COLOR CORRECTION AND IMAGE BLENDING FOR PANORAMA STITCHING VIA EXTREME POINT MATCHING OF LUMINANCE HISTOGRAM

Yaping Liu, Jian Yao†, Mi Zhang, Li Li

School of Remote Sensing and Information Engineering, Wuhan University, P.R. China
†Email: jian.yao@whu.edu.cn

ABSTRACT
In this paper, we propose a novel image blending method for panorama stitching from multiple images based on the extreme point matching of the luminance histogram. To reduce the brightness differences between images, an automatic adjustment of contrast is firstly applied on individual images based on their own color histograms. Both the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) with respect to luminance histograms are secondly utilized to obtain extreme points of the luminance histograms that are capable of best illustrating statistical characteristics of the image, which are robustly matched according to two predefined rules to construct the relationship between two involved images. Thirdly a pixel-wise color correction is employed in the whole image and a simple alpha transition correction strategy is used to ensure color continuity in the overlap regions between images. Finally a multi-band blending with optimal seamlines is adopted to create a seamless panorama. Experimental results show that our method is capable of blending street view panorama with high quality and faultless color continuity, which obviously out-performs the open-source software Enblend and the popularly used commercial software PTGui.

1. INTRODUCTION
Panoramic photography is a brand new technique that can obtain images with large field of view and present broader scenes in a more appealing way. It is quite popular among landscape, cityscape and architectural photographers both at home and abroad. The panoramic image is often acquired by mosaicking a series of images captured by several cameras with special lens, such as panoramic, fish-eye and wide-angle lens. The mosaicking procedure includes finding the overlap regions and spatial relationship between images, namely global alignment, and blending the images involved to get rid of artifacts and then to produce visual-appealing and informative wide-angle panorama. In this paper, we will concentrate on the image blending issue based on the assumption that the alignment and optimal seamline detection have been figured out already.

So far, there are three kinds of most popular blending strategies [1]. One is the transition smoothing method (also known as alpha blending and feathering) [2] that conceals seamlines by fusing overlap regions of two involved images. Another is optimal seamline detection method [3, 4] that seeks for the best placement of the seamlines which can minimize luminance or/and gradient changes along the path as much as possible. The last one is called hybrid method, which combines the above-mentioned two methods by fusing image regions around the optimal seamlines.

The most familiar blending algorithm is Poisson blending [5] that prefers to color contrast rather than intensity values which is just the same to human visual system. However, the Poisson blending is too complex that can be hardly used in the blending of multiple images with large size. To simplify the computational complexity of traditional blending methods, Chang et al. [6] proposed a low-complexity stitching algorithm which can even deal with the situation when the two photos have zooming and rotation actions. Li et al. [7] blended the overlap regions by using the average weighted method, which is quite simple and efficient. Nevertheless, it always results in blurring and ghosting artifacts that degrade the mosaic to some extent. Besides, the weighting coefficients in alpha blending largely depend on the determination of patch boundaries. In order to improve this, Xiong et al. [8, 9] created a single channel mask to provide weighting coefficients for blending images and therefore obtained the panorama, which did not need to compute boundaries of overlap regions and required low computational and memory costs. With the popularization of panoramic photography in mobile phone, the demands for time-consuming cost and image quality are increasing. Pulli et al. [10] studied the high-resolution panoramic image creation in the mobile system. They employed a coarse-to-fine method for image registration and an optimal seamline detection method to avoid the ghosting effect. Many researchers also dedicated themselves to the improvement of existing algorithms. Popovic et al. [11] discussed the Gaussian blending technique for the omni-directional view reconstruction and decreased the high light intensity variations by using the modified Gaussian algorithm. Juan et al. [12] transformed the multi-band blending (pyramid blending) method, which has been verified to be particular effective in panoramic image blending without obvious blurring and ghosting. Considering the presence of moving objects in scenes, Mills and Dudek [4] suggested heuristic seamline selection technique in gradient and intensity domains to decide which pixels were picked out from one of the two involved images, and then smoothly blended to get the final image. The blending algorithms are also applied in image-based object modeling [13] and texture atlas for 3D modeling [14].

To create a seamless panorama uniform in color and brightness from multiple overlapped images, we propose an efficient color correction via matching extreme points of luminance histograms to reduce color differences between images and apply a multi-band image blending with optimal seamlines in this paper. The detailed description of our proposed method is presented in Section 2. Experimental results on street view panoramic images as shown in Section 3 sufficiently demonstrate that our proposed method can create high-quality panoramas with faultless color and brightness continuity between images, which illustrates obvious superiority over the open-source software Enblend† and the popularly used commercial software PTGui‡. The conclusion is drawn in Section 4.

†http://enblend.sourceforge.net/
‡http://www.ptgui.com/
2. ALGORITHM

For an arbitrary image, both the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) with respect to luminance histograms describe the statistical characteristics and color distribution features in a more visually stimulating way, which are extremely informative and worth studying. In probability theory and statistics, the PDF is a function that describes the relative likelihood for a random variable to take on a given value while the CDF describes the probability that a real-valued random variable with a given probability distribution will be found to have a value less than or equal to this random variable. In this paper, the random variables in both PDF and CDF denotes the intensities of pixels in an 8-bit grayscale image. The main idea of our proposed method is that the PDF and CDF of the overlap regions in the first image should be approximately equal to those in the second image overlapped with the first image with respect to their curve trends. We match the extreme points selected from the two PDFs (i.e., luminance histograms) and then interpolate the intensities of pixels that fall in between the adjacent two extreme points. In addition, the PDFs play an important role in the matching of extreme points when the number of existing extreme points that we find in the PDFs is not sufficient. The intensities of all the pixels in the two images are modified afterwards, not only for the pixels in the overlap regions, but also in the non-overlap regions. Subsequently, the multi-band blending with optimal seamlines is implemented to produce a visual appealing and informative panoramic image.

2.1. Automatic Contrast Adjustment

Image blending sometimes fails when there are large intensity differences in color space between images, which can be caused by either natural illumination changes or different exposure modes. In order to make sure that multiple images have the similar contrast, which can produce satisfactory blending results, multiple channels of individual images are automatically adjusted in contrast. The histograms of a color image are calculated firstly in each channel, namely, R, G, and B channels, respectively. Let I be a single-channel image with an image size of $W \times H$ and $Z = \{I(i)\}_{i=1}^N$ be a set of one dimensional sorted intensities of all valid pixels in I in the ascending order where $N$ denotes the total number of valid pixels in I and $I(i)$ represents the intensity of the $i$-th pixel in I. The minimal intensity $I_{\min}$ and the maximal intensity $I_{\max}$ of I are defined as:

$$I_{\min} = I([N \times (1-c\%)]), \quad I_{\max} = I([N \times (1-\epsilon\%)])$$

where $[x]$ denotes the upper integer of a real value $x$ and c is a small percentage value in the range of $(0, 0.5)$ ($\epsilon = 0.1$ was empirically used in this paper), which can be used to skip over a part of the residual maximal and minimal intensities due to the fact that these pixels may be caused by noises and information lacking in most cases. The minimal and maximal intensity values of the R, G and B channels of a color image are denoted as $I_{\minR}, I_{\minG}, I_{\minB}$ and $I_{\maxR}, I_{\maxG}, I_{\maxB}$, respectively. Therefore, any intensity $I$ of the R, G and B channels of a color image will be linearly modified as:

$$I' = \begin{cases} 
0, & \text{if } I \leq V_{\min} \\
255 \times (L - V_{\min})/(V_{\max} - V_{\min}), & \text{if } V_{\min} < I < V_{\max} \\
255, & \text{if } I \geq V_{\max}.
\end{cases}$$

In the same way, all the images for creating a panorama will be automatically adjusted in contrast, which will reduce the brightness differences between images.

2.2. Finding Extreme Points

Only valid pixels in the overlap regions between images are considered for statistic analysis, which are obtained from their mask images with the same size precisely aligned beforehand in a common coordinate system. The overlap regions in two images are represented as A and B, respectively. To make a better description of the information hidden behind the image, we convert A and B from the RGB color space to the HSI color space. Then the PDFs and CDFs are calculated in the intensity channels for A and B, which are denoted as PDF$_A$, PDF$_B$, CDF$_A$, and CDF$_B$, respectively.

To robustly find extreme points in PDF$_A$ and PDF$_B$, these two PDFs are smoothed first by a Gaussian function to suppress possible noisy pixels. The initial local extreme points can be easily obtained from the smoothed PDF$_A$ and PDF$_B$. In an ideal situation, the extreme points should be uniformly distributed in the color space. However, most of the extreme points are relatively centralized in some cases, which will lead to information redundancy due to that multiple extreme points are selected out to represent the similar image statistical information. To avoid the situation mentioned above, we further check out all the initial extreme points by local window suppression. Let $\{I_A^i\}_{i=1}^N$ be the intensities of $K$ extreme points $\{P_A^i\}_{i=1}^N$ in PDF$_A$, which are sorted in the ascending order. Given an extreme point $P_A^i$, we generate a neighborhood range $[L_A - w, L_A + w]$ centered on the corresponding intensity $I_A^i$ with the size of $2w + 1$. If there are more than one extreme points located in this neighborhood range, the extreme point with the highest frequency in PDF$_A$ will be retained while $P_A^i$ will be eliminated. All the initial extreme points are examined in this way and the retained extreme points are used for following matching. The final extreme points of PDF$_A$ and PDF$_B$ are represented as $\{P_A^i\}_{i=1}^{N_A}$ and $\{P_B^j\}_{j=1}^{N_B}$, where $N_A$ and $N_B$ are the numbers of extreme points in PDF$_A$ and PDF$_B$, respectively.

2.3. Matching Extreme Points

The extreme points in the luminance histograms sufficiently reflect image statistical characteristics in brightness. To efficiently adjust the color of two adjacent images with an overlap region, one way is to correct the color of these two images to ensure that the extreme points in the luminance histograms (or PDFs) computed from the overlap regions of the color-corrected images are the same. To reliably match these extreme points $\{P_A^i\}_{i=1}^{N_A}$ and $\{P_B^j\}_{j=1}^{N_B}$ in PDF$_A$ and PDF$_B$ of two adjacent images A and B, we define a cost function to measure the matching similarity of two extreme points $P_A^i$ and $P_B^j$ as:

$$Cost(P_A^i, P_B^j) = \frac{L_A^i + L_B^j}{2L_{\max}} \times \frac{\min(L_A^i, L_B^j)}{\max(L_A^i, L_B^j)} \times \frac{\max(C_A^i, C_B^j) - C_{\min}}{C_{\max} - C_{\min}},$$

where $L_A^i$ and $L_B^j$ are the intensities of $P_A^i$ and $P_B^j$, respectively, and $L_{\max}$ is the maximal intensity of all the extreme points in PDF$_A$ and PDF$_B$. $C_A^i$ and $C_B^j$ denote the cumulative values of $P_A^i$ and $P_B^j$ in CDF$_A$ and CDF$_B$, respectively. $C_{\max}$ and $C_{\min}$ represent the maximal and minimal cumulative values of all the extreme points in CDF$_A$ and CDF$_B$, respectively. The above cost function judges the two extreme points from the view of both PDF and CDF. The higher the function value is, the more likely the two points are matched. Based on this cost definition, a $N_A \times N_B$ matching cost matrix $M = [M_{ij}]_{N_A \times N_B}$ is introduced. If the cost function $Cost(P_A^i, P_B^j)$ is smaller than a predefined threshold $\theta$, the value
of \( M_i \) in the matching cost matrix \( M \) is directly set to 0, otherwise \( M_{ij} = \text{Cost}(P_i, P_j) \). For \( P_A \), we can obtain \( N_B \) matching costs \( \{M_{ij}\}_{j=1}^{N_B} \) with respect to \( N_B \) extreme points in \( \text{PDF}_B \). Some extreme point \( P_B^k \) in \( \text{PDF}_B \) is considered as the optimally matched point of \( P_A \) only when \( M_{ik} \) reaches the maximal value among \( \{M_{ij}\}_{j=1}^{N_B} \). In order to have a more stable matching result, we further examine \( P_A^k \) and \( P_B^k \) according to:

\[
|C_{max} - C_{min}| > |C_A(L_A + \delta) - C_A(L_A - \delta)| + \theta,
\]

where \( \delta \) is a given threshold to generate a range centered with \( L_A \) and \( L_B \). Hence, \( C_A(L_A + \delta) \) and \( C_A(L_A - \delta) \) denote the corresponding cumulative values in \( \text{CDF}_A \) with the intensities of \( (L_A + \delta) \) and \( (L_A - \delta) \), respectively. The same meanings are for \( C_B(L_B + \delta) \) and \( C_B(L_B - \delta) \). \( \theta \) is an empirical value to balance the equation. We use this formula to further make sure that \( P_A^k \) and \( P_B^k \) are matched with respect to CDFs. Only when \( P_A^k \) and \( P_B^k \) both satisfy the above mentioned two conditions, we regard \( P_A^k \) and \( P_B^k \) as an optimal match. To obtain a more reliable matching between extreme points in \( \text{PDF}_A \) and \( \text{PDF}_B \), we start to seek the optimal match of the extreme point with the maximal matching cost value in \( M \). If an optimal match is found, these two matched extreme points are removed and the corresponding cost matrix \( M \) is accordingly updated. In this iterative way, a set of extreme point matches are found.

If the number of extreme point matches is less than a predefined threshold, we will search more matching pairs with the help of \( \text{CDF}_A \) and \( \text{CDF}_B \). We collect the \( H \) uniformly distributed points \( \{C_{A1}^{(i)}\}_{i=1}^{H} \) and \( \{C_{B1}^{(i)}\}_{i=1}^{H} \) from \( \text{CDF}_A \) and \( \text{CDF}_B \), respectively. Different from the PDF, the collected points in the CDF are relatively with less image information. So the redundant points in CDF should be avoided in the following sampling. The same number of sampling points in \( \text{CDF}_A \) and \( \text{CDF}_B \) are uniformly selected in accordance with the cumulative density value. Let \( L_A^i \) and \( L_B^i \) be the intensities of the \( i \)-th sampling points \( C_A^i \) and \( C_B^i \) in \( \text{CDF}_A \) and \( \text{CDF}_B \), respectively. If there are no extreme point match found in the ranges \( L_A^i - \varepsilon, L_A^i + \varepsilon \) and \( L_B^i - \varepsilon, L_B^i + \varepsilon \) where \( \varepsilon \) is a given threshold in advances, the current sampling points \( C_A^i \) and \( C_B^i \) are added into the matching set as a new point match.

### 2.4. Color Correction and Multi-Band Blending

The extracted matching points in the overlap regions are then applied to correct the intensities of the two adjacent images, including the pixels in non-overlap regions. Let \( \{Q_A^i\}_{i=1}^{N} \) and \( \{Q_B^i\}_{i=1}^{N} \) be the final matching points in the overlap regions \( A \) and \( B \) with the point number of \( N \). Based on the matching results, the intensity of \( Q_A^i \) and \( Q_B^i \) is modified to \( (L_A^i + L_B^i)/2 \) where \( L_A^i \) and \( L_B^i \) denote the intensities of \( i \)-th matching points \( Q_A^i \) and \( Q_B^i \), respectively. In addition, the intensity \( L_B^i \) of an unmatched point \( P_A^i \) in \( A \) is linearly corrected as:

\[
\tilde{L}_B^i = \frac{L_A^i - L_A^i}{L_A^i - L_A^i} + \frac{L_A^i - L_A^i}{L_A^i - L_A^i},
\]

where \( L_A^i \in (L_A^i, L_A^i) \), \( L_A^i \) and \( L_A^i \) denote the intensities of two matching points in \( A \) that are closest to \( P_A^i \), and \( \tilde{L}_A^i \) and \( \tilde{L}_A^i \) are the corresponding corrected intensities. In order to produce a smooth and gradual transition from non-overlap regions to overlap ones, the alpha correction method is conducted as:

\[
L = (1 - \alpha)L_A + \alpha \tilde{L}_A,
\]

where \( \alpha \) is a function that related to the distance between the pixel and the seamline, which ranges between 0 and 1. The smaller the distance is, the larger the \( \alpha \) is. \( L_A \) is the intensity of a certain original pixel in non-overlap regions while \( L_A \) is the corrected intensity of the corresponding pixel based on the above mentioned correction method. All the pixels in another image will be processed in the same way.

The Laplacian pyramid blending [15] is applied in the multi-band blending. One modified contribution of our method here is that the mask image \( W \) in each level pyramid fusion is constructed by the weight both related to the valid regions and seamlines of the corresponding input image based on the following pixel-by-pixel multiplication:

\[
W = W_M \cdot W_S,
\]

where \( W_M \) is the weight in accordance with the valid regions of the image. To be specific, for a certain point, if the distance \( d \) to the boundary of the valid regions is larger than the threshold \( \mu \), the label is 1, otherwise is \( 1/d \). The \( W_S \) is the weight map in accordance with seamlines, that is to say, if the point is located in the stitching region of some image, the label is 1, otherwise is 0.

### 3. EXPERIMENTAL RESULTS

To evaluate the effectiveness and superiority of our proposed method, we compared our blending results with those of the open-source software Enblend and the commercial software PTGui, which were both popularly used in the field of panoramic photography. The street view images were obtained by a mobile vehicle platform with 6 cameras with wide-angle lens and precisely aligned afterwards. In total, the color correction was firstly applied on 6 image pairs, consisting of 5 image pairs from 5 adjacent horizontal camera view pairs and 1 image pair between the top camera view and the horizontal camera ones, and the multi-band blending was then applied on these 6 corrected images for creating a whole panorama. Our method was implemented with C++ under Windows and tested in a computer with an Intel Core i7-4770 at 3.4GHz with 16GB RAM.

Fig. 1 shows the visual comparison between Enblend and our proposed method on three street view panoramas taken from different scenes with buildings, vegetations, cars and pedestrians under different illumination conditions. It is quite obvious that the color of...
Fig. 1. The visual comparison between blended panoramas by Enblend (Top) and our proposed method (Bottom).

Fig. 2. The visual comparison between blended panoramas by PTGui (Top) and our proposed method (Bottom).

our blending results are much more natural when compared with Enblend, especially in the sky, due to that all the pixels were corrected according to the matched extreme points by interpolation in our proposed method instead of only pixels in the overlap regions, as processed in Enblend. The details of roads and buildings are preserved to the greatest extent as well as color continuity. In addition to superiorities mentioned above, our blending results are more faultless in local details than Enblend. For example, the roofs in the top-right panorama created by Enblend as shown in Fig. 1 are blurred for unknown reason. The visual comparison between blended panoramas by PTGui and our proposed method are shown in Fig. 2. The blended panoramas created by PTGui are visually similar to ours except for a few flaws in the sky. In addition, we observed that the blended panoramas created by our proposed method have more seamless blending results in overlap regions mainly due to the use of reliable alignment and optimal seamlines detected via graph cuts.

The weight map $W_S$ with respect to semline regions of the left panorama shown in Fig. 1 is presented in Fig. 3(Left) as well as the corresponding blended panorama without the use of $W_S$ in Fig. 3(Right). Different colors as shown in Fig. 3(Left) correspond to six different involved images with the optimal seamlines found in advances via graph cuts. Through visual comparison, we can figure out that $W_S$ plays a crucial role in balancing color intensity between the adjacent two involved images and smoothing seamlines to produce a good-quality and natural looking panorama. To have a better understanding of our proposed method, the final extreme points of luminance histograms of two adjacent involved images in the left panorama shown in Fig. 1 is plotted in Fig. 4 in which those 21 sets of matched points were then employed to correct the intensities of the whole image. Due to the limit of pages, more experimental results and analysis are presented at http://cvrs.whu.edu.cn/projects/panoBlending/.

4. CONCLUSION

In this paper, we have developed a novel image blending method for panorama stitching based on extreme point matching of luminance histograms of two involved images via probability density functions and cumulative distribution functions. To reduce the brightness differences between images, an automatic contrast adjustment is firstly applied in each channel of original images. The extreme points of luminance histograms were secondly selected out according to a series of rules and regarded as feature points for matching of two luminance histograms. Then, all the pixels in the image, not only in the overlap regions but also in the non-overlap ones, were corrected successively with the guidance of those matched extreme points. The alpha weights of image regions around seamlines were introduced in the multi-band blending to further improve blending results with respect to blurring and color incontinuity. Comparative experiments were conducted to verify the effectiveness and superiority of our proposed method over Enblend and PTGui.
5. REFERENCES


