MANIFOLD ALIGNMENT BASED COLOR TRANSFER FOR MULTIVIEW IMAGE STITCHING

Yuntao Qian∗ Danping Liao∗ Jun Zhou†

∗ College of Computer Science, Zhejiang University, Hangzhou 310027, China
† School of Information and Communication Technology, Griffith University, Nathan, Australia

ABSTRACT

In multiview image stitching, color transfer removes all color inconsistencies between different views under different illumination conditions and camera settings to make the stitching more seamless or visually acceptable. This paper presents a manifold alignment method to perform color transfer by exploring manifold structures of partially overlapped source and target images. Manifold alignment projects a pair of source and target images into a common embedding space in which not only the local geometries of color distribution in the respective images are preserved, but also the corresponding pixels in overlapped area across two images are pairwise aligned. Under this new space, color transfer can be considered as a matching problem between different manifolds, i.e. the color of each target pixel is replaced by the color of a source pixel that is nearest to this target pixel in this new space. Compared with other techniques in the literature, the proposed method makes full use of both the correspondences in overlapped area and the intrinsic color structures in the whole stitching scene so that a favorable performance is achieved.

Index Terms— color transfer, manifold alignment, multiview images, color distribution

1. INTRODUCTION

Color transfer is widely used in computational photograph and computer vision [1, 2]. When images captured from different views are stitched together, geometric registration is used to align them for generating a panorama photograph. However, these component images are not always stationary, causing a rugged transition between neighbors due to inconsistency of environment illumination, exposure variance, white balance and Gamma correction of camera. Color transfer addresses color or intensity disparity between them as if their composition is a single image. Color transfer for multiview image stitching can be defined as: given two partially overlapped images \( I_s \) and \( I_t \), one is source and the other is target, discover a mapping function \( f(C(i, j)) = C'(i, j) \) that changes the color characteristics of the target image to match that of the source image, where \( C(i, j) \) is the color of pixel \((i, j)\) in the target image. This definition is easily extended to more than two multi-view images.

Several color transfer methods have been presented. Reinhard et al. [3] proposed a simple but very effective global method based on the statistics of color distribution that is calculated by taking into account all pixels in a pair of images. The transferred color at pixel \((i, j)\) of target \( I_t \) is \( f(C(i, j)) = \mu_s + \frac{\sigma_s}{\sigma_t}(C(i, j) - \mu_t) \) where \((\mu_s, \sigma_s)\) and \((\mu_t, \sigma_t)\) are the means and standard deviations of the global color distributions of the source and target images in the \(a*b*\) color space. Pitie et al. [4] proposed a more advanced method that iteratively uses a one-dimensional probability density function (pdf) of color distribution to achieve a color transfer mapping. The advantage of these two methods is that some correspondences between target and sources images are not necessarily to be known in advance. However, their mapping functions are not accurate enough for image stitching. Multi-band image blending method [5], on the other hand, restricts color transfer to the overlapped area, which can provide a smooth transition in the stitching, but the other parts of scene are not changed so that color incoherence is still obvious.

In recent years, a main technical trend in color transfer for multiview image stitching is to utilize the correspondences in the overlapped area to form a mapping function, and then apply this function to all scenes. If a parametric model of color transfer is assumed, the learning of mapping function becomes a parameter estimation problem. In [6], the affine model and the linear model are assumed respectively, so their mapping functions are described as \( f(C(i, j)) = M \cdot C(i, j) \), where \( M \) is an affine matrix or an arbitrary \(3 \times 3\) matrix. \( M \) can be estimated by multiple linear regression. Sometimes the precise model of color transfer is difficult to obtain due to the complicate illumination environment and camera setting. Parameter free approaches directly search a color vector for each pixel in target image by a mapping table. Joint histogram of pairwise pixels in the overlapped area is always used to get robust color transfer mapping [7, 8]. More detailed overview of color transfer can be found in [9].
This paper addresses color transfer from the aspect of manifold alignment. Manifold alignment is one of the important frames for transfer learning, which establishes correspondence between data sets by a mapping function in the mutual embedding space [10, 11]. Therefore, color transfer can be seen as a transfer learning problem that applies the knowledge of color distribution of the source image to reconstruct the color distribution of the target image. In the proposed method, manifold alignment uses some corresponding pixel-pairs to align the manifolds of source and target images in an embedding space. Compared with other color transfer techniques that use either simple statistics of whole image or correspondences in overlapped area to estimate the transfer function, this paper provides a new way to combine color distribution information of the overlapped area and the corresponding images in order to derive a color mapping function.

2. MANIFOLD ALIGNMENT

Assume two data sets are represented by the matrices $S \in \mathbb{R}^{n_s \times p_s}$ and $T \in \mathbb{R}^{n_t \times p_t}$. $n_s$ and $n_t$ are the numbers of samples, and $p_s$ and $p_t$ are the numbers of features. Manifold alignment uses the corresponding sample-pairs between different data sets to find two transformation functions $f_s : R^{p_s} \rightarrow R^{p_s}$ and $f_t : R^{p_t} \rightarrow R^{p_t}$, which leads to implicit alignment of different data manifolds in the mutual embedding space $R^{p_t}$.

To achieve this goal, the first step is to construct the weighted graph that represents two input manifolds and their correspondences. In the graph of each data set, the nodes/samples $i$ and $j$ are connected by an edge if $i$ is among $k$ nearest neighbors of $j$ or $j$ is among $k$ nearest neighbors of $i$. The weight of an edge can be calculated by heat kernel.

$$W(i,j) = e^{-\frac{\|x_i-x_j\|^2}{\sigma}} \quad (1)$$

Now we combine these two individual graphs into a joint graph by constructing the edges between the nodes $i$ and $j$ from the different data sets. Typically, $W(i,j) = 1$ if the samples $x_i$ and $x_j$ form a corresponding pair. As a result, the size of the weighted adjacency matrix of this joint graph is $(n_s + n_t) \times (n_s + n_t)$. The diagonal of the matrix is zero, while other entries $W(i,j)$ are

$$W(i,j) = \begin{cases} 
\alpha_1 W(i,j) & \text{if } x_i \text{ and } x_j \text{ are both from the same data set} \\
\alpha_2 W(i,j) & \text{if } x_i \text{ and } x_j \text{ are a corresponding pair from two sets respectively} 
\end{cases} \quad (2)$$

The second step is to construct the cost function for manifold alignment. Here, a general loss function is defined as

$$H(f_{st}) = \sum_{i,j} \|f_{st}(x_i) - f_{st}(x_j)\|^2 W(i,j) \quad (3)$$

where $f_{st}$ is the joined function of $f_s$ and $f_t$ that are applied to data sets $S$ and $T$ respectively, and $W$ is the joint adjacency matrix. Eq.3 will derive a simplified form by the joint graph Laplacian $L = D - W$ where $D$ is a diagonal matrix with $D(i,i) = \sum_j W(i,j)$.

$$H(f_{st}) = \text{tr}(F'LF) \quad (4)$$

where $F$ is an $(n_s + n_t) \times q$ matrix that represents two data sets in the embedding space. If the manifold project is linear, $f_s$ and $f_t$ are two transformation matrices $F_s$ and $F_t$ with the sizes of $p_s \times q$ and $p_t \times q$ respectively, i.e. $f_s(x_i) = x_i F_s$ and $f_t(x_i) = x_i F_t$. Therefore, the cost function for the linear manifold alignment is described as

$$H(f_{st}) = \sum_{i,j} \|x_i F - x_j F\|^2 W(i,j) \quad (5)$$

where $F$ is a joint matrix of $F_s$ and $F_t$, and its size is $(p_s + p_t) \times q$. The cost function Eq.3 is simplified as

$$H(f_{st}) = \text{tr}(F'X'LDXF) \quad (6)$$

where $X$ is an $(n_s + n_t) \times (p_s + p_t)$ joint data matrix in which $S$ is on the upper left corner, and $T$ on the lower right corner. This optimization problem can be solved by Laplacian eigenmaps, i.e., $F$ is constructed by the minimum eigenvectors of the generalized eigenvector problem [11]

$$X'LDX = \lambda X'DXF \quad (7)$$

Linear alignment sacrifices the ability to align arbitrarily warped manifolds, but it is much faster than nonlinear alignment. In this paper, linear method is used in all experiments.

3. MANIFOLD ALIGNMENT BASED COLOR TRANSFER

Given a partially overlapped image pair $I_s$ and $I_t$ from multi-view images, manifold alignment method makes use of the correspondences in the overlapped area to perform implicit alignment of two manifolds embedding over source and target images respectively, so that all correspondences between two images can be estimated. Using the notations in section 2, the source and target images $I_s$ and $I_t$ are represented in their color spaces respectively by two matrices $S$ and $T$ in which the numbers of rows $n_s$ and $n_t$ are their respective image sizes and the numbers of columns $p_s$ and $p_t$ are the respective color dimensions. If their color spaces are spanned by each pixel’s $RGB$ or $I0\beta$ values, $p_s = p_t = 3$. And if the color space of a pixel is a $3 \times 3$ neighborhood centered at this pixel, $p_s = p_t = 3 \times 9$. Manifold alignment based color transfer algorithm is as follows:

1. Find the weighted adjacency matrices, $W_s$ and $W_t$ of $S$ and $T$ using $k$ nearest neighbors and heat kernel.
2. Construct the joint weighted adjacency matrix $W$ using $W_s$ and $W_t$ along with the corresponding pixel-pairs in overlapped area with Eq. 2, and then obtain the Laplacian matrix $L$.

3. Solve the optimization problem by eigenmaps to get the linear transformation matrices $F_s$ and $F_t$, and then the position of each pixel of two images in the $q$-dimensional embedding space is calculated.

4. Find the counterpart in source image for each pixel of target image, i.e., search the aligned pixel of source image for each pixel of target image in the embedding space, and then use the color of this source pixel to substitute the color of the target pixel.

Two points need to be mentioned for step 4: Firstly, if $p_s = p_t = q = 3$, this algorithm can be considered as a parametric model based method whose mapping function from an original target color $C(i,j)$ to its source color is

$$f(C(i,j)) = (F_t \cdot F_s^{-1}) \cdot C(i,j) = M \cdot C(i,j) \quad (8)$$

where $M$ is a $3 \times 3$ matrix representing the linear transformation of the three color channels. Therefore, Eq.8 can substitute step 4. If $p_s$ and $p_t$ are larger than $q$, this algorithm is a non-parametric method because no particular parametric format of color mapping function can be obtained, and step 4 plays a role of look-up table that maps the color levels from target to source image.

Secondly, manifold project only preserves local geometries but not the global geometries of the input manifolds. It does not prevent dissimilar colors in the original space from being neighbors in the embedding space. In other words, it only encourages similar colors in the original space to be neighbors in the embedding space. This could lead to some errors in determining correspondences between two manifolds in the embedding space. To solve this problem, we look for $k$ nearest neighbor pixels in the source image for each target pixel instead of only one nearest pixel. Then we select one of them as the corresponding pixel according to their SIFT features [12], i.e., a pixel with the most similar SIFT descriptor as this target pixel is chosen.

4. EXPERIMENTS

Color transfer for multiview image stitching requires geometric registration as preprocessing. Since registration is beyond the focus of this study, in the experiments we assume the images under study have been well geometrically aligned. The test images are generated as follows: Firstly, two panorama photographs are selected. Then, the first photograph is divided into two overlapped clips while the second into three clips. Finally, the left clips are designated as the source images, and the other clips are set as the target images by changing their color distributions. The performance is evaluated by comparing the transferred target image with its original clip via PSNR (Peak Signal to Noise Ratio), and visually perceiving the consistency of the stitching image.

In all experiments, the parameters $\alpha_1 = \alpha_2 = 1$, and the color space is $I^{(Y)}$. For the first photograph (Fig.1(a)), we transform the right clip into two target images with bad quality: one simulates low-illumination (Fig.1(c)) that is converted from the original clip by a linear transformation, and the other simulates color distortion (Fig.1(d)) that is based on a nonlinear transformation. Fig.1(e-f) are the results of color correction by Reinhard’s method [3], in which the overlap-
ping area is used to calculate the color transform function and then this function is used on the whole target image. Fig.1(g-h) are the results of our approach with $p_s = p_t = q = 3$. Fig.1(i-j) are the results with $p_s = p_t = 3, q = 1$. It is clear that the proposed method can provide very favorable performance that the transferred images are almost the same as the original clips, but the results of Reinhard’s method have obvious inconsistency between left and right parts of the stitched image. The nice performance of the proposed method is mainly due to that the local structure of color distribution is investigated to build the color transfer mapping, even the true color mapping between target and source images is nonlinear.

The second photograph (Fig.2(a)) is used to illustrate that manifold alignment is not limited to two data sets, and the multiview image stitching is not limited to two images either. The goal of color transfer in this experiment is to correct the colors of middle (Fig.2(b)) and right clips (Fig.2(c)) with the left clip as the source image. These two target clips are generated by linear and nonlinear color transformations respectively. Two methods can be used to complete this task. The first is sequential method that transfers the colors of the left clip into that of middle clip, and then transfers the colors of the corrected middle clip to the right clip. The second is one-off method that is a simple extension of two-dataset manifold alignment, in which the weighted adjacency matrix $W$ is constructed by the combination of the similarities between the pixels from the same image and the connections between the corresponding pixels from different images in overlapping areas. Fig.2(e-f) and (g-h) are the results of these two methods with $q = 3$ and $q = 1$, respectively. It can be observed that both of them have seamless stitching effect without visual artifacts. Compared to the one-off method, the sequential approach has low computation and storage cost, but the color transfer error may be accumulated if the number of clips for stitching is large.

5. CONCLUSIONS

Manifold alignment provides a more flexible and promising solution for color transfer than traditional approaches. In the proposed method, color mapping function is determined by the underlying manifolds of the images and their pairwise correspondences, so that the intrinsic structure of their color distribution can be preserved in the transfer procedure. Various algorithms based on manifold alignment can be developed to solve different color transfer problems. For example, some applications of color transfer, such as colorizing gray scale images and restoring the color of old photos, have not overlapped area between source and target images. In these cases, the correspondences might be obtained by finding similar regions between source and target images [13], or by determining the desired coordinates of some pixels in the embedding space. Extending the proposed method to other color transfer problems is our next task.

Fig. 2. Experimental results on the second photograph. (d-h): PSNR is for the middle and right clips.
6. REFERENCES


