

Generalized Content-Preserving Warps for Image Stitching

BMVC 2018 Submission # 401

Abstract

Local misalignment caused by global homography is a common issue in image stitching task. Content-Preserving Warping (CPW) is a typical method to deal with this issue, in which geometric and photometric constraints are imposed to guide the warping process. One of its essential condition however, is colour consistency, and an elusive goal in real world applications. In this paper, we propose a Generalized Content-Preserving Warping (GCPW) method to alleviate this problem. GCPW extends the original CPW by applying a colour model that expresses the colour transformation between images locally, thus meeting the photometric constraint requirements for effective image stitching. We combine the photometric and geometric constraints and jointly estimate the colour transformation and the warped mesh vertexes, simultaneously. We align images locally with an optimal grid mesh generated by our GCPW method. Experiments on both synthetic and real images demonstrate that our new method is robust to colour variations, outperforming other state-of-the-art CPW-based image stitching methods.

1 Introduction

Given two images capturing a same scene, the objective of image stitching is to combine them into a panorama with a wider field of view (FOV). At the same time, the stitching result must be free from obvious misalignment artifacts. In order to achieve this purpose, Brown and Lowe [2] proposed to stitch images through global homography, which was estimated by Direct Linear Transform (DLT) using matched keypoints in an overlapping region. Li *et al.* [12] utilized dual-feature correspondences to estimate global homography, which achieved a more accurate and more robust estimation result. Chang *et al.* [3] proposed to estimate the global homography by deep learning. They adopted features learned from convolutional neural network (CNN), and designed a neural network layer to perform the inverse compositional Lucas-Kanade algorithm [1] to estimate the homography parameters. Methods based on global homography however, are prone to suffer from misalignment artifacts, as these methods assume that scenes are located on a single plane or the motion of camera is pure rotation, conditions that are hard to meet in many practical applications.

Content-Preserving Warping (CPW) methods therefore were proposed to overcome the shortcoming of methods using global homography. The CPW framework was first proposed by Liu and Jin [15] for the purpose of video stabilization and then widely applied in image stitching. These methods formulated the local alignment of images as a mesh deformation problem and applied various types of constraints to guide the mesh deformation process. Generally, two types of constraints were considered: geometric and photometric constraints. CPW methods can be divided into two categories based on the type of constraint.

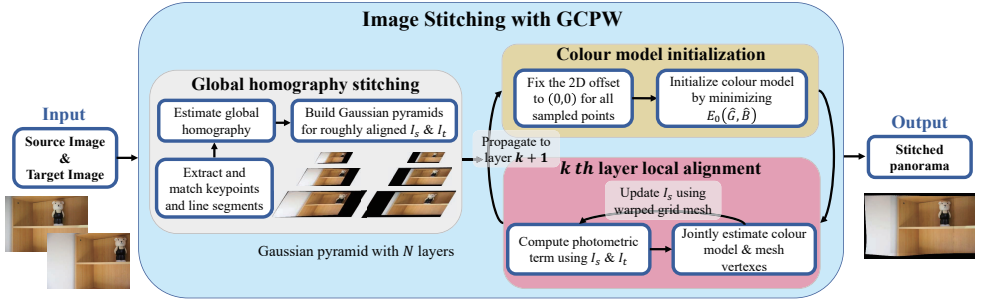


Figure 1: An overview of our image stitching method with GCPW. From given image pair, we first roughly align them with a global homography. Next, we combine photometric and geometric constraints in the GCPW framework and design a robust optimization scheme to estimate the colour model and the mesh vertexes, simultaneously. At last, images are locally aligned based on the optimized grid mesh to create a stitched panorama.

For methods with geometric constraint, research [10, 13, 20] detected and matched keypoints in an overlapping region and constrained each pair of matched keypoints to be warped to the same position during mesh deformation. Guo *et al.* [7] proposed self-adaptive thresholds to detect keypoints in order to collect sufficient amount of point correspondences for adequate warping guidance. Similarly, Zhang *et al.* [21] designed a local-homography-based outlier rejection procedure to preserve point correspondences located on different planes, while Chen and Chuang [4] took the aligned mesh vertexes generated by APAP [19] as uniformly distributed matched keypoints in an overlapping region. These methods however, still fail when scenes were lack of textures. To deal with this problem, Li *et al.* [12] and Xiang *et al.* [18] resorted to line segment correspondences. They combined point and line constraints in a CPW framework to obtain better stitching results than methods using only point constraints.

Geometric constraints however, are not robust enough. Although the combination of keypoints and line segments has significantly improved alignment quality, these two kinds of geometric constraints have limitations. For one, keypoints are often distributed with spatial bias in an overlapping region [11]. For another, the extraction of line segments is difficult in images with small structures or tiny gradients.

Lin *et al.* [14] therefore proposed to introduce photometric constraint into the CPW framework to stitch images. They sampled a set of points at a fixed interval in an overlapping region and required overall photometric differences of all sampled points to be minimized by mesh deformation. This process is just similar to the one of optical flow estimation [6], which usually requires images to obey the colour consistency assumption [5]. Images with significant brightness or colour variations are common however, which would be resulted by complex scenes or various illumination conditions. Under these circumstances, we must correct colour differences before adopting a photometric constraint to stitch images. Many techniques can be applied to compensate image colour differences. These includes: invariant feature extraction [9] and colour consistency correction [17], but these methods usually work as pre-processing operations, and may fail to correct the differences thoroughly or may deteriorate the original image content, which would affect the effectiveness of the photometric constraint and produce poor stitching results.

In this paper, we propose a *Generalized Content-Preserving Warping* (GCPW) method to stitch images with colour variations. The proposed GCPW extends the original CPW by appending an affine colour model that expresses the colour transformation from a source image to a target image locally. Our extension to CPW makes the photometric constraint robust to significant colour variations. Figure 1 presents an overview of our image stitching scheme incorporating GCPW. We first roughly align images with a global homography that is estimated from dual features [12]. Next, we locally align them using the proposed GCPW. Specifically, we combine geometric and photometric constraints in the GCPW framework. Furthermore, we designed a robust optimization procedure that jointly estimates the affine colour model and the warped mesh vertexes. At last, images are locally aligned according to the optimized grid mesh. We tested our proposed method on both synthetic and real images. The results from these experiments show that the proposed method is robust to colour variations, producing better alignment quality than other state-of-the-art methods.

2 Image Stitching with GCPW

Given two images I_1 and I_2 , we first roughly align them with a global homography to obtain I_s and I_t . In the classical CPW framework, the local alignment of I_s and I_t is then formulated into a problem of mesh deformation, in which various geometric and photometric constraints are imposed to align images and preserve image content at the same time. Lin *et al.* [14] demonstrated that the photometric constraints can improve the alignment quality, but this method requires that images obey the colour consistency assumption. We therefore propose the GCPW framework (Section 2.1), and design a photometric constraint beyond colour consistency (Section 2.2). We combine the proposed photometric constraint with geometric constraints in the GCPW framework to locally align images by mesh deformation (Section 2.3). Considering the high degree of freedom (DoF) of our proposed GCPW model, we designed an optimization scheme to estimate the colour model and the mesh vertexes simultaneously (Section 2.4).

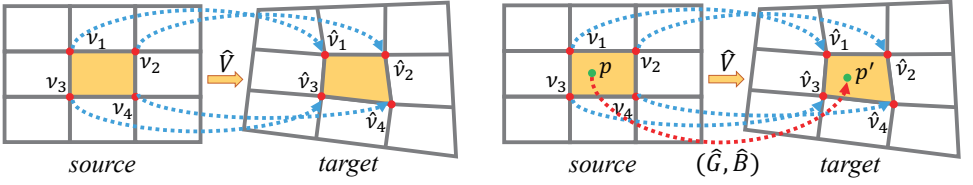
2.1 The Proposed GCPW

In this section, we introduce the proposed GCPW model. Figure 2 shows a comparison between CPW and GCPW. Specifically, I_s is first divided into an $m \times n$ uniform grid mesh ($m = n = 32$ in this paper), whose vertexes are denoted as \mathbf{V} . In contrast to CPW, which only estimates warped mesh vertexes $\hat{\mathbf{V}}$, our proposed GCPW optimizes $\hat{\mathbf{V}}$ and the local colour transformation from I_s to I_t ($\hat{\mathbf{G}}, \hat{\mathbf{B}}$) at the same time. We assume that pixels within a quad share a same colour transformation. Similar to [8], we use an affine model to express the colour transformation from I_s to I_t within a quad, where $\hat{\mathbf{G}}$ and $\hat{\mathbf{B}}$ denote the colour gain and the colour bias respectively. For typical *RGB* images, we convert them into *YCbCr* colour space and compute colour models for each image channel separately.

The proposed GCPW typically has two advantages. Firstly, it considers the colour transformation from I_s to I_t . Therefore, it measures the photometric error normally even when I_s and I_t have significant colour differences. Secondly, the photometric error computed by GCPW retains the quadratic property, which ensures that the photometric constraint is easily combined with other widely used constraints to align images locally.

2.2 Photometric Constraints Beyond Colour Constancy

For two roughly aligned images I_s and I_t , our objective is to locally align them by warping I_s to I_t via mesh deformation. Lin *et al.* [14] proved that photometric constraint is an important



(a) The CPW framework [15]

(b) The proposed GCPW framework

Figure 2: A comparison between CPW and GCPW. (a) The original CPW only estimates warped mesh vertexes $\hat{\mathbf{V}}$. (b) Our proposed GCPW jointly estimates warped mesh vertexes $\hat{\mathbf{V}}$ and the local colour model $(\hat{\mathbf{G}}, \hat{\mathbf{B}})$. The colour model in GCPW expresses the colour transformation from source image to target image (e.g., from p to p').

complementary to guide this process. The required colour constancy condition however, is not always satisfied in practice. We therefore propose a photometric constraint beyond colour consistency. Specifically, we uniformly sample a set of points at a fixed interval (three pixels in this paper) both horizontally and vertically in an overlapping region. For each sampled point \mathbf{q} in I_s , we want to find a 2D offset $\tau(\mathbf{q})$ that minimizes the photometric difference between \mathbf{q} in I_s and $\mathbf{q} + \tau(\mathbf{q})$ in I_t . Considering the colour variation between I_s and I_t , we measure their photometric error as:

$$\|(\mathbf{G}^i(\mathbf{q}) \times I_s^i(\mathbf{q}) + \mathbf{B}^i(\mathbf{q})) - I_t^i(\mathbf{q} + \tau(\mathbf{q}))\|^2, \quad (1)$$

where i indexes the image channel. $I_s^i(\mathbf{q})$ is the intensity of I_s^i at \mathbf{q} , and $I_t^i(\mathbf{q} + \tau(\mathbf{q}))$ is the intensity of I_t^i at $\mathbf{q} + \tau(\mathbf{q})$. $\mathbf{G}^i(\mathbf{q})$ and $\mathbf{B}^i(\mathbf{q})$ denote the colour gain and bias of the quad that \mathbf{q} is located. Since I_s and I_t are roughly aligned, we assume that $\tau(\mathbf{q})$ is small and set its initial value as $(0, 0)$. Using the first-order Taylor expansion of $I_t^i(\mathbf{q} + \tau(\mathbf{q}))$, the photometric error defined by Eq. 1 can be expressed as:

$$E_c(\tau(\mathbf{q})) = \|(\mathbf{G}^i(\mathbf{q}) \times I_s^i(\mathbf{q}) + \mathbf{B}^i(\mathbf{q})) - (I_t^i(\mathbf{q}) + \nabla I_t^i(\mathbf{q})\tau(\mathbf{q}))\|^2, \quad (2)$$

where $\nabla I_t^i(\mathbf{q})$ is approximated by image gradient of I_t^i at \mathbf{q} .

2.3 Local Alignment by Mesh Deformation

In the proposed GCPW framework, we combine the proposed photometric constraint with other widely used constraints to align images locally, and define the objective function as follows:

$$E(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}}) = \lambda_1 E_p(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}}) + \lambda_2 E_g(\hat{\mathbf{V}}) + \lambda_3 E_s(\hat{\mathbf{V}}) + \lambda_4 E_c(\hat{\mathbf{G}}, \hat{\mathbf{B}}), \quad (3)$$

where E_p is the proposed photometric constraint, and E_g is the geometric constraint and similar to the one in [12]. E_s and E_c are smoothness constraints, in which E_s is used to constrain the similarity transformation of the warped mesh [15] and E_c is the term designed to preserve the smoothness of the local colour model in GCPW. The associated weights are denoted by λ_1 , λ_2 , λ_3 and λ_4 respectively. ($\lambda_1 = 100.0$, $\lambda_2 = \lambda_4 = 1.0$ and $\lambda_3 = 0.5$ in our implementation).

Photometric term For each sampled point \mathbf{q} , we express its corresponding photometric error according to Eq. 2. In order to associate it with mesh deformation, we further parameterize the 2D offset $\tau(\mathbf{q})$ by:

$$\tau(\mathbf{q}) = \hat{\mathbf{q}} - \mathbf{q}, \hat{\mathbf{q}} = \sum_{k=1}^4 w_k \hat{\mathbf{v}}_k, \quad (4)$$

where $\hat{\mathbf{q}}$ is the warped position of \mathbf{q} , w_k are bilinear interpolation coefficients computed by expressing \mathbf{q} with four initial mesh vertexes \mathbf{v}_k^q that enclose \mathbf{q} , and $\hat{\mathbf{v}}_k^q$ are the warped

positions of \mathbf{v}_q^k . $E_p(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}})$ is computed by summing up $E_c(\tau(\mathbf{q}))$ over all sampled points and all image channels:

$$E_p(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}}) = \sum_{i=1}^3 \sum_{\mathbf{q}} E_c(\tau(\mathbf{q}^i)), \quad (5)$$

Geometric term Similar to [12], we adopt geometric constraint based on dual-feature correspondences. Thus, our geometric term can be divided into the point component $E_f(\hat{\mathbf{V}})$ and the line component $E_l(\hat{\mathbf{V}})$.

For each pair of matched keypoints $(\mathbf{p}, \mathbf{p}')$, we restrict the warped position of \mathbf{p} to being close to \mathbf{p}' , which is achieved by minimizing $D_p = \|\hat{\mathbf{p}} - \mathbf{p}'\|^2$. For each pair of matched line segment $\mathbf{l} = [a, b, c]$ and $\mathbf{l}' = [a', b', c']$, we uniformly sample keypoints along \mathbf{l} at a fixed interval (five pixels in this paper) and stipulate that the total distance from these keypoints to \mathbf{l}' is minimized. We traverse all pairs of keypoints and line segments to get the point component $E_f(\hat{\mathbf{V}})$ and line component $E_l(\hat{\mathbf{V}})$. More details about $E_f(\hat{\mathbf{V}})$ and $E_l(\hat{\mathbf{V}})$ can be found in [12]. The overall geometric term is:

$$E_g(\hat{\mathbf{V}}) = E_f(\hat{\mathbf{V}}) + E_l(\hat{\mathbf{V}}). \quad (6)$$

Similarity transformation term The same similarity transformation constraint [15] is used in our implementation. Specifically, each quad $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]$ is divided into two triangulations $\triangle \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3$ and $\triangle \mathbf{v}_2 \mathbf{v}_3 \mathbf{v}_4$. We express \mathbf{v}_1 of $\triangle \mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3$ in the local coordinate system defined by \mathbf{v}_2 and \mathbf{v}_3 as follows:

$$\mathbf{v}_1 = \mathbf{v}_2 + u(\mathbf{v}_3 - \mathbf{v}_2) + v\mathbf{R}_{90}(\mathbf{v}_3 - \mathbf{v}_2), \quad \mathbf{R}_{90} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad (7)$$

where u and v are computed from original positions of $\mathbf{v}_1, \mathbf{v}_2$ and \mathbf{v}_3 . We require that \mathbf{v}_1 be represented by \mathbf{v}_2 and \mathbf{v}_3 using the same local coordinates (u, v) before and after warping. The overall similarity transformation term is defined as:

$$E_s(\hat{\mathbf{V}}) = \sum_i^N \|\hat{\mathbf{v}}_1^i - (\hat{\mathbf{v}}_2^i + u(\hat{\mathbf{v}}_3^i - \hat{\mathbf{v}}_2^i) + v\mathbf{R}_{90}(\hat{\mathbf{v}}_3^i - \hat{\mathbf{v}}_2^i))\|^2, \quad (8)$$

where N is the total number of triangulations in the grid mesh.

Colour smoothness term A colour smoothness term is designed to constrain the smoothness of colour models within an eight-connected neighboring region. Let Q be a quad of the grid mesh, and Q' be one of its neighboring quads. Let $y = gx + b$ and $y = g'x + b'$ be their affine colour models on one image channel. We sample a set of intensity values in normalized intensity range (e.g., sample values within $[0, 1]$ at an interval of 0.1), and require the sampled values to remain close after the neighboring two affine transformations, which is achieved by minimizing $D_c = \sum_i \|(gx_i + b) - (g'x_i + b')\|^2$, where x_i denotes sampled intensity value.

For each quad of the grid mesh, we constrain the similarities between its colour model with those of all neighboring quads over three image channels. The colour smoothness term is defined as:

$$E_c(\hat{\mathbf{G}}, \hat{\mathbf{B}}) = \sum_{n=1}^3 \sum_i \sum_{j \in \Omega_i} D_c(i, j, n), \quad (9)$$

where n indexes image channels, Ω_i denotes the eight-connected neighborhood of quad Q_i .

2.4 Robust Optimization

Since the proposed image stitching method with GCPW estimates warped mesh vertexes $\hat{\mathbf{V}}$ and a local colour model $(\hat{\mathbf{G}}, \hat{\mathbf{B}})$ at the same time, the high degree of freedom (DoF) makes the optimization result unstable. Therefore, a two-step optimization pipeline is designed: We first initialize $(\hat{\mathbf{G}}, \hat{\mathbf{B}})$ using sampled points. After that, $\hat{\mathbf{V}}$ and $(\hat{\mathbf{G}}, \hat{\mathbf{B}})$ are jointly estimated by

minimizing Eq. 3. As for the colour model initialization, we fix the $\tau(\mathbf{q})$ in Eq. 5 at (0,0) and initialize the colour model by minimizing the following function:

$$E_0(\hat{\mathbf{G}}, \hat{\mathbf{B}}) = E_p(\hat{\mathbf{G}}, \hat{\mathbf{B}}) + E_c(\hat{\mathbf{G}}, \hat{\mathbf{B}}) + E_r(\hat{\mathbf{G}}, \hat{\mathbf{B}}), \quad (10)$$

where E_p and E_c have been defined in Eq. 5 and Eq. 9. E_r is a regular term that we add for initialization, and defined as: $\sum_{Q_i \cap (I_s \cap I_t) = \emptyset} \|\hat{\mathbf{G}}(Q_i) - 1.0\|^2 + \|\hat{\mathbf{B}}(Q_i) - 0.0\|^2$. E_r restricts the colour gain to being close to 1.0 and the colour bias to being close to 0.0 when the quad Q_i is not overlapped with an overlapping region in the image.

Both the initialization function $E_0(\hat{\mathbf{G}}, \hat{\mathbf{B}})$ and the final objective function $E(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}})$ are quadratic and can be easily minimized by any sparse linear system. In order to handle images with large displacements, we built a three-layer Gaussian pyramid and adopted a coarse-to-fine optimization scheme. For each layer of the pyramid, as shown in Figure 1, we minimize $E_0(\hat{\mathbf{G}}, \hat{\mathbf{B}})$ to initialize the colour model, and then iteratively optimize $E(\hat{\mathbf{V}}, \hat{\mathbf{G}}, \hat{\mathbf{B}})$. As the grid mesh gets updated after each iteration, the image content of I_s is re-rendered according to the updated mesh. The iteration terminates when the grid mesh becomes stable. We warp I_s based on the optimized grid mesh to locally align I_s and I_t .

3 Experiments and Results

In this section, we conduct experiments on both synthetic and real images to demonstrate the superiority of our proposed image stitching method with GCPW. All experiments are conducted with default parameter settings as discussed in previous sections of this paper. For convenience, in these comparative experiments, we do not adopt any post-processing methods but simply blend the warped I_s and I_t by the intensity average to obtain the final stitching result.

3.1 Experiments on Synthetic Images

We performed experiments on synthetic images with different degrees of colour variation to demonstrate the effectiveness and robustness of our proposed image stitching method. We collected images that have nearly no colour differences, which are obtained from publicly available datasets [4, 12, 19] or captured by ourselves. In order to synthesize image pairs with different colour variations, for each pair of images, we used a colour model similar to the model in [16] to generate 16 pairs of images, which explicitly have different degrees of colour variation (from 1 to 16, the colour difference varies from smallest to largest). More details about this generation process are presented in the supplementary material.

We stitched these synthetic image pairs with our method and compared the results with those produced by three different methods: DF-W, LSH+MF-W, and ROS+MF-W. The dual-feature warping (DF-W) method [12] locally aligns I_s and I_t using dual-feature correspondences, and we tested it with our own re-implementation. For the second and third tested methods, we combined the photometric constraint proposed in [14] with DF-W in the CPW framework to get the multiple-feature warping (MF-W) method. As the photometric constraint in [14] requires images to obey colour consistency assumption that are not satisfied in our synthetic images, we therefore tested MF-W cooperated with two types of pre-processing operations to get our another two tested methods. In the first operation, we detected invariant features using the locality sensitive histogram (LSH) [9] and extracted photometric constraint based on LSH feature images. We refer to the second tested method as LSH+MF-W. In the second operation, we performed colour consistency correction by the colour re-mapping optimization scheme (ROS) [17] and imposed photometric constraint using corrected images. We refer to the third tested method as ROS+MF-W.

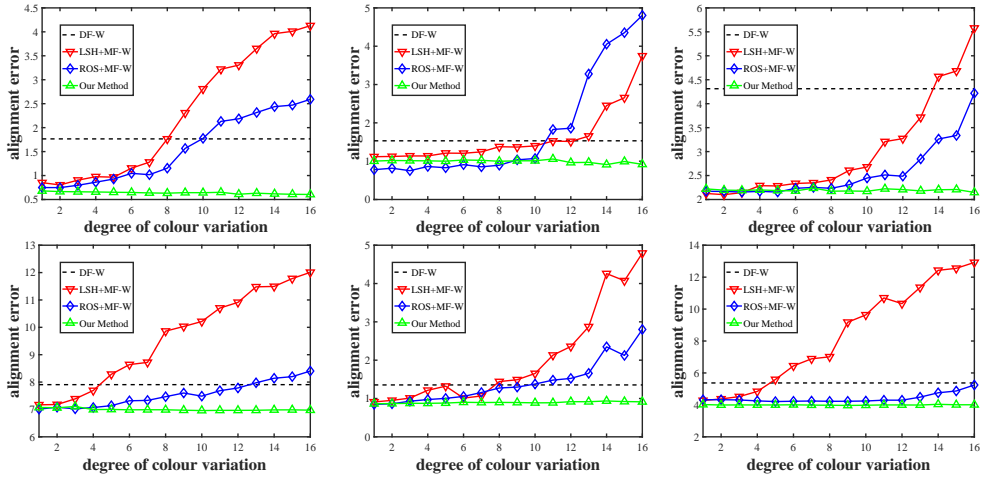


Figure 3: Comparisons on alignment errors of stitching results produced by different methods on synthetic images with different degrees of colour variations.

In the experiment, we use the photometric term as the only variable, and fix point and line correspondences among image pairs in different degrees of colour variations. In order to measure alignment quality quantitatively, we use the same error metric as in [12, 14], which computes the RMSE of one minus normalized cross correlation over a $W \times W$ window ($W = 5$ in this paper) for pixels in the overlapping region. As the quality measurement is relevant to specific intensities of image content, for the sake of fairness, all alignment error values were computed based on original images. We estimated the optimal grid mesh using the synthetic images and warped the original images with the optimized grid mesh to compute the alignment error.

Figure 3 presents six groups of comparative results, and more results are presented in our supplementary material. When colour difference is modest, LSH+MF-W, ROS+MF-W and our method usually produce fewer errors than DF-W, demonstrating that the photometric constraint is beneficial to better alignment quality. However, the errors in LSH+MF-W and ROS+MF-W results become larger as the colour difference increases, as LSH and ROS fail to compensate colour differences between images when the colour variation becomes larger and more complex. In contrast, as the degree of colour variation varies from the smallest to the largest, the alignment error in our method is stable, demonstrating that our proposed method is robust to significant colour differences. Our proposed method combines photometric and geometric constraints to produce better stitching results than DF-W, and it stably obtains good alignment results even for images with complex colour variations.

3.2 Experiments on Real Images

We conducted experiments on real images with significant colour differences, which were directly collected from publicly available datasets [4, 12, 19]. In the experiment, The proposed method was compared with three geometric-based methods (global homography, CP-W [15], and DF-W [12]) as well as three photometric-based methods (MF-W, LSH+MF-W, and ROS+MF-W). Table 1 presents the alignment error results on ten pairs of images. The proposed GCPW method usually yielded the lowest alignment error values.

| | only geometric constraint | | | geometric and photometric constraints | | | |
|-----------|---------------------------|-------|-------|---------------------------------------|----------|-------------|--------------|
| method | Global | CPW | DF-W | MF-W | LSH+MF-W | ROS+MF-W | GCPW |
| 01 | 12.26 | 12.22 | 11.54 | 16.41 | 10.73 | 8.40 | 5.61 |
| 02 | 7.60 | 9.41 | 4.19 | 5.57 | 4.10 | 3.18 | 1.71 |
| 03 | 10.42 | 9.66 | 4.71 | 10.70 | 4.29 | 3.80 | 3.32 |
| 04 | 10.26 | 7.56 | 5.17 | 8.23 | 4.64 | 4.16 | 3.74 |
| 05 | 10.63 | 10.04 | 8.59 | 8.27 | 8.31 | 7.46 | 8.07 |
| 06 | 7.15 | 9.53 | 4.14 | 7.80 | 6.49 | 5.04 | 3.46 |
| 07 | 5.28 | 5.15 | 2.66 | 2.19 | 4.05 | 2.26 | 1.91 |
| 08 | 17.13 | 18.87 | 13.19 | 10.97 | 9.74 | 10.58 | 9.18 |
| 09 | 10.90 | 9.31 | 5.52 | 10.37 | 7.68 | 4.59 | 4.44 |
| 10 | 20.30 | 17.99 | 11.98 | 15.77 | 12.96 | 12.38 | 10.39 |

Table 1: Alignment errors on ten image pairs for each evaluated methods. Our method usually yields the lowest alignment error.



Figure 4: Comparisons with three geometric-based methods: global homography, CPW [15] and DF-W [12]. Yellow arrow indicates the misalignment and ghost artifacts. With the photometric guidance, our GCPW method outperforms other three methods.

Figure 4 presents two groups of result that intuitively compare our method with global homography, CPW, and DF-W. As we can see, the global homography fails to produce satisfactory results because of the model limitation. The CPW and the DF-W still suffer obvious misalignments as they only use geometric constraints to align images locally. In contrast, our method combines geometric constraint with photometric constraint in the proposed GCPW framework. The experimental results show that our method produces the best alignment quality compared with these three geometric-based methods.

In order to further demonstrate the superiority of our method, Figure 5 presents another two groups of result that compare the proposed GCPW with MF-W, LSH+MF-W, and ROS+MF-W. We can observe that in results produced by MF-W, LSH+MF-W and

ROS+MF-W, there are severe ghosts and misalignments. The MF-W fails because it cannot cope with images with huge colour variations. Although LSH+MF-W and ROS+MF-W overcome this limitation through some pre-processing operations, nevertheless, they usually do not compensate the colour difference fully. Therefore, these two methods still suffer from severe misalignment artifacts. In contrast, our method does not adopt any pre-processing operation. It models the colour transformation in the GCPW framework and jointly estimates a local colour model and an optimal grid mesh at the same time. The results demonstrate that our method outperforms other methods when stitching images with significant colour variations. More comparative results are presented in our supplementary material.

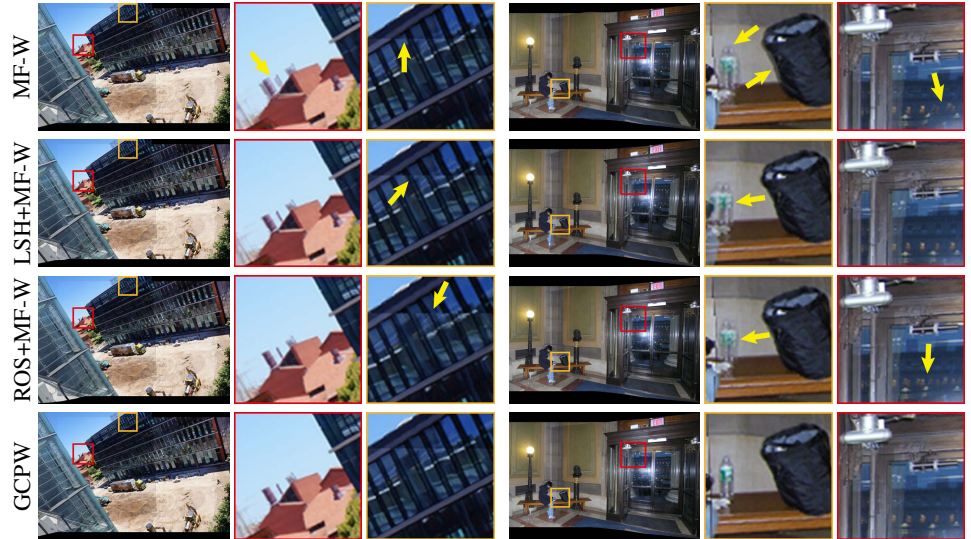


Figure 5: Comparisons with MF-W, LSH+MF-W and ROS+MF-W on real images. Our method produces better stitching result than other three methods.

4 Conclusion

In this paper, we propose the GCPW framework to stitch images with colour variations. The GCPW extends the original CPW by appending a local colour model that expresses the colour transformation from the source image to the target image. This extension makes the photometric constraint robust to significant colour variations and thus can be applied to improve alignment quality. To stitch images with GCPW, we roughly align them with a global homography. Next, we combine geometric and photometric constraints in the proposed GCPW framework, in which the warped mesh vertexes and the affine colour model are optimized jointly. The images are aligned locally based on the warped grid mesh. We conducted experiments on both synthetic and real images. For synthetic experiment, images with different degrees of colour variation are tested, and results demonstrate that our method is robust to significant colour difference. For experiment on real images, although the colour model that applied in GCPW is simple, it is flexible enough to express complex colour transformations that occur in real world. Comparative results show that our method stably produces better stitching results than other state-of-the-art methods.

References

- [1] Simon Baker and Iain Matthews. Lucas-kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56(3):221–255, 2004.
- [2] Matthew Brown and David G Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007.
- [3] Che-Han Chang, Chun-Nan Chou, and Edward Y Chang. Clkn: Cascaded lucas-kanade networks for image alignment. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [4] Yu-Sheng Chen and Yung-Yu Chuang. Natural image stitching with the global similarity prior. In *European Conference on Computer Vision (ECCV)*, pages 186–201, 2016.
- [5] Oliver Demetz, Michael Stoll, Sebastian Volz, Joachim Weickert, and Andrés Bruhn. Learning brightness transfer functions for the joint recovery of illumination changes and optical flow. In *European Conference on Computer Vision (ECCV)*, pages 455–471, 2014.
- [6] Denis Fortun, Patrick Bouthemy, and Charles Kervrann. Optical flow modeling and computation: a survey. *Computer Vision and Image Understanding*, 134:1–21, 2015.
- [7] Heng Guo, Shuaicheng Liu, Tong He, Shuyuan Zhu, Bing Zeng, and Moncef Gabbouj. Joint video stitching and stabilization from moving cameras. *IEEE Transactions on Image Processing*, 25(11):5491–5503, 2016.
- [8] Yoav HaCohen, Eli Shechtman, Dan B Goldman, and Dani Lischinski. Non-rigid dense correspondence with applications for image enhancement. *ACM transactions on Graphics*, 30(4):70, 2011.
- [9] Shengfeng He, Qingxiong Yang, Rynson WH Lau, Jiang Wang, and Ming-Hsuan Yang. Visual tracking via locality sensitive histograms. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2427–2434, 2013.
- [10] Jie Hu, Dong-Qing Zhang, Heather Yu, and Chang Wen Chen. Multi-objective content preserving warping for image stitching. In *IEEE International Conference on Multi-media and Expo (ICME)*, pages 1–6, 2015.
- [11] Kyungdon Joo, Namil Kim, Tae-Hyun Oh, and In So Kweon. Line meets as-projective-as-possible image stitching with moving dlt. In *IEEE International Conference on Image Processing (ICIP)*, pages 1175–1179, 2015.
- [12] Shiwei Li, Lu Yuan, Jian Sun, and Long Quan. Dual-feature warping-based motion model estimation. In *IEEE International Conference on Computer Vision (ICCV)*, pages 4283–4291, 2015.
- [13] Kaimo Lin, Nianjuan Jiang, Loong-Fah Cheong, Minh Do, and Jiangbo Lu. Seagull: Seam-guided local alignment for parallax-tolerant image stitching. In *European Conference on Computer Vision (ECCV)*, pages 370–385, 2016.

- [14] Kaimo Lin, Nianjuan Jiang, Shuaicheng Liu, Loong-Fah Cheong, and Minh Do2 Jiang-bo Lu. Direct photometric alignment by mesh deformation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2405–2413, 2017.
- [15] Feng Liu, Michael Gleicher, Hailin Jin, and Aseem Agarwala. Content-preserving warps for 3d video stabilization. *ACM Transactions on Graphics*, 28(3):44, 2009.
- [16] Jaesik Park, Yu-Wing Tai, Sudipta N Sinha, and In So Kweon. Efficient and robust color consistency for community photo collections. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 430–438, 2016.
- [17] Menghan Xia, Jian Yao, Renping Xie, Mi Zhang, and Jinsheng Xiao. Color consistency correction based on remapping optimization for image stitching. In *IEEE International Conference on Computer Vision Workshop (ICCVW)*, pages 2977–2984, 2017.
- [18] Tianzhu Xiang, Gui-Song Xia, Xiang Bai, and Liangpei Zhang. Image stitching by line-guided local warping with global similarity constraint. *arXiv preprint arXiv:1702.07935*, 2017.
- [19] Julio Zaragoza, Tat-Jun Chin, Michael S Brown, and David Suter. As-projective-as-possible image stitching with moving dlt. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2339–2346, 2013.
- [20] Fan Zhang and Feng Liu. Parallax-tolerant image stitching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3262–3269, 2014.
- [21] Guofeng Zhang, Yi He, Weifeng Chen, Jiaya Jia, and Hujun Bao. Multi-viewpoint panorama construction with wide-baseline images. *IEEE Transactions on Image Processing*, 25(7):3099–3111, 2016.