A Robust Region-wise Colour Correction Method for Stereo Matching

Qing Ran , Wenjing Zhao , Jieqing Feng *

State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou, 310058, P.R. China
* Author for correspondence (jqfeng@cad.zju.edu.cn)

Abstract

Significant colour discrepancies between stereo images have severe impacts on stereo matching algorithms, which depend on colour similarity. To address this problem, a region-wise colour correction method is proposed in this paper. First, the source image, which is to be corrected, is segmented into a set of regions using the mean-shift method. SIFT features are extracted from both the source image and the reference image, and then matched. The matched SIFT pairs are refined by exploiting epipolar geometry constraint. Based on the segmentation result and the refined SIFT matches, the regional correspondences between two stereo images are estimated. Then, each matched region pair is used to compute a local colour correction function. A set of colour weight maps is calculated for these functions. Last, to alleviate the colour transformation discontinuities along the region boundaries and facilitate a smooth colour correction globally, the corrected colour is obtained by combining the colour correction functions using the colour weighting maps. We apply the proposed local colour correction method to stereoscopic images before conducting stereo matching. The results indicate that the proposed algorithm can effectively and robustly alleviate the colour discrepancies, and improve the accuracy of stereo matching.

1 Introduction

Stereo matching is one of the key steps for estimating depth information in stereo vision. Numerous stereo matching algorithms have been proposed [20, 21]. Each algorithm requires a matching cost to measure the similarity of image locations. These matching costs are based on the assumption, that pixels corresponding to the same scene point have similar or ideally the same value in the stereo images [12]. However, this assumption is difficult to be satisfied in practical applications. Various factors, such as slightly different camera settings, varying lighting conditions or non-Lambertian surfaces of objects in a scene, etc, inevitably cause radiometric differences between stereo views.

The simplest matching cost assume same or similar intensities at the corresponding scene points, and more complicated costs, explicitly or implicitly, allows a certain degree of radiometric differences. Recent studies for stereo matching focus on dealing with the robustness of matching cost under various conditions. Bleyer et al. [2] and Chambon [3] evaluated different colour spaces via the matching costs, i.e. absolute difference (AD) and squared difference (SD), for improving the accuracy of stereo matching. Census transform [31], the approximation of mutual information [32] and other matching costs [11, 16], have been proposed to take radiometric differences into account. However, none of the current prevalent matching costs can handle strong radiometric differences successfully, according to the survey by Hirschmuller and Scharstein [12]. The stereo matching algorithms may suffer from these colour discrepancies, and leads to the matching ambiguities. Our aim is to correct significant colour discrepancies and improve the colour consistency between two stereo images, to improve the accuracy of stereo matching finally.

A typical method for alleviating colour discrepancies is to calibrate the discrepancies by employing a professional colour calibration pattern before image acquisition [8, 13]. Under a common colour model, identical colours have the same value. Although this solution is often workable, the operations of this colour calibration scheme are inconvenient and infeasible in some cases. Once the illumination is altered, the entire tedious procedure must be repeated. Such colour calibration works well for indoor scenes. However, in cases which images are retrieved from the Internet [9], calibrating colours in advance is impossible.

Alternative approaches involve colour correction methods, such as colour transfer, colour mapping, colour correction, and other. The differences among these methods are not technically distinct [29]. The purpose of these techniques is to establish the photometric correspondence between the colours of one image and the colours of another image. The image that is adopted as a reference is referred to as the reference image, whereas the image to be corrected is referred to as the source image. Thus, the colour correction is a process of transforming the colours of the source image so that its colour appearance is consistent with that of the reference image.

Colour correction is an effective way for balancing colour discrepancies; it has been extensively adopted in many computer vision and computer graphic applications. Although many colour correction methods exist, few methods have been specially extended to stereo vision applications to address the colour inconsistency problem. Compared with the general cases, stereo images represent the same scene taken from two adjacent shooting angles. The majority of the pixels in one image have corresponding pixels in the other image, which is a strong prerequisite for colour correction.

In this paper, a robust region-wise colour correction approach, which produces consistent colour appearances between stereo images, is proposed. The main idea is to compensate the colour discrepancies region by region. Instead of addressing the complex problem of consistent segmentations between two input images [23, 26], regional correspondences are established by mapping the segmentation of the source image to the reference image. For each pair of corresponding regions, a local colour correction function is computed by collecting the local colour statistics. To eliminate colour transformation discontinuities across regions, a weighted combination scheme of all the local colour correction functions is utilized to generate the corrected source image. Both quantitative and qualitative comparison and analyses prove that our algorithm can effectively alleviate the colour discrepancies between two stereo images and improve the accuracy of stereo matching algorithms.

The following points highlight the contributions of this paper:

- A robust region-wise colour correction approach is proposed to solve the colour inconsistency problem between two stereo images and to meet the requirements of stereo matching.
- A practical regional correspondences estimation is proposed. Based on image segmentation and matched features, the corresponding region in the reference image can be approximately estimated for each region in the segmented source image.
- A weighted colour correction scheme is proposed. The local colour corrections are weighted to deal with the colour discontinuities across the regions in the corrected image.

(a) (b) (c) (d)

Figure 1: *Example of "Art". (a) A reference image; (b) A source image; (c) The corrected source image by using the global algorithm [18]; (d) The corrected source image by using our algorithm. Note that (c) has considered the colour transformations of all the regions as a whole. Clearly, the colour details of the statue region have been lost.*

The remainder of this paper is organized as follows; Section 2 describes the related works on colour correction. Section 3 details the proposed robust region-wise colour correction method. Implementation results and comprehensive comparisons are provided in Section 4. The conclusions and future works are presented in Section 5.

2 Related Work

According to a comprehensive evaluation by Xu and Mulligan [29], colour correction approaches can be categorized into two classes - non-parametric methods and parametric methods. Non-parametric methods [25, 30] focus on the construction of a look-up table to directly record the mapping of a complete range of colour values. The look-up table is usually computed from the correspondence between image features or pixels in the overlapped area of two images; this process is prone to be influenced by noise or outliers. Due to the problem of robustness, non-parametric methods are generally more complex and more unpredictable than parametric methods. Therefore, we follow the parametric methods to discuss the colour correction problem.

The parametric methods assume that the colour mapping function is based on the statistics of the colour distribution in the images. They are classified into two types: global approaches and local approaches. Colour transfer is a global approach, which was proposed by Reinhard et. al [18]. In this method, the images are transformed to $l\alpha\beta$ colour space, and the colour correction function is defined as:

$$\bar{c}^{src}(i,j) = \mu^{ref} + \frac{\sigma^{ref}}{\sigma^{src}}(c^{src}(i,j) - \mu^{src}) \tag{1}$$

where $\bar{c}^{src}(i,j)$ and $c^{src}(i,j)$ are the corrected value and original value, respectively, at the pixel (i,j) of the source image; the variables μ and σ represent the mean and standard deviation, respectively, of the global colour distribution; and the superscripts ref and src indicate the reference image and the source image, respectively. Fig. 1(c) shows an example using Reinhard's global method. Without simultaneous consideration of colour statistical contribution and spatial information, this method tends to produce unsatisfying results with unnatural scene details and colour distribution, e.g. the loss of details of the statue in Fig. 1(c). The colour correction function proposed by Reinhard et al. [18], which is calculated by considering different colour regions as a whole in fact, is inadequate for discerning different colour statistics for complex scenes. Therefore, some local approaches [17, 23, 26] have been proposed to improve the colour correction results.

The method proposed by Tai et al. [23] was recommended as the first option among local colour correction methods in the evaluation proposed by Xu et al. [29]. The input images are probabilistically segmented using Gaussian Mixture Models (GMMs) based on their soft segmentation algorithm [24]. Then, correspondence was established to map each Gaussian component in the source image to certain Gaussian components in the reference image. Finally, the corrected colour of each pixel was obtained as a combination of the intermediate results that were produced by each Gaussian component pair according to each pixel's probability. However, the assumption of GMMs will restrict the flexibility of segmentation. Small regions may not be properly segmented if they don't contain a sufficient number of pixels to form a distinctive Gaussian component. Furthermore, the expectation maximization segmentation is computationally demanding. Other colour correction methods include computing the brightness transfer from 2D joint histograms [15] and a 2D tensor voting scheme [14]. For more detailed descriptions of different colour correction methods, readers can refer to the comprehensive survey [6, 29].

Stereo matching estimates the pixel correspondences in two stereo images, and then compute the disparity of the scene. The colour similarity is usually adopted as matching costs. Unfortunately, colours are often not well matched due to several inevitable factors [12], such as non-Lambertian surfaces, non-diffused illumination and etc. To reduce matching ambiguities caused by colour discrepancies, Wang et al. [27] proposed a local colour correction method. In their method, the source image was segmented into several regions and the SIFT features were matched between stereo images using an optical-flow based matching algorithm. However, the optical flow based matching method is regarded to be sensitive to colour discrepancy in the images [1]. Then, the colour of each pixel in a segmented region was corrected by adding a constant value, which is the average of colour discrepancies of the matched SIFT features in that region. Thus, the local colour correction function in their method can be considered to be constant, which does not sufficiently sample colour informations in that region, and is inadequate for the complex scenes. Furthermore, a naive region wise colour correction may lead to colour discontinuities across the regions. As an improvement, we will employ all pixels in a segmented region to compute a weighted linear local colour correction function to address these problems.

3 Region-wise Colour Correction for Stereoscopic images

We assume that input stereo images satisfy the epipolar geometry constraint as a prerequisite. Without a loss of generality, the left image in stereo images is assumed to be the reference image - denoted as I_{ref}, and the right image is assumed to be the source image I_{src}.

The proposed approach consists of three consecutive steps. First, regional correspondences are estimated based on the image segmentation and spatial relationship derived from matched features. Second, local colour transfer is applied to each region pair

via a local colour correction function. Third, to eliminate the colour transformation discontinuities across the regions, a colour weight map is computed for every region. The final corrected colour of each pixel is calculated using the local colour functions weighted by the colour weight maps. Details of each stage are described in detail below.

3.1 Coarse Regional Correspondences

In this section, we present a method to establish the coarse regional correspondences between input stereo images. The source image is segmented into regions. Each region in the source image is projected to the reference image by a local spatial transformation, which is derived from matched SIFT features within this region. In this way, the coarse regional correspondences are established.

First, the mean-shift algorithm [5] is adopted to segment the source image into a set of regions $\{{R_1}^{src}, {R_2}^{src}, ..., {R_N}^{src}\}$. Each segmented region consists of several similar colours. In our method, the segmentation should be carefully tuned for the region size, colour discontinuity characteristics, etc. By carefully adjusting the parameters, we could control the size and the number of regions, and the boundaries of segmented regions will be consistent with the colour discontinuities in the source image. In this way, each segmented region can be roughly regarded as a single plane, which is meaningful to estimate the spatial transformations between two stereo images later. Compared with other segmentation methods [7, 22], the mean-shift algorithm is more robust to noises and has lower computation complexity considering the large size of the test images.

After segmentation, the SIFT features are extracted and then matched between the two stereo images. The invariance of the SIFT features to viewpoint variation and illumination change, guarantees that we can obtain robust matches between two stereo images even with significant colour discrepancies. Fig. 2(a) shows the result of initial SIFT matches. Then, the initial matches are refined with the epipolar geometry constraint [10] as follows:

$$Q_{ref} F Q_{src}^T = 0 \tag{2}$$

where F is the fundamental matrix between two stereo images, and Q_{ref}, Q_{src} are the matched SIFT features, respectively. The matching SIFT pairs that do not satisfy Eqn. (2) will be considered as outliers. Considering robustness and accuracy, we used the Sampson distance, rather than the algebraic distance, in both fundamental matrix estimation and outlier rejection in Eqn. (2). The removed outliers shown in Fig. 2(b), obviously, these features are matched inaccurately.

(a) (b)

Figure 2: *Illustration of SIFT matches. (a) Initial SIFT matches; 2321 matches. (b) 105 mismatched SIFT features matches (outliers) are detected by Eqn.(2). 2216 corrected SIFT matches are preserved.*

Let ${R_k}^{src}$ be the k-th segmented region in the source image I_{src}, and ${R_k}^{ref}$ be the corresponding region in the reference image I_{ref} to be determined. We assume that ${R_k}^{src}$ and ${R_k}^{ref}$ are two projections of the same scene plane S_k in two images from two different viewpoints. Therefore, the relationship between R_k^{src} and R_k^{ref} can be described as a homography matrix induced by S_k [10]. Then, ${R_k}^{ref}$ is determined by projecting all the pixels in ${R_k}^{src}$ into the reference image via the homography matrix, which is formulated as follows:

$${R_k}^{ref} = P_k({R_k}^{src}) \tag{3}$$

$P_k(\cdot)$ can be estimated through a set of SIFT feature pairs $\{{q_i}^{src}, {q_i}^{ref}\}$ between ${R_k}^{src}$ and ${R_k}^{ref}$. We calculated a numerical solution of $P_k(\cdot)$ by minimizing the back-projection error in Eqn. (4)

$$\sum_{i=1}^{n} \|{q_i}^{ref} - P_k({q_i}^{src})\|^2 \tag{4}$$

where n is the number of matched SIFT feature pairs. There are three cases in a homography matrix estimation:

1) $n >= 4$. Based on the RANSAC method, we try many different random subsets of the matches to estimate a set of $P_k(\cdot)$s by minimizing the back-projection error above. For each subset, the estimated homography matrix $P(\cdot)$ is assessed by the number of inliers. The best one is then used again to estimate the final homography matrix for all the inliers by minimizing the back-projection error above.

2) $0 < n < 4$. In this case, there is no reasonable numerical solution to minimize Eqn. (4). As a feasible solution, the homography matrix degrades to a 2D translation $(\Delta x, \Delta y)$. In practise, these regions always contain small number of pixels which can be completely considered as the frontal planes, so that the above simplification is reasonable and feasible.

3) $n = 0$. We search the nearest matched SIFT feature with minimum Euclidean distance to this unmatched region's centre, which is the centre of the axis-aligned bounding box of this region. Then, the homography matrix of the region which contains the nearest matched SIFT feature will be applied to this unmatched region.

After homography matrix estimations, a set of regional correspondences $\{{R_k}^{src}, {R_k}^{ref}\}$ is obtained by applying Eqn. (3). In this way, most of pixels in the reference image are assigned with labels to be related to their corresponding regions in the source image. Those pixels assigned more than one label or no label, are considered to be in the occluded regions.

Fig. 3 shows the consequent regional correspondences between two stereo images. In order to demonstrate the reliability of the estimation of regional correspondences, the back-projected segmentation of the source image (Fig. 3(c)) is compared with its

(a) (b) (c) (d)

Figure 3: *Regional correspondences for the example "Art". (a) The mean-shift segmentation of the source image. (b) The segmentation of the source image (a) is projected onto the reference image via estimated homography matrices, where the black areas are the unmatched areas in the reference image. (c) The estimated segmentation of the reference image is backwardly projected onto the source image. (d) The illustration of difference between (a) and (c), where the blue pixels, the red pixels and the black pixels represent the correct correspondences, the incorrect correspondences and the unmatched areas respectively. In our experiments, we found that the unmatched areas were almost overlapped with the occluded areas of the source image in subsequent stereo matching.*

mean-shift segmentation (Fig. 3(a)). The comparison result is illustrated in Fig. 3(d). Pixels with the same segmented labels are shown in blue, whereas the different segmented labels are shown in red. Note that the incorrect correspondences are mainly located along the boundaries of regions.

3.2 Local Colour Correction Functions

In the previous step, the regional correspondences between two stereo images were established. For a region pair $\{{R_k}^{src}, {R_k}^{ref}\}$, a local colour correction function is derived:

$$\bar{c}_k{}^{src}(i,j) = {\mu_k}^{ref} + \frac{{\sigma_k}^{ref}}{{\sigma_k}^{src}}({c_k}^{src}(i,j) - {\mu_k}^{src}) \tag{5}$$

${c_k}^{src}(i,j)$ and $\bar{c}_k{}^{src}(i,j)$ denote the colour of pixel (i,j) in the region ${R_k}^{src}$ before correction and after correction respectively. All pixels in ${R_k}^{src}$ are collected to calculate the mean ${\mu_k}^{src}$ and the standard deviation ${\sigma_k}^{src}$ locally. It is same for ${\mu_k}^{ref}$ and ${\sigma_k}^{ref}$ in the corresponding region ${R_k}^{ref}$. Through this correction function, the colour discrepancy between two regions is compensated locally in a linear manner.

The colour correction is performed in the $l\alpha\beta$ colour space. $l\alpha\beta$ colour space is one type of luminance-chrominance system which is very close to human perception, and it minimizes correlation among its colour channels for natural scenes so that can reduce the artefacts caused by channel correlation potentially [18]. For convenience, the colour correction functions can be re-formulated as a function of several parameters as follows:

$$\bar{c}^{src}(i,j) = f(\mu^{ref}, \mu^{src}, \frac{\sigma^{ref}}{\sigma^{src}}, c^{src}(i,j)) \tag{6}$$

In global approaches, the first three parameters in $f(\cdot)$ are constant for the whole image. Due to different surface reflection properties and non-uniform illumination, the colour correction functions should be different for each region. We compare $(\mu^{ref}, \mu^{src}, \frac{\sigma^{ref}}{\sigma^{src}})$ of the l-channel in each segment region with the global constant value as an example. The comparison result is shown in Fig. 4. Obviously, the parameters are not constant for all segmented regions. The global parameters can only provide a rough approximation of the local parameters. In the proposed approach, the parameters are constant only for each segmented region.

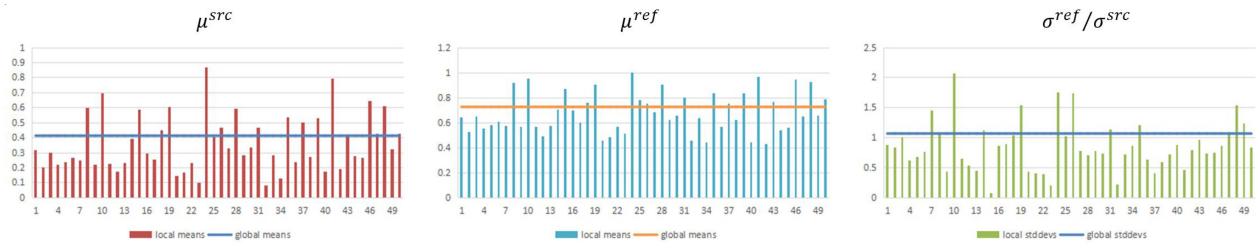


Figure 4: *Colour correction parameters comparison between the global approach and the proposed local approach for "Art". Due to limited text scope, only the first 50 regions are shown. The global parameters, as shown by the straight line, remain constant. There are large difference between global values and local values.*

3.3 Weighted Colour Correction

Due to occlusion and perspective distortion between two stereo images, the estimations of spatial relationship from feature matches are coarse and the regional correspondences may be not precise (Fig. 3(d)). As a consequence, the local colour correction result will be discontinuous in a naive region-by-region manner, that is the corrected colour of each pixel is computed by the local colour correction function of the region to which the pixel belongs, as shown in Fig. 5(a). To eliminate the colour transformation discontinuities and facilitate a smooth colour correction result, a weighted colour correction scheme is presented in this section.

The basic idea is to average the local colour correction functions weighted by colour and spatial distance, which is an improved Colour Influence Map (CIM). The CIM used by Oliveira et al. [17] is a colour weight mask for each region that measures the similarity between each colour pixel (i, j) and the mean colour of the region. Pixels with a colour similar to the mean colour of the current region are assigned large weights. Without consideration of spatial distance, the pixel (i, j) will be influenced by all regions. Thus, the weighted average will produce an over-smoothed result. Obviously, the regions far from the pixel (i, j) will have almost no influence on the colour of the pixel (i, j). In our method, CIM is improved by adding a spatial factors as follows:

$$\mathrm{CIM}_k(i,j) = exp(-\frac{c(i,j) - \mu_k}{2\alpha^2}) \times W_k(i,j) \tag{7}$$

where

$$W_k(i,j) = exp(-\frac{dist((i,j), R_k)}{2\beta^2}) \times F_k(i,j) \tag{8}$$

$W_k(i, j)$ is the spatial weight function, where the distance between the pixel (i, j) and the region R_k is defined as $dist((i, j), R_k) = min\{dist((i, j), r), r \in R_k\}$. To reduce the computation cost, only the neighbour regions of R_k are taken into consideration. $F_k(i, j)$ is the mask, where if the pixel (i, j) is in the current region R_k or in the adjacent regions of R_k, the $f_k(i, j)$ is 1; otherwise, 0. In Eqn. (7), α is the range bandwidth used to measure the colour similarity and β is the spatial range bandwidth used to measure the distance. Fig. 5(b)- 5(d) give an illustration of the improved CIMs for two selected regions in the source image.

(a) (b) (c) (d)

Figure 5: *Illustration of improved colour influence maps. (a) The colour correction result without using the weighted colour correction. (b) Two selected regions, shown in blue and green, respectively. (c) The colour influence map for region "blue". (d) the colour influence map for region "green". In (c) and (d), the brighter the pixel is, the larger the influence of this region on the pixel is. The black colour pixels represent no influence.*

Finally, the corrected colour $\bar{c}^{src}(i, j)$ of a pixel (i, j) in the source image is calculated by averaging locally corrected colours, weighted by the improved CIMs, as follows:

$$\bar{c}^{src}(i,j) = \frac{\sum_{k=1}^{n} \mathrm{CIM}_k(i,j) \times f_k(\mu^{src}, \mu^{ref}, \frac{\sigma^{ref}}{\sigma^{src}})}{\sum_{k=1}^{n} \mathrm{CIM}_k(i,j)} \tag{9}$$

where n is the number of the segmented regions of the source image, $f_k(\mu^{src}, \mu^{ref}, \frac{\sigma^{ref}}{\sigma^{src}})$ is the local colour correction of the k-th region, and $\mathrm{CIM}_k(i, j)$ is the weight of pixel (i, j) in the colour influence map of the k-th region.

4 Experiments Results, Evaluations and Discussion

The proposed algorithm is tested on the Middlebury benchmark dataset [19]. 30 pairs of stereo images are employed in our experiments. The two images in each stereo pair are taken under different light conditions or with different exposure settings., which contain significant colour discrepancies. The left image is chosen as the reference image, and the right image is chosen as the source image. An image in the dataset can be regarded as the ground truth of the source image, which is taken under the same settings with the reference image. The ground truth disparity map for each pair of stereo images is used for later quantitative evaluation. All the experiments are executed on a computer with an Intel Core i5 CPU 3GHz and 4GB memory. The execution time for a pair of stereo images whose resolution is 1390×1110 is approximately 30 seconds. Most of the execution time is consumed on the mean-shift segmentation and SIFT feature matches extraction.

The mean shift parameters are carefully tuned. A large region may contain too many different colours and would result in a correction result similar to its global counterpart (Fig. 6(a)). In contrast, a small region will be sensitive to the inaccurate estimation of spatial transformations, which may lead to an inappropriate correction result (Fig. 6(b)). According to the analysis by Comaniciu and Meer [4], we set the parameters of the mean-shift algorithm to $\sigma_s = 25$, $\sigma_r = 9$, $M = 300$ for all the tested images. The parameters of the improved colour influence map computation include the colour range α and the spatial width β. Experimental results show that the values of α and β have small impacts on the colour correction results, as shown in Fig. 6(c). In our experiments, α is set to between 20 and 40; β is set to 10% of the image diagonal. These empirical values of the parameters can result in satisfied correction results for all the source images in the dataset.

We compare our method with three previous methods - the global approach [18], the optical-flow based local approach [27] and the EM based local approach [23]. GL, OF and EM are used as abbreviations for these three methods respectively, and n-RW, RW are for the proposed method without and with using the weighted colour correction scheme respectively, in the reminder of the paper. The experimental results will be evaluated from three aspects.

4.1 Subjective Evaluation

First, the colour correction results are assessed in a subjective manner. We believe that human perception is a meaningful way to evaluate the results qualitatively because humans are the final evaluator. In Table. 1, the corrected source images produced

by four methods for "Art" and "Moebius" are shown. It is visible that the colour consistencies between two stereo images have been greatly improved after colour correction. The proposed method can produce better results (Table. 1 6^{th} col) than others, especially for regions which contain a dominant colour. All colour correction results can be found in dropox_share_page

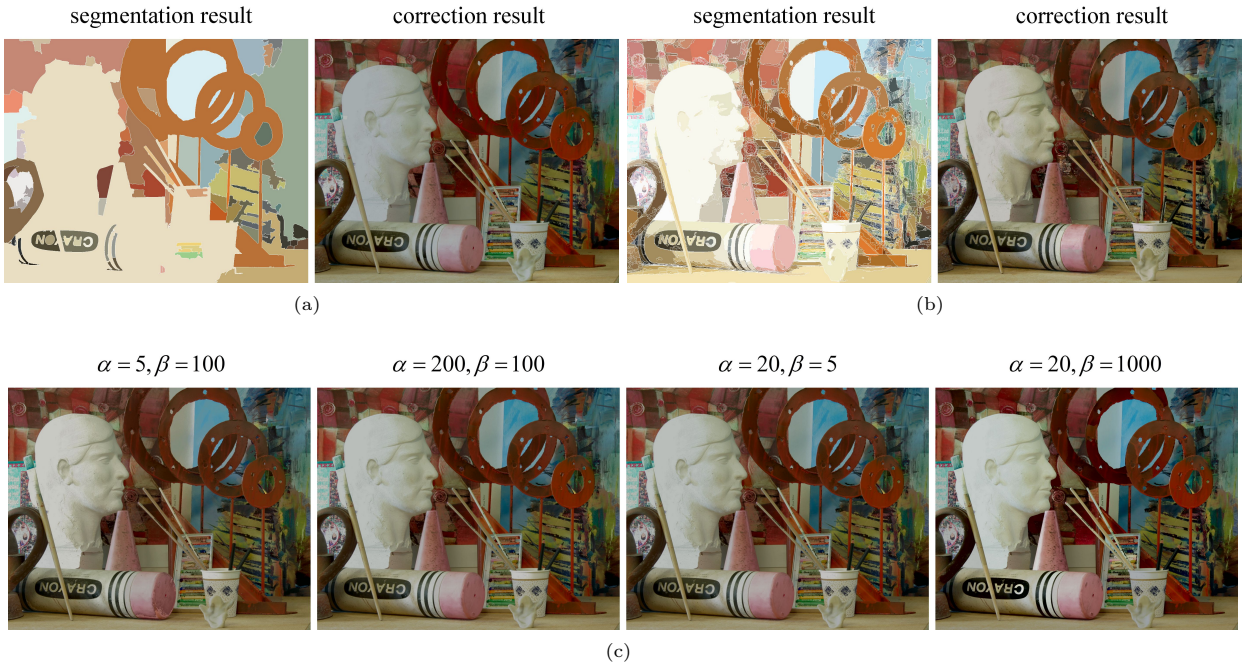


Figure 6: *Segmentation and colour correction results with different parameters. (a) and (b) The segmentation and colour correction results with ($\sigma_s = 25$, $\sigma_r = 18$, $M = 600$) and ($\sigma_s = 5$, $\sigma_r = 3$, $M = 50$), respectively. There are 90 segmented regions in (a) and 4181 segmented regions in (b). (c) The colour correction results for different (α, β).*

A user study is designed to compare the proposed method with GL, OF and EM qualitatively. For each dataset, the reference image and four colour corrected source images are displayed in a pre-calibrated monitor simultaneously. The corrected source images are displayed in a randomized order. Then the participants need to choose the most similar corrected source image to the reference image. The participants' choices are recorded to calculate the percentage of which method produces the most similar results. We recruited 30 young people without visual weakness to attend the user study. Nearly 93.1% of the users chose the results of the proposed method as the most similar one, which demonstrates the feasibility of the proposed method in terms of subjective evaluation.

		colour correction results			
source image	ground-truth	GL	OF	EM	RW
					

Table 1: *Comparison of colour correction results for "Art" and "Moebius". From left to right, the six images are the original source image, the ground truth image, four corrected image via GL, OF, EM and RW, respectively.*

4.2 Colour Similarity Evaluation

To evaluate the colour correction results quantitatively, the metric used by Oliveria et al. [17] is applied to measure the colour similarity improvement after colour correction, which is defined as follows:

$$CC_{method} = 1 - \frac{CS_{method}}{CS_{base}} \tag{10}$$

The criterion CS stands for colour similarity, which is defined as the three-channel Euclidean distance between the corrected

source image and the ground truth image.

$$CS = ||c_{gt}(i,j) - c(i,j)|| \tag{11}$$

CS_{base} denotes the colour similarity between the ground truth and the original source image. CS_{method} denotes the colour similarity between the ground truth and the corrected source image after a colour correction method, where *method* takes the value in $\{GL, OF, EM, n-RW, RW\}$. Thus, the metric CC_{method} is used to describe the improvement ratio of colour similarity between the ground truth and the corrected source image. The larger the metric CC_{method} is, the better the method is in compensating colour discrepancies.

Fig. 7 shows the colour correction improvement ratio of different methods. The horizontal axis refers to the index of a pair of stereo images, valued from 0 to 29. The vertical axis refers to the improvement ratio of colour similarity after being magnified 100 times. It is visible that can correct the colour discrepancy well. First, GL tends to produce unnatural looking results, especially for the scene details, because it considers colour discrepancies in different regions as a whole. Second, OF computes a constant colour discrepancy as a local correction function. If the colours of SIFT matches are not dominant in the region, then the method will lead to incorrect correction result. The proposed method employs all pixels in the local region to compute a linear local function, rather than sampled ones. Obviously, it is more robust than OF. When the colour variations of the image are relative small, i.e., containing a small number of the regions, EM will perform the segmentation similarly with the proposed method. If the colour variations of the image are large, the GMM computation in the EM will be not convergent due to a large amount of components, which may lead to incorrect correction result. In order to assess the proposed weighted colour correction scheme, the colour correction results without the weighted scheme (n-RW) are implemented and compared with other methods as well. As shown in Fig. 7, among five colour correction methods, the proposed method and EM perform better than the other ones. But for the images containing abundant colour variations, the proposed method performs better than EM.

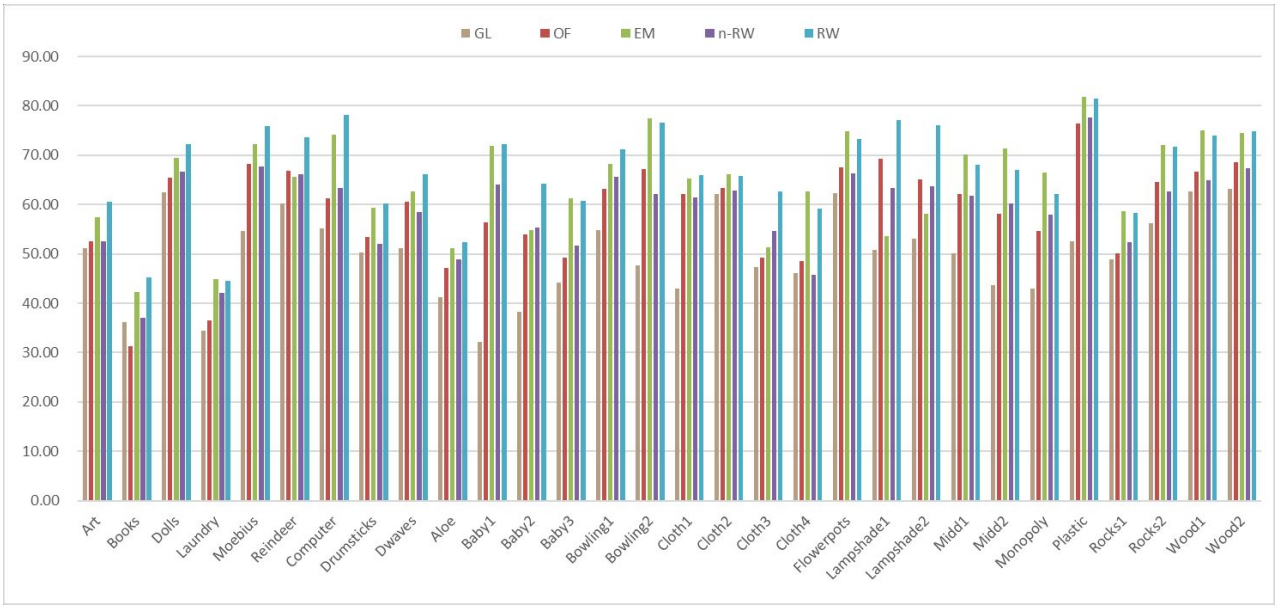


Figure 7: *Comparison of the improvement ratio of colour similarity of colour correction results using five methods.*

4.3 Stereo Vision Evaluation

We also evaluate our approach in terms of stereo matching. The objective is that for two stereo images with colour discrepancies, the disparity results of stereo matching can be improved after the proposed colour correction algorithm. The stereo matching algorithm [28], which is also used in the stereo vision evaluation of the OF local method, is employed to calculate the disparity results for all datasets. The matching cost is the max absolute difference of three colour channels, which would be influenced by colour discrepancy. An inter-regional cooperative optimization has been employed to minimize the matching costs. It is one of the best ranked algorithms in the Middlebury Evaluation Website [19].

First, the reference image and the original source image, as a stereo pair, are used to calculate a disparity map D. Then, the reference image and the corrected source image are used to calculate another disparity map D^c. Disparity maps for the "Art" and "Moebius" are shown in Table. 2, while other results are shown in dropox_share_page. Influenced by the colour discrepancy, the disparity results without colour correction (Table. 2 2^{nd} col) have much noise and are quite different from the ground-truth disparity (Table. 2 1^{st} col). In contrast, because of the resemblance of colour discrepancies, smoother and more correct disparity results are obtained after using colour correction.

To evaluate the colour correction methods quantitatively, a pixel that differ from the true disparity value by more than 1 is regarded as a bad pixel. We calculated and compared the percentage of bad pixels in D and D^c. Table. 3 lists the percentages of bad pixels without correction and after correction by the five methods for 27 datasets ("Computer", "Drumsticks" and "Dwaves" are not evaluated because no ground-truth disparities are provided). In Table. 3, PP denotes the bad pixels percentage of original disparity result; PP_{GL}, PP_{EM}, PP_{OF}, PP_{n-RW} and PP_{RW} denote the bad pixels percentages of disparity results after colour correction by GL, EM, OF, n-RW and RW.

As shown in Table. 3, all percentages of bad pixels in the test datasets are decreased notably after the proposed colour correction method is used. Consistent with the colour similarity comparison results, the disparity results via the proposed method and EM outperform those via the other methods. But the disparity results via EM tend to be noisier, while the disparity results via the proposed method are smoother, especially for the images containing abundant colour variations (Table. 2). Moreover, even without using weighted colour correction, the percentage of bad pixels by the proposed method performs better than that

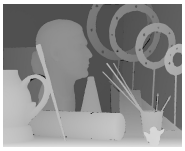
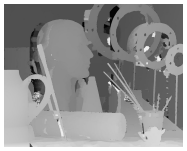
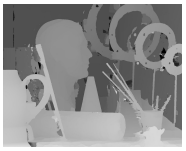
ground-truth	without correction	disparity results after colour correction			
		GL	OF	EM	RW
					
					

Table 2: *Disparity results of "Art" and "Moebius".*

by OF. It can be concluded that the local colour correction function in the proposed method are more reasonable than that in OF.

Some complicated similarity measures maybe handle colour discrepancies well, e.g., the census transform [31]. In order to test the necessary of colour correction, we also implemented the stereo matching algorithm via a more robust matching cost "AD-census", which is the combination of the absolute differences measure and the census transform. According to the implementation results in Fig. 8, we can find that, the percentage of bad pixels of disparity results via the proposed local colour correction method has decreased accordingly, compared with those without local colour correction. At least for those matching costs tested in our paper, we can draw that it is a good choice to compensate the colour discrepancies before stereo matching.

	Bad Pixels Percentage (%)					
Dataset	PP	PP_{GL}	PP_{OF}	PP_{EM}	$PP_{n\text{-}RW}$	PP_{RW}
Art	44.05	31.55	21.46	17.99	20.87	15.30
Books	45.43	32.30	27.20	23.56	29.04	21.11
Dolls	32.09	20.68	17.52	12.96	15.21	9.29
Laundry	48.07	22.02	20.52	16.57	21.62	17.32
Moebius	38.48	28.22	23.61	16.39	18.35	10.15
Reindeer	45.97	33.63	22.16	23.11	21.95	13.57
Aloe	36.76	26.46	17.76	12.53	15.16	10.52
Baby1	21.70	15.26	10.97	9.27	14.37	9.98
Baby2	23.18	20.93	19.29	17.62	18.11	12.23
Baby3	24.40	12.92	11.14	10.08	11.41	10.84
Bowling1	32.40	26.02	20.52	18.58	20.16	17.21
Bowling2	22.94	17.05	11.89	8.22	13.74	9.67
Cloth1	70.18	40.24	29.16	21.64	31.39	19.56
Cloth2	45.35	38.58	29.80	25.32	28.87	27.85
Cloth3	39.61	24.95	19.19	16.53	15.64	6.96
Cloth4	41.79	33.68	30.39	22.38	31.01	24.17
Flowerpots	30.10	29.01	17.22	12.37	19.16	13.83
Lampshade1	35.26	21.42	16.72	27.21	22.04	15.05
Lampshade2	28.90	23.06	19.07	21.24	19.89	13.28
Midd1	36.08	29.69	22.64	17.53	20.79	17.16
Midd2	55.24	52.58	33.84	31.12	35.73	30.94
Monopoly	33.94	26.32	20.89	16.60	18.52	15.98
Plastic	47.50	33.09	24.97	24.98	25.21	22.12
Rocks1	27.90	16.10	14.46	11.84	15.04	11.25
Rocks2	16.11	10.21	6.95	6.19	7.23	5.67
Wood1	46.13	32.71	19.25	15.74	21.09	16.97
Wood2	16.44	10.41	6.12	4.92	13.64	5.08

Table 3: *Percentages of bad pixels in disparity results. The proposed method results in better disparity results for the majority of the tested examples than other methods.*

5 Conclusion and Future Work

In this paper, we proposed a local colour correction method to address the colour discrepancies between two stereo images. The proposed method segmented the source image into a set of regions and compensated for colour discrepancies by using a weighted combination of all regions. In this way, the global colour distribution of the corrected source image is close to the reference image while the local colour information is maintained. The detailed discussion and comparison show that the proposed method can

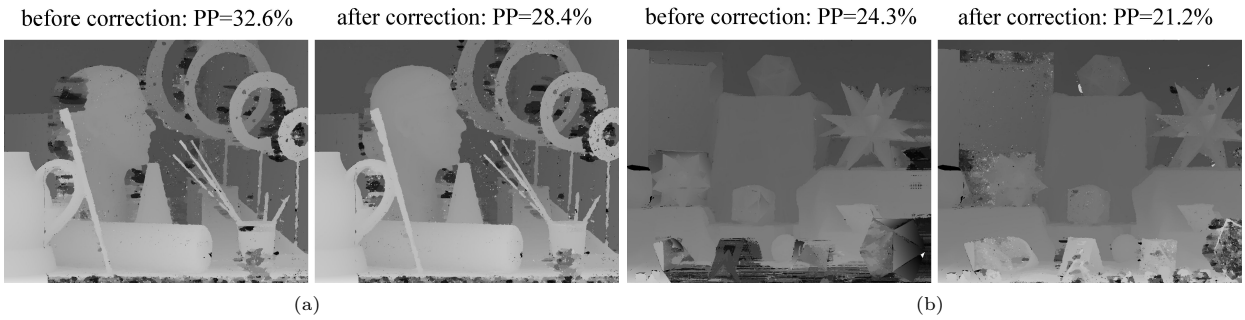


Figure 8: *The disparity results by using a robust matching cost "AD-census" before colour correction and after colour correction, for "Art" and "Moebius".*

improve the accuracy of disparity results in stereo matching.

The performance of the proposed approach relies on a reasonable segmentation and reliable SIFT feature matches. We assume that each segmented region corresponds to a scene plane. Under this assumption, the spatial relationship between a segmented region and its desirable corresponding region can be approximated using a homography matrix. As a result, the estimated regional correspondences are approximate. Thus seeking a consistent segmentation between the stereo images and finding the better regional correspondences will be the central focus of our future works. Finally, although our method is applied to stereo images currently, it can be extended for multi-view application in a straightforward way.

6 Acknowledgements

omitted for review

References

[1] Simon Baker, Daniel Scharstein, JP Lewis, Stefan Roth, Michael J Black, and Richard Szeliski. A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1):1–31, 2011.

[2] Michael Bleyer, Sylvie Chambon, Uta Poppe, and Margrit Gelautz. Evaluation of different methods for using colour information in global stereo matching approaches. *Int. Society for Photogrammetry and Remote Sensing*, pages 63–68, 2008.

[3] Sylvie Chambon and Alain Crouzil. Colour correlation-based matching. *International Journal of Robotics and Automation*, 20(2):78–85, 2005.

[4] Dorin Comaniciu and Peter Meer. Mean shift analysis and applications. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1197–1203. IEEE, 1999.

[5] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619, 2002.

[6] Hasan Sheikh Faridul, Tania Pouli, Christel Chamaret, Jürgen Stauder, Alain Trémeau, Erik Reinhard, et al. A survey of color mapping and its applications. *Eurographics State of the Art Report, Strasbourg*, 2014.

[7] Pedro F Felzenszwalb and Daniel P Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.

[8] Graham D Finlayson, Michal Mackiewicz, and Anya Hurlbert. Colour correction using root-polynomial regression. *Image Processing, IEEE Transactions on*, 24(5):1460–1470, 2015.

[9] Yasutaka Furukawa, Brian Curless, Steven M Seitz, and Richard Szeliski. Towards internet-scale multi-view stereo. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1434–1441. IEEE, 2010.

[10] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[11] Yong Seok Heo, Kyoung Mu Lee, and Sang Uk Lee. Robust stereo matching using adaptive normalized cross-correlation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(4):807–822, 2011.

[12] Heiko Hirschmuller and Daniel Scharstein. Evaluation of cost functions for stereo matching. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[13] Adrian Ilie and Greg Welch. Ensuring color consistency across multiple cameras. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1268–1275. IEEE, 2005.

[14] Jiaya Jia and Chi-Keung Tang. Tensor voting for image correction by global and local intensity alignment. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(1):36–50, 2005.

[15] Seon Joo Kim and Marc Pollefeys. Robust radiometric calibration and vignetting correction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(4):562–576, 2008.

[16] Xing Mei, Xun Sun, Mingcai Zhou, Shaohui Jiao, Haitao Wang, and Xiaopeng Zhang. On building an accurate stereo matching system on graphics hardware. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 467–474. IEEE, 2011.

[17] Miguel Oliveira, Angel D Sappa, and Vitor Santos. Unsupervised local color correction for coarsely registered images. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 201–208. IEEE, 2011.

[18] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Colour transfer between images. *IEEE Computer graphics and applications*, (5):34–41, 2001.

[19] Daniel Scharstein and Richard Szeliski. Middlebury stereo vision page. *Online at http://www. middlebury. edu/stereo*, 2002.

[20] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, 47(1-3):7–42, 2002.

[21] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519–528. IEEE, 2006.

[22] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.

[23] Yu-Wing Tai, Jiaya Jia, and Chi-Keung Tang. Local color transfer via probabilistic segmentation by expectation-maximization. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 747–754. IEEE, 2005.

[24] Yu-Wing Tai, Jiaya Jia, and Chi-Keung Tang. Soft color segmentation and its applications. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(9):1520–1537, 2007.

[25] Mehrdad Panahpour Tehrani, Akio Ishikawa, Shigeyuki Sakazawa, and Atsushi Koike. Iterative colour correction of multicamera systems using corresponding feature points. *Journal of Visual Communication and Image Representation*, 21(5):377–391, 2010.

[26] Haoxing Wang, Longquan Dai, and Xiaopeng Zhang. Consistent segmentation based colour correction for coarsely registered images. In *Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on*, pages 319–324. IEEE, 2013.

[27] Qi Wang, Xi Sun, and Zengfu Wang. A robust algorithm for color correction between two stereo images. In *Computer Vision–ACCV 2009*, pages 405–416. Springer, 2010.

[28] Zeng-Fu Wang and Zhi-Gang Zheng. A region based stereo matching algorithm using cooperative optimization. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[29] Wei Xu and Jane Mulligan. Performance evaluation of color correction approaches for automatic multi-view image and video stitching. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 263–270. IEEE, 2010.

[30] Kenji Yamamoto, Tomohiro Yendo, Toshiaki Fujii, Masayuki Tanimoto, and David Suter. Colour correction for multi-camera system by using correspondences. In *ACM SIGGRAPH 2006 Research posters*, page 73. ACM, 2006.

[31] Ramin Zabih and John Woodfill. Non-parametric local transforms for computing visual correspondence. In *Computer Vision-ECCV'94*, pages 151–158. Springer, 1994.

[32] C Lawrence Zitnick, Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. High-quality video view interpolation using a layered representation. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 600–608. ACM, 2004.