. Let the Singular Value Decomposition (SVD) of the matrix \mathbf{Q} be $\mathbf{Q} = \mathbf{W}\mathbf{D}\mathbf{V}^T$. The optimal

²In literature, this upper triangular matrix is usually denoted by \mathbf{R} according to the name of this factorization i.e., QR factorization.

Algorithm 2: Random rotation matrix sampling

Input : N :size of the $N \times N$ rotation matrix; σ :standard deviation of the normal distribution

Output: A sampled rotation matrix \mathbf{R}

```

1 for  $i \leftarrow N$  to 2 do
2   Sampling  $\mathbf{a}_i \in \mathbb{R}^i$  from normal distribution  $\mathcal{N}(\mathbf{0}_i, \sigma^2\mathbf{I}_i)$ ;
3   Compute Householder transformation  $\hat{H}_i$  reducing  $\mathbf{a}_i$  to  $r_i\mathbf{e}$ ;
4    $H_i \leftarrow \text{diag}(\mathbf{I}_{N-i}, \hat{H}_i)$ ;
5 end
6  $\mathbf{D} \leftarrow \text{diag}(\text{sign}(r_N), \text{sign}(r_{N-1}), \dots, \text{sign}(r_2))$ ;
7  $\mathbf{Q} \leftarrow \mathbf{D}H_NH_{N-1} \dots H_2$ ;
8  $\mathbf{W}\mathbf{D}\mathbf{V}^T \leftarrow \text{SVD}(\mathbf{Q})$ ;
9  $\mathbf{R} \leftarrow \mathbf{W} \text{diag}(1, 1, \dots, 1, \det(\mathbf{W}\mathbf{V}^T))\mathbf{V}^T$ ;

```

rotation matrix can then be computed as follow:

$$\mathbf{R} = \mathbf{W} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & \det(\mathbf{W}\mathbf{V}^T) \end{bmatrix} \mathbf{V}^T \quad (11)$$

It can be easily proved that this projection provides proper rotation matrix i.e., $|\mathbf{R}| = (|\mathbf{W}||\mathbf{V}^T|)^2 = 1$ because $|\mathbf{W}| = |\mathbf{V}| = \pm 1$

By combining the orthogonal matrix sampling with projection onto $SO(N)$ manifold, we obtain the scheme for sampling the rotation matrix which is illustrated in Algorithm 2. The multi-dimensional rotation matrix can then be randomly sampled and used in the probability distribution transfer.

4. EXPERIMENTAL RESULTS

4.1 Image to image probability distribution transfer

In this experiment, we demonstrate the effectiveness of the multi-dimensional probability distribution transfer. The performance of the proposed method was evaluated in terms of quality and quantity.

In order to show the evolution of the distribution transfer, we used the two benchmark images from stereo dataset i.e., cones and teddy images³ because of their color variations. In Figure 1, the evolution of the bivariate distribution of the red and green bands is illustrated. The Figure shows the contour plots of color distribution functions at some iterations, which were estimated by kernel density function technique, particularly Epanechnikov kernel. It can be observed that the distribution transfer can transform the shape of the source distribution function to match with that of target one.

The proposed method was compared with the method using uniform angle sampling (Inamdar et al., 2008) e.g., uniform Euler angle distribution in 3D case. The performance indicator is the similarity between the transformed probability distribution and the target one. A popular similarity metric called Kullback-Leibler (KL) divergence is utilized which is defined as follows:

$$D_{KL}(c_s \| c_t) = \int_{\mathbf{u}} c_s(\mathbf{u}) \log \left(\frac{c_s(\mathbf{u})}{c_t(\mathbf{u})} \right) d\mathbf{u}. \quad (12)$$

³<http://vision.middlebury.edu/stereo/data/scenes2003/>

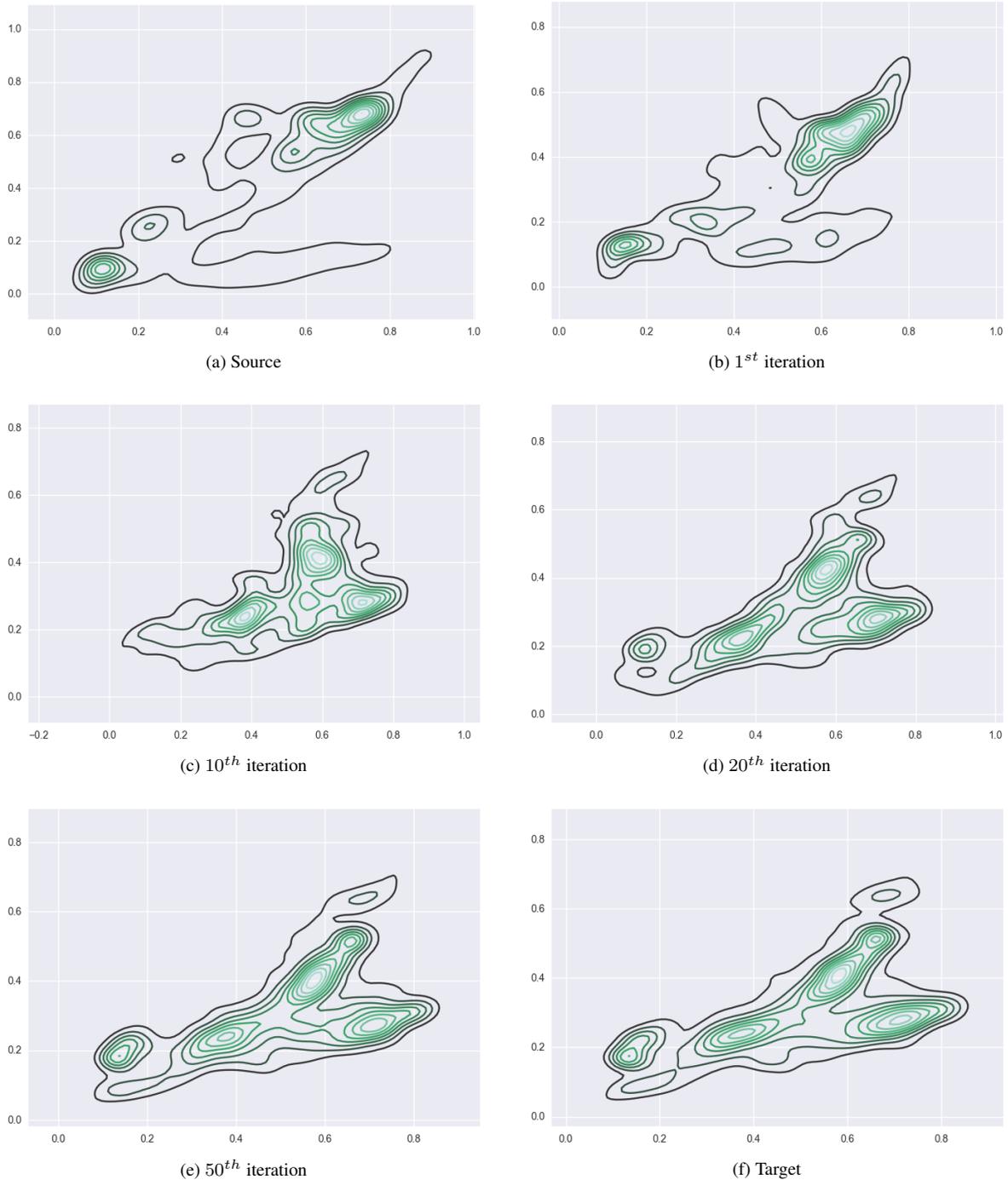


Figure 1: The evolution of source probability distribution function. The bivariate probability distribution of the red and green bands of the source data is shown in Figure 1a. The target probability distribution function is presented in Figure 1f.

It can be noticed that the KL divergence is not symmetric i.e., $D_{KL}(c_s, c_t) \neq D_{KL}(c_t, c_s)$. In order to make the metric symmetric, the average of the KL divergence is used:

$$D_{avg} = \frac{D_{KL}(c_s \| c_t) + D_{KL}(c_t \| c_s)}{2}. \quad (13)$$

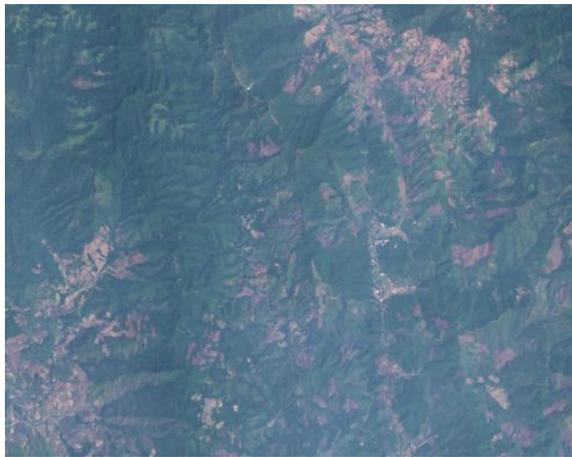
The images used in this comparison are Landsat 8 imageries, see Figure 2. The simulation was repeated 100 times and the averaged log Kullback-Leibler divergence at each iteration is shown in Figure 3. It can be observed that the proposed rotation sampling scheme yields better convergence than that of rotation angle sampling. The reason is that the rotation angle sampling can have

unintentionally bias in the rotation matrix samples.

The proposed method was also compared with the band-by-band distribution transfer. In this experiment, the performance was evaluated using Kolmogorov-Smirnov test which is a non-parametric test for determining whether two samples are drawn from the same distribution. If the Kolmogorov-Smirnov statistics is small, we therefore cannot reject the null-hypothesis that those two samples were drawn from the same distribution. The Kolmogorov-Smirnov statistics from the experiment are tabulated in Table 1. It can be observed that the multidimensional distribution transfer yields lowest Kolmogorov-Smirnov statistics.



(a)



(b)

Figure 2: Landsat 8 imageries used for comparing the performance of the proposed rotation sampling method and uniform rotation angle sampling.

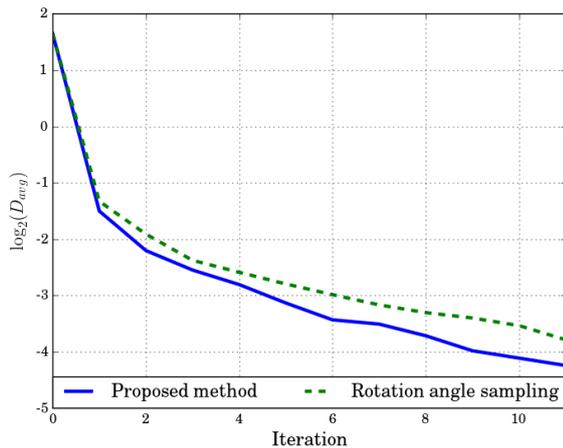
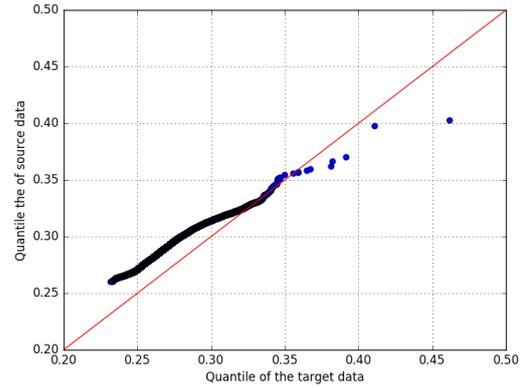
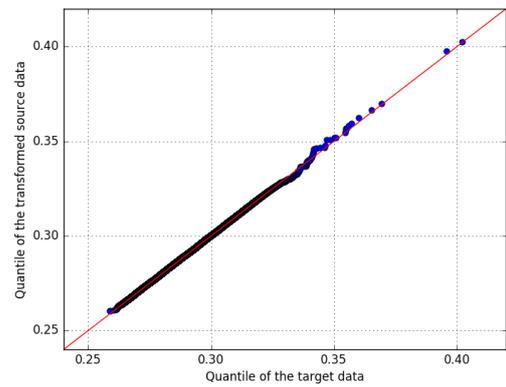


Figure 3: The averaged log Kullback-Leibler divergence from 100 simulations.

The performance is also analyzed visually using Q-Q plot, where Q stands for quantile. If two sampled are drawn from the same distribution, they are expected to line up on the line of identity. In Figure 4, the Q-Q plots of the red band with and without distribution transfer are illustrated. It can be concluded that the source



(a) Without distribution transfer



(b) With distribution transfer

Figure 4: The Q-Q plot of the red band of images illustrated in Figure 2.

	Original source data	1D transformation	ND transformation
Band 2	0.145	0.015	0.005
Band 3	0.040	0.003	0.002
Band 4	0.179	0.009	0.001
Band 5	0.214	0.003	0.0005
Band 6	0.151	0.005	0.0007
Band 7	0.095	0.008	0.0008
Band 8	1	0.046	0.210

Table 1: Kolmogorov-Smirnov statistics

and target data are not drawn from the same probability distribution function. After distribution transfer, the distribution of the transformed source data is similar to that of target data.

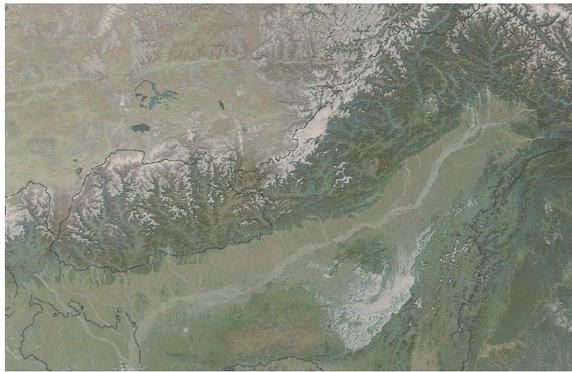
4.2 Data Gaussianization

In some applications, the data distribution has to be transformed into a parametric distribution function, particularly, the Gaussian or normal distribution. A popular approach is to use the power transformation e.g., Box-Cox transformation (Osborne, 2010), which is a parametric transformation. In this experiment, the distribution transfer was applied to transform the shape of the distribution function to that of normal distribution.

To transform the source images probability distribution function, the mean (μ) and the covariance matrix (Σ) of the source data



(a)



(b)

Figure 5: The MODIS imagery before and after data Gaussianization.

were computed. The target data was then drawn from the normal distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. The source data was then iteratively transformed to be distributed with the normal distribution $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. The source data used in this experiment was the true-color MODIS⁴ image of the semi-arid Tibetan Plateau, illustrated in Figure 5a. The image after distribution function transformation is shown in Figure 5b.

To quantitatively test the normality of the transformed source data, the kurtosis and skew were used (Ghasemi and Zahediasl, 2012). The kurtosis was used for measure the peakedness of the probability distribution function. The asymmetry of the distribution function was measured by skew. The shape of a distribution function is similar to that of normal distribution function, if the kurtosis is close to 3 and skew is negative. To pass the normality test with 95% confidence, the p-value must be greater than 0.05. Namely, by not reject the hypothesis of normality, no significant departure from normality is found. The statistics for normality test are tabulated in Table 2. It can be concluded that the multidimensional distribution transfer can transform the source data to be normally distributed.

The normality of the data can also be visually inspected using pairwise bivariate distribution plot shown in Figure 6. Particularly, the univariate distributions on each variable (color band) are shown on diagonal axis; the bivariate distributions between two variables are shown in off-diagonal elements. It can be observed that distribution transfer can change the shape of the non-normal distribution function to that of normal one.

It is also possible to check that the desired parameters of normal distribution function are obtained using the Kullback-Leibler divergence for multivariate normal distributions. Let the desired

⁴<http://modis.gsfc.nasa.gov/>

normal distribution be $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$, and the obtained normal distribution $\hat{\mathcal{N}}(\hat{\boldsymbol{\mu}}, \hat{\Sigma})$. Therefore, (12) can be simply modified:

$$D_{KL}(\hat{\mathcal{N}}\|\mathcal{N}) = \frac{1}{2} \left(\text{trace}(\Sigma^{-1}\hat{\Sigma}) + \ln \left(\frac{|\Sigma|}{|\hat{\Sigma}|} \right) - k + (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})^\top \Sigma^{-1} (\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \right), \quad (14)$$

The desired and obtained means are as following:

$$\begin{aligned} \boldsymbol{\mu} &= [0.2935 \ 0.3244 \ 0.2181]^\top \\ \hat{\boldsymbol{\mu}} &= [0.2938 \ 0.3248 \ 0.2183]^\top, \end{aligned} \quad (15)$$

and the desired and obtained covariance matrices:

$$\begin{aligned} \Sigma &= \begin{bmatrix} 0.0445 & 0.0365 & 0.0387 \\ 0.0365 & 0.0316 & 0.0329 \\ .03870 & 0.0329 & 0.0365 \end{bmatrix} \\ \hat{\Sigma} &= \begin{bmatrix} 0.0447 & 0.0366 & 0.03889 \\ 0.0366 & 0.0318 & 0.03311 \\ 0.0389 & 0.0331 & 0.03665 \end{bmatrix}. \end{aligned} \quad (16)$$

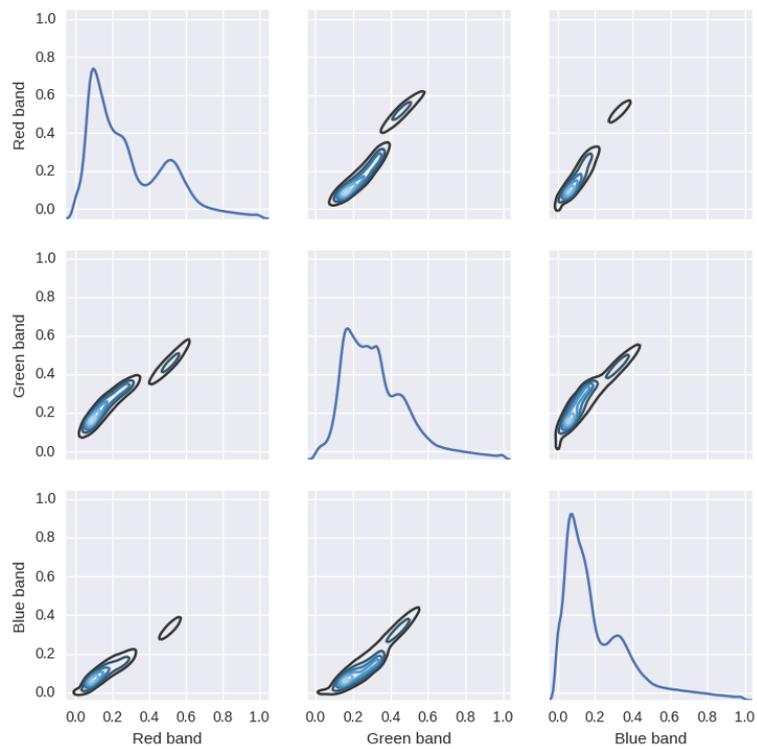
By using the average of $D_{KL}(\hat{\mathcal{N}}\|\mathcal{N})$ and $D_{KL}(\mathcal{N}\|\hat{\mathcal{N}})$, the similarity score is then 7.32×10^{-6} .

5. CONCLUSIONS

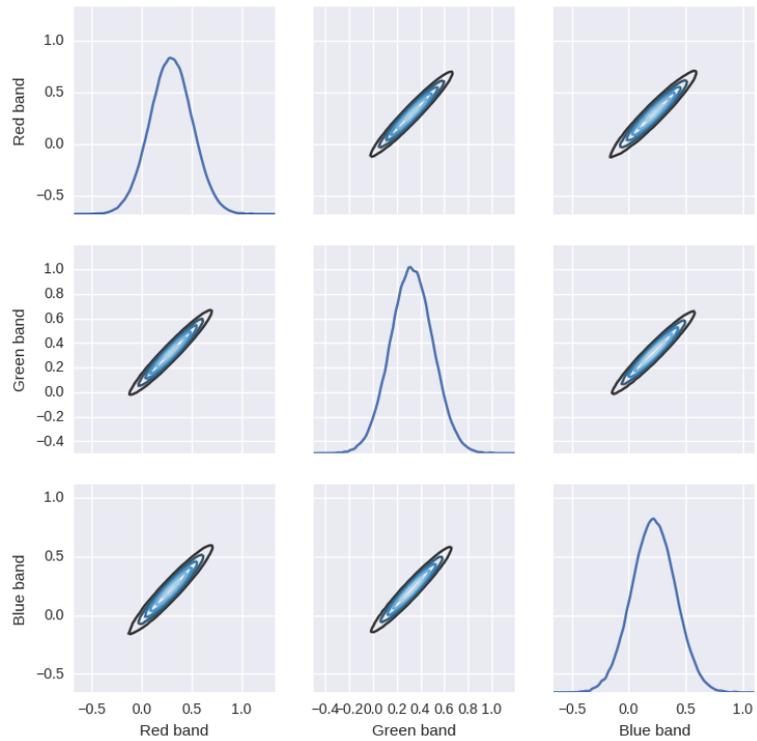
In this paper, a rotation matrix sampling scheme for distribution function transfer was presented. We proposed to use the Householder transformation to uniformly sampling the orthogonal matrices first and then transform the orthogonal matrices to rotation matrices. Using the advantage of Householder transformation for sampling orthogonal matrix, the rotation matrix can be sampling more efficient than using the rotation angle sampling. This efficiency was demonstrated in the experimental evaluation. The performance of the distribution transfer was tested with two applications i.e., image to image distribution transfer and data Gaussianization. It was demonstrated that the use multidimensional distribution transfer yield better results than that of 1D distribution transfer.

REFERENCES

- Ghasemi, A. and Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism* 10(2), pp. 486489.
- Inamdar, S., Bovolo, F., Bruzzone, L. and Chaudhuri, S., 2008. Multidimensional probability density function matching for pre-processing of multitemporal remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on* 46(4), pp. 1243–1252.
- Kanatani, K., 1994. Analysis of 3-d rotation fitting. *IEEE Trans. Pattern Anal. Mach. Intell.* 16(5), pp. 543–549.
- Kuffner, J., 2004. Effective sampling and distance metrics for 3d rigid body path planning. In: *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, Vol. 4, pp. 3993–3998.
- Laparra, V., Camps-Valls, G. and Malo, J., 2011. Iterative gaussianization: From ica to random rotations. *Neural Networks, IEEE Transactions on* 22(4), pp. 537–549.
- Mitchell, J. C., 2008. Sampling rotation groups by successive orthogonal images. *SIAM Journal on Scientific Computing* 30(1), pp. 525–547.



(a) Before transformation



(b) After data Gaussianization

Figure 6: Visualization of the pairwise relationships between two variables.

	Kurtosis			Skew			p-value		
	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
Original source data	3.057	4.339	5.199	0.874	1.081	1.497	0	0	0
Transformed source data	3.009	3.011	3.007	1.25×10^{-3}	-1.894×10^{-3}	-8.286×10^{-3}	0.422	0.290	0.627

Table 2: Experimental results for data normality test

Osborne, J. W., 2010. Improving your data transformations: Applying the box-cox transformation. *Pract. Assess. Res. Eval.* 15, pp. 1–9.

Pitié, F. and Kokaram, A., 2007. The linear monge-kantorovitch linear colour mapping for example-based colour transfer. *IET Conference Proceedings* pp. 23–23(1).

Pitié, F., Kokaram, A. C. and Dahyot, R., 2007. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding* 107(12), pp. 123 – 137.

Rabin, J., Ferradans, S. and Papadakis, N., 2014. Adaptive color transfer with relaxed optimal transport. In: *Image Processing (ICIP), 2014 IEEE International Conference on*, pp. 4852–4856.

Reinhard, E., Adhikhmin, M., Gooch, B. and Shirley, P., 2001. Color transfer between images. *Computer Graphics and Applications*, IEEE 21(5), pp. 34–41.

Stewart, G. W., 1980. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM Journal on Numerical Analysis* 17(3), pp. 403–409.

Xiao, X. and Ma, L., 2006. Color transfer in correlated color space. In: *Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications, VRCIA '06*, pp. 305–309.