

EECS 559 Project Report

Robust Global Homography Estimation and Correction

Yingwen Tan and Ziyuan Huang*

April 2023

1 Introduction

Image stitching is one of the most successful applications in Computer Vision that combines multiple photographic images with overlapping fields of view to produce a segmented panorama or high-resolution image. It is a well-studied topic in the field of Computer Vision. The very first image stitching solutions are feature-based methods, proposed by David L. Milgram in 1975 [1], where geometric registration is used to combine overlapping images into a photomosaic. In [2], Yaseen et al applied the direct linear transformation (DLT) method technique to calculate homography based on invariant features given by Scale-Invariant Feature Transform (SIFT) algorithm. Based on the motion model involving line and point, a warping was introduced by Li et al [3] to preserve geometrical and structural information of the scenes. However, these methods depend highly on the quality of feature matching. In this paper, we want to convert the image stitching problem into a non-linear optimization problem. In this piece, we study a nonlinear optimization that gives estimates for the accurate global homography matrix using the potentially inaccurate local homography matrices generated by classical algorithms.

The problem is significant in that it improves the existing methods of local homography estimations. Current panorama construction algorithms rely on pairwise homography transformation, which is myopic and pruned to extreme points and outliers. It also links together the optimization problem and a standard problem in computer vision.

2 Related Works

Image stitching, or the creation of a panorama from a set of overlapping images, is a well-studied topic with widespread applications. Milgram [1] introduced the method of combining a set of overlapping images into a panorama picture by geometric registration. In 2003, Brown et al. [4] uses recognition techniques based on invariant local features to select matching images. However, these methods often fail when processing large parallax and often generate artifacts. This shortness was reduced and refined using local warping guided by local motion by [5]. In 2014, Zhang et al. [6] proposed a hybrid alignment model by combining homographies and content-preserving warping to handle parallax while objectional local distortion is avoided. This paper presents a global stitching method to handle parallax based on the observation that input images do not need to be perfectly aligned over the whole overlapping region for stitching. Li et al. [3] also, present a similar idea by considering both keypoint and line segment correspondence as data-term. The problem with these methods is that with high degrees of freedom, the accuracy of the model heavily relies on feature correspondence. To overcome this issue, the photometric constraint is used in those works. Chen et al. [7] calculated global homography based on line and point features. After that, they use the local warping method to integrate both photometric and geometric constraints. In this paper, we only focus on global homography estimation and propose a related optimization method that attempts to mitigate the misalignments that

*All co-authors equally contributed to the project. Yingwen Tan mainly worked on programming and simulations and Ziyuan Huang mainly worked on idea development and theoretical proofs. The project was written jointly and each person engaged in the writing of all sections.

originated from the drawbacks described before. To our knowledge, it is the first attempt to model the image stitching problem as non-linear optimizations on the stitched plane to compute global homography matrices.

3 Problem Statement

3.1 Homogeneous Coordinate and Homography Transformation

First of all, we need to introduce the geometric transformation that relates the images to mosaics. The homogeneous coordinates or projective coordinates, which was introduced by August Ferdinand Möbius [8] in 1827, represent a point in Cartesian Coordinates(X, Y) using (x, y, w) in Homogeneous Coordinates. And X and Y in Cartesian are re-expressed with x, y and w in Homogeneous as

$$X = \frac{x}{w}, Y = \frac{y}{w}, w \neq 0.$$

In Homogeneous Coordinates, all transformations of 2D images could be written as matrix multiplication.

$$\mathbf{u} = \mathbf{H} \cdot \mathbf{x} \quad (1)$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = H \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

where

$$u = u'/w', v = v'/w', w' \neq 0.$$

Also, notice that any non-zero scaling of a homogeneous coordinate effectively represents the same point on the 2D plane.

3.2 Global Homography Correction

Random Sample Consensus (RANSAC) is a well-known algorithm for pairwise homography estimation. Let I represents the set of images taken from different angles of a specific scene. For each pair of images $(q, t) \in I \times I$, the RANSAC algorithm generates a local homography matrix H_q^t that transforms the image q onto the plane of image t , and a set of matching inlier point correspondences Ψ_q^t . However, this local homography is usually subject to misalignment [9] due to poor generalizability. *Our goal is to obtain a global homography matrix based on the existing pairwise results of RANSAC by solving a generalization-aimed optimization problem.* Our proposal follows similar ideas of the pioneering work by Saha et al. [9] that identifies a global, instead of pairwise, homography matrix by solving a constrained non-convex optimization problem. We first select a subset of image pairs $\Theta \subseteq I \times I$ such that for each $(q, t) \in \Theta$, t is the best-matched image to q . Image t is “best-matched” to q if the average distance between points in their matching inlier point correspondence Ψ_{qt} is the smallest, compared to images other than t . Namely, we attempt to match every image only to its best-matched peer. Thus, with an abuse of notation, we can write $\Theta(q)$ as the best-matched image for the image q . Let X_q^1 and X_q^2 be the data matrices corresponding to the point correspondence $\Psi_q^{\Theta(q)}$. The i -th column of X_q^1 is the i -th feature point of the image q and that of X_q^2 is the i -th feature point of the image $\Theta(q)$. Thus, our goal is to find a set of global homography matrices $\mathcal{H} := \{H_q \in \mathbb{R}^{3 \times 3} : q \in I\}$ by minimizing the following objective

$$J(\mathcal{H}, \mathbf{s}) := \sum_{q \in I} \left\| [H_{\Theta(q)}]^{-1} H_q X_q^1 - X_q^2 \text{diag}(\mathbf{s}) \right\|_F^2 + \lambda \cdot \left\| [H_{\Theta(q)}]^{-1} H_q - H_q^{\Theta(q)} \right\|_F^2, \quad (3)$$

where $\mathbf{s} \in \mathbb{R}^N$ is the vector of scaling factors. The appearance of \mathbf{s} is justified by the fact that any scaling of a homogeneous coordinate actually represents the same point (see previous subsection). This is in general a non-convex optimization problem. Each $H_q \in \mathcal{H}$ denotes homographic transformation from a local image plane to a global image plane (a canvas) shared by all images. A visual explanation is provided in Figure 1

The first term in Eq. (3) can be interpreted as the quality of global image matching, while the second term is the closeness to the local (pairwise) homography matrix. The parameter λ balances the tradeoff between these terms. Since the local homography estimation by RANSAC is likely to contain misalignment, we intend to set λ as small as possible. However, as we will show in later sections, the difference in scales between the first and second term in Eq. (3) raises an issue of λ selection. For example, if the number of training data, N , is very large, the Frobenius norm of the latter term (a 3×3 matrix) is heavily dominated by the former (a $3 \times N$ matrix with $N \gg 3$). In that case, we need to select the value of λ carefully in order to achieve the desired results.

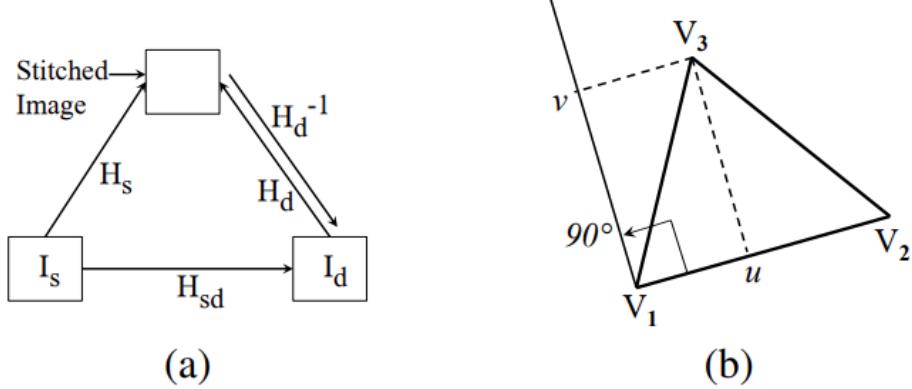


Figure 1: (a) Relation between global and pairwise homography matrices. Here I_s , I_d corresponds to two matched image I_θ^1 , I_θ^2 . Global homography matrix H_{sd} corresponds to H_θ . And pairwise homography matrices H_s , H_d corresponds to H_θ^1 , H_θ^2 . (b) V_3 is the linear combination of point V_1 and V_2 .

We hope to get the global homograph matrix H_θ the features of image I_θ^1 and I_θ^2 from RANSAC. Using the method of non-linear optimization, we can get the "optimal" pairwise homography matrices H_θ^1 and H_θ^2 that meet the constraint.

4 Method

In this section, we propose two optimization reformulations of Eq. (3) as well as associated algorithms to calculate the optimal value. Each reformulation is specifically tied to an application domain of global homography estimation, i.e., image stitching and panorama construction.

4.1 Image Stitching

Panorama problems with two images are also called image stitching problems. We propose an alternating direction method for the reduced problem.

Suppose without generality that $\Theta = \{(1, 2)\}$. Then, denoting $M := H_2^{-1}H_1$. In two-image cases, we may assume $H_2 = I$ without loss of generality. Thus, the global homography matrix of the second image is I and that of the first image is given by M . We consider the following equality constraint problem

$$\min_{M, s} \|MX_1^1 - X_1^2 \text{diag}(s)\|_F^2 + \lambda \|M - H_1^2\|_F^2, \quad (4)$$

where $s \in \mathbb{R}^N$ is the scaling factor associated with the homogeneous coordinate.

Notice that this problem is not in general convex, but we can observe that it is individually convex (even strongly convex) for each of the optimization variables involved. This inspired us to perform iterative updates to variables M and s . The idea is to update each variable individually while keeping the other one fixed. This yields a sequence of very easy-to-solve subproblems and we hope that it converges to some optima of the overall problem and is also fast enough. To update the manufactured matrix $M^{(k+1)}$, we calculate the partial derivative of the objective function w.r.t. M , i.e.,

$$2(MX_1^1 - X_1^2 \text{diag}(s))(X_1^1)^T + 2\lambda(M - H_1^2). \quad (5)$$

Thus, by setting it to zero, we obtain

$$M^{(k+1)} = [X_1^2 \operatorname{diag}(\mathbf{s}^{(k)})(X_1^1)^T + \lambda H_1^2](\lambda I + X_1^1(X_1^1)^T)^{-1}. \quad (6)$$

Fix $M^{(k+1)}$, the objective function is separable w.r.t. the scaling factor \mathbf{s} , i.e.,

$$\min_{\mathbf{s}} \sum_{i=1}^N \|\mathbf{a}_i - \mathbf{x}_i s_i\|_2^2, \quad (7)$$

where \mathbf{a}_i and \mathbf{x}_i are respectively the i -th columns of the matrix $M^{(k+1)}X_1^1$ and X_1^2 . In light of this, we derive a closed form of the update on \mathbf{s} :

$$\mathbf{s}^{(k+1)} = \operatorname{diag}\left((X_1^1)^T M^T X_1^2\right) ./ (\|X_1^2\|_2^2)_{i=1}^N, \quad (8)$$

where $./$ denotes element-wise division and $A_{:,i}$ stands for the i -th column of the matrix A .

Proposition 4.1. *Fixing \mathbf{s} , M given by Eq. (6) is a minimizer. Fixing M , \mathbf{s} given by Eq. (8) is also a minimizer.*

Proof. The first claim is clearly supported by the fact that the objective function is convex in M and $M^{(k+1)}$ is obtained by setting the gradient w.r.t. M to zero. To show the second claim. We again take the gradient w.r.t. \mathbf{s} of the equivalent problem in Eq. (7) and set to zero, i.e.,

$$\mathbf{x}_i^T (s_i \mathbf{x}_i - \mathbf{a}_i) = 0 \implies s_i = \frac{\mathbf{x}_i^T \mathbf{a}_i}{\|\mathbf{x}_i\|_2^2}. \quad (9)$$

By stacking them into a column vector, we arrive at the closed form in Eq. (8). \square

Therefore, we can construct an alternating-direction algorithm for solving this simple image-stitching example.

Algorithm 1 An alternating direction algorithm for image stitching problem.

```

1: procedure IMAGESTITCHESTIMATION( $X_1^1, X_1^2, H_1^2$ )
2:   Initialize  $M^{(0)}$  and  $\mathbf{s}^{(0)}$ .
3:   for  $k = 1, 2, \dots$  until converge do
4:     Compute  $M^{(k+1)}$  using Eq. (6).
5:     Compute  $\mathbf{s}^{(k+1)}$  using Eq. (8).
6:   end for
7:   Recover global homography matrices by  $H_1 \leftarrow M$  and  $H_2 \leftarrow I$ .
8: end procedure

```

4.2 Panorama Construction

We extend the two-image stitching algorithm 1 to multi-image stitching problems or panorama construction problems.

We introduce auxiliary matrices M_1, \dots, M_N . For each $i \in I$, $M_i = (H_{\Theta(i)})^{-1} H_i$. Then, the problem is equivalent to the following,

$$\min_{(M_i, \mathbf{s}_i)_{i=1}^N} \sum_{i \in I} \|M_i X_i^1 - X_i^2 \operatorname{diag}(\mathbf{s}_i)\|_F^2 + \lambda \cdot \|M_i - H_i^{\Theta(i)}\|_F^2 \quad \text{s.t. } M_i = (H_{\Theta(i)})^{-1} H_i, \forall i \in I. \quad (10)$$

The basic idea is to ignore all the equality constraints and solve N optimization problems w.r.t each pair of (M_i, \mathbf{s}_i) , since Eq. (10) is clearly separable in each optimization variable. Without loss of generality, we assign the identity global homography matrix to the first image, and all other global homography matrices can be recovered by matrix multiplication (as is explained in Figure 1-(a)). In algorithms,

Algorithm 2 An Extended Alternating Direction Methods for Panorama Construction

```

1: procedure PANORAMACONSTRUCTION( $(X_i^1, X_i^2, H_i^{\Theta(i)})_{i=1}^N$ )
2:   Initialize  $(M_i^{(0)})_{i=1}^N$  and  $s_i^{(0)}$ .
3:   for  $k = 1, 2, \dots$  until converge do
4:     for  $i = 1, \dots, N$  do
5:       Compute  $M_i^{(k+1)}$  using Eq. (6)
6:       Compute  $s_i^{(k+1)}$  using Eq. (8)
7:     end for
8:   end for
9:   Recover global homography matrices by the recurrence  $H_1 \leftarrow I$  and  $H_i \leftarrow H_{\Theta(i)} M_i$ .
10: end procedure

```

5 Convergence Analysis

The convergence of both application domains relies on the convergence of the alternating direction method described by Eq. (6) and (8). Let's denote the objective function as $f(M, s)$. Then, f is obviously an individually convex function for each M and s . In fact, it is strongly convex for M and, with high probability, for s . The strong convexity for s requires the sufficient and necessary condition that $[X_1^2]_{:i} \neq \mathbf{0}$ for all i , which holds in almost all situations. According to Proposition 4.1, both updates are optimal given the other variable fixed. Then, the proof of convergence is straightforward.

Proposition 5.1. *Both algorithms 1 and 2 converge to local minima.*

Proof. . Let $\mathbf{z}_0 := (M^{(0)}, s^{(0)})$. Define $\mathbf{z}_{2k+1} := (M^{(k+1)}, s^{(k)})$ and $\mathbf{z}_{2k+2} := (M^{(k+1)}, s^{(k+1)})$ as the sequence of iterates generated by Eq. (6) and (8). Then, by the optimality condition for both updates, we have $f(\mathbf{z}_{2k+2}) \leq f(\mathbf{z}_{2k+1}) \leq f(\mathbf{z}_{2k})$. Since $f(\mathbf{z})$ is bounded below by 0, this sequence always has a convergent subsequence. The limit of convergence is in general a stable point, but in light of the strong convexity in both directions, it has to be a local minimum. \square

We now show that the algorithms converge faster than $\mathcal{O}(1/k)$.

Proposition 5.2. *The convergence is faster than $\mathcal{O}(\frac{1}{k})$.*

Proof. We reexamine equations (7), (8), and (9). For each i , we have

$$s_i \mathbf{x}_i = \frac{\mathbf{a}_i^T \mathbf{x}_i}{\|\mathbf{x}_i\|_2^2} \mathbf{x}_i = \mathbf{a}_i^T \left(\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|_2} \right) \frac{\mathbf{x}}{\|\mathbf{x}_i\|_2}, \quad (11)$$

which is the projection of the vector \mathbf{a}_i onto the direction $\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|_2}$. Let $\mathcal{P}_i(\mathbf{a}_i)$ be this projection operator and $\mathcal{P}(A) := [\mathcal{P}_i(\mathbf{a}_i)]_{i=1}^N$ be the horizontal stack of those, where \mathbf{a}_i is the i -th column of the matrix A . Let $\mathcal{Q}(A) := A - \mathcal{P}(A)$ be the projection operator onto the orthogonal space.

Consider the objective values at the iterates $\{(M^{(k)}, s^{(k)})\}_{k \in \mathbb{N}}$ and denote this sequence of objectives as $\{f_k\}_{k \in \mathbb{N}}$. Then, f_k can be computed as

$$f_k = \left\| \mathcal{Q}(M^{(k)} X_1^1) \right\|_F^2 + \lambda \left\| M^{(k)} - H_1^2 \right\|_F^2. \quad (12)$$

It can be viewed as a minimization problem over projections with the Frobenius penalty. Due to the non-expansiveness of projection operators, the gradient of the function $f : M \mapsto \left\| \mathcal{Q}(MX_1^1) \right\|_F^2 + \lambda \left\| M - H_1^2 \right\|_F^2$ possesses a Lipschitz constant $L < \infty$. Even if the objective is non-convex, the smoothness of the gradient still implies the convergence of the series $\sum_{i=1}^N \left\| \nabla f(M^{(k)}) \right\|_F < \infty$ (we have seen it in HW2). Therefore, the convergence is faster than $\mathcal{O}(\frac{1}{k})$. \square

In fact, the function f defined in the proof above is locally strongly convex near any local minima. Thus, the convergence can be actually boosted to linear convergence or even faster. Most of our simulations show

that both the algorithms actually converge in less than 10 iterations, which coincides with common intuition of the speed of coordinate descent methods.

6 Evaluation Metric

We evaluate our model using the difference in signal strengths on the stitched plane. Given H_1, \dots, H_N as the learned global homography transformation that map images onto the stitched plane. Let $I_i(\mathbf{u})$ be the original pixel intensity of the i -th image at location \mathbf{u} . Then, supposing \mathbf{x} is a location on the stitched plane, then the transformed pixel intensity of the i -th image at that location is $I_i(H_i^{-1}\mathbf{x})$. The idea of the evaluation metric is to measure the difference in pixel intensities on all overlapping regions on the stitched plane. Namely, denoting \mathcal{K}_{ij} as the set of overlapping points between image i and j , the error metric is given by

$$E(H_1, \dots, H_N) = \sum_{i,j \in I \times I} \sum_{\mathbf{x} \in \mathcal{K}_{ij}} \rho(I_i(H_i^{-1}\mathbf{x}) - I_j(H_j^{-1}\mathbf{x})), \quad (13)$$

where $\rho(\cdot)$ is a difference metric. Typical choices of $\rho(\cdot)$ include squared difference $(\cdot)^2$ and absolute difference $|\cdot|$, but both of them, especially the former, are very sensitive to outliers. This is detrimental as in signals in images are very complex and noisy. Here we adopt a solution from Black and Rangarajan [] who discussed a variety of noise-robust difference metrics, including the Geman-McClure function

$$\rho_{GM}(x) = \frac{x^2}{1 + x^2/a}, \quad (14)$$

where a is the *outlier threshold* parameter. This function mimics the growth dynamics that behave like quadratic functions for small values but grows increasingly slowly when x moves towards large displacement from the origin. The appropriate value of a is derived using statistics estimations according to [], i.e., the median of absolute differences multiplied by 1.4. We will take this function as our error metric $\rho(\cdot)$.

7 Numerical Result

We adopt the As-Projective-As-Possible Image Stitching dataset [10]. It consists of 11 scenes, each of which contains at least 2 images taken from different angles. We split the dataset into two categories according to our two application tasks. For the image stitching problem, we used scenes of *boat*, *bridge*, *prague*, and *uttower*; and for the panorama construction problem, we used scenes of *hill*, *ledge*, and *pier*. For the panorama problem, we only selected scenes with 3 images due to computing power constraints. We provide an overview of all these images in Figure 2.

For the input of the RANSAC algorithm, We use the ORB algorithm for general feature point detection and brute force matching with cross-checks to generate initial matches.

The device we utilized is the standard CPU on Google Colab with a maximum of 12.7GB RAM. Our code is available at [GitHub link](#).

7.1 Image Stitching Problems

For the image stitching problem, we compare the error metrics between our learned global homography matrices (“Correction”) and those obtained from classical RANSAC algorithms. The comparison table is listed as follows.

Table 1: Image Stitching Error and Values of Parameter λ .

Methods	<i>boat</i>	<i>bridge</i>	<i>prague</i>	<i>uttower</i>
λ	0.2	0.2	0.7	0.2
RANSAC	40.53	15.67	49.93	108.31
Correction	40.29	12.56	28.44	106.63

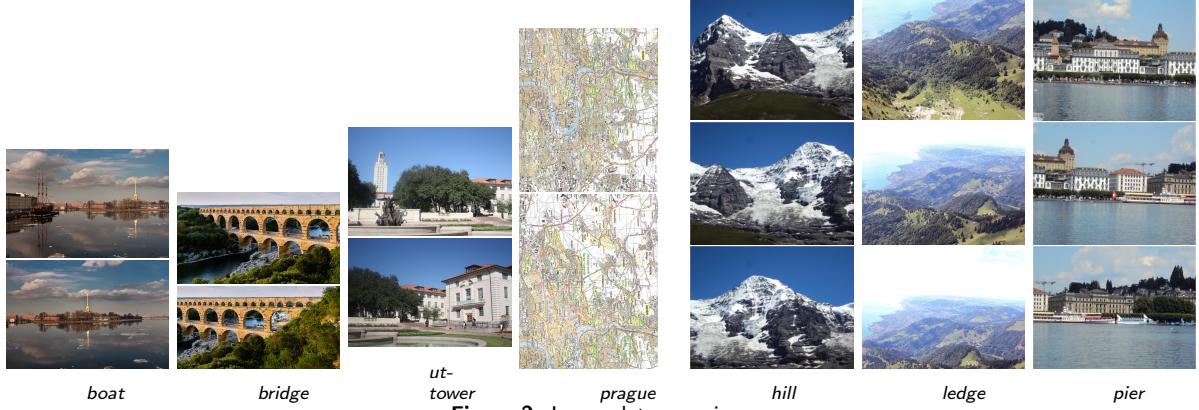


Figure 2: Image data overview.

We see from Table 1 that the correction of global homography matrices indeed improves the alignment for most of the scenes, although the amount of improvement may not be significant (the smallest numerical improvement is only 0.24 for the scene *boat*). We demonstrate visually the results of *bridge* and *prague* in Figure 3 and 4. The stitching looks very seamless, though there are still visible misalignments at the stitching boundary. We did not recognize significant visual improvements in the stitching results in addition to the reported numbers, which aligns with the previous observation that the improvements are not very significant.

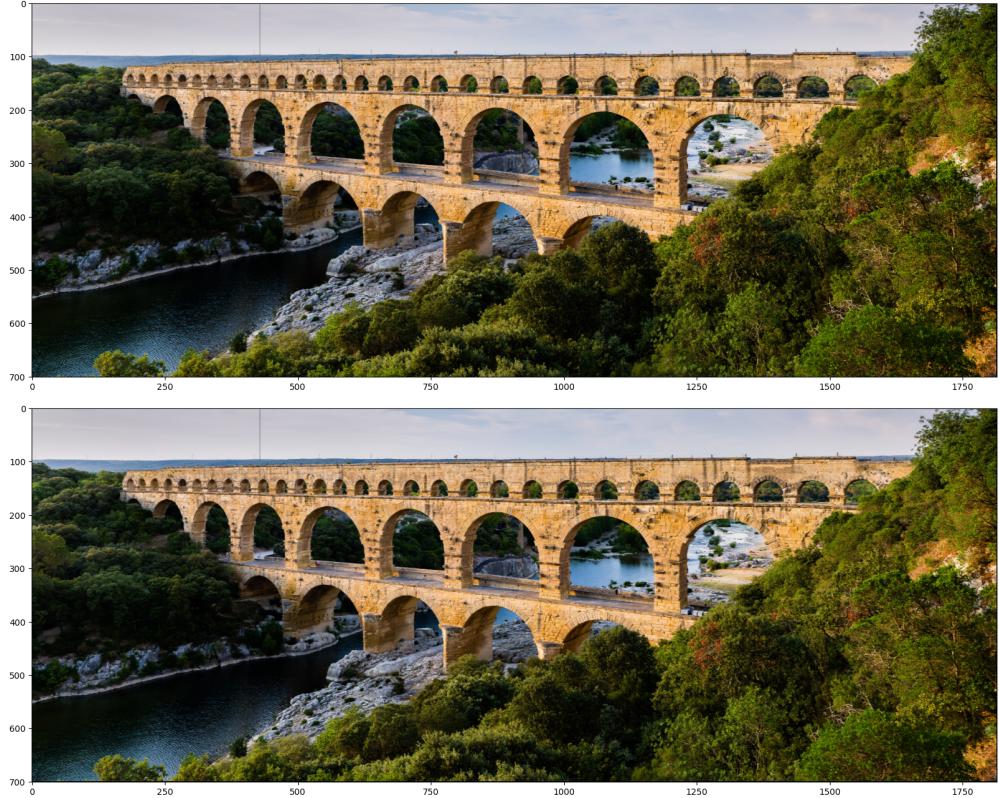


Figure 3: Comparison between RANSAC stitching (above) and our proposed stitching (below) on *bridge*. Both stitchings look almost the same. Thus, our proposed correction method does not demonstrate significant visual improvements in the stitching task.

Another important observation is that the quality improvement depends largely on the parameter λ . For example, the *prague* scene achieves a much lower error rate when λ is set to 0.7 instead of standard 0.2 in which case the correction error is over 49. These situations frequently occur in bright images or image pairs with a large number of matched feature points. We conjectured that it is because both cases make the first

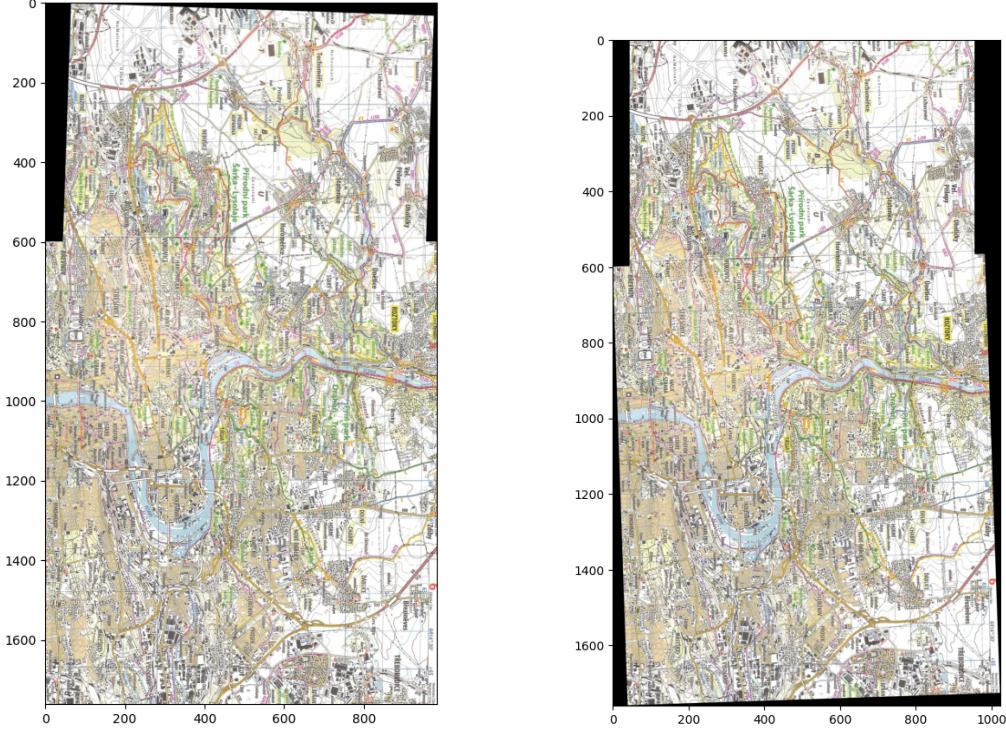


Figure 4: Comparison between RANSAC stitching (left) and our proposed stitching (right) on *prague*. Both images display similar quality albeit different orientations. The intersection of image boundaries in the left figure looks less natural than that in the right one. Despite this minor difference, both images look exactly the same.

term in Eq. (4) so large that the influence of the second (penalty) term is effectively ignored. Thus, we sometimes need to raise the value of λ to account for these issues.

8 Panorama Construction

We performed panorama construction for scenes of *hill*, *ledge*, and *pier* shown in Figure 2. We perform the same comparisons with the RANSAC algorithm as in the previous section.

Table 2: Panorama Construction Error and Values of Parameter λ .

Methods	<i>hill</i>	<i>ledge</i>	<i>pier</i>
λ	0.6	1.2	0.7
RANSAC	83.83	102.24	18.26
Correction	19.44	17.40	17.53

We demonstrate all three recovered panoramas and their quality of stitching in Figure 5.

Notice that we observe very similar results as in the image stitching problem. We indeed see improvements using our proposed inlier correction methods. Sometimes these improvements are very significant, e.g., *hill* and *ledge*. We attach particular interest in this significance of improvements. We consider this phenomenon a success of the global homography estimation. The RANSAC algorithm provides a very good estimate for every pair of images. That's the reason why our inlier correction did not significantly improve the image stitching task. However, in the context of panorama construction, some regions could be overlapped by more than two images and thus the pairwise homography estimation seems too myopic. Therefore, our correction methods are able to outperform the RANSAC algorithm on this task by taking all of the global homography matrices, i.e., H_q for $q \in I$, into account.

Unfortunately, those improvements are subject to appropriate choices of the parameter λ since our experiments show that all scenes can reach a higher error with a different λ value. Nevertheless, the choices

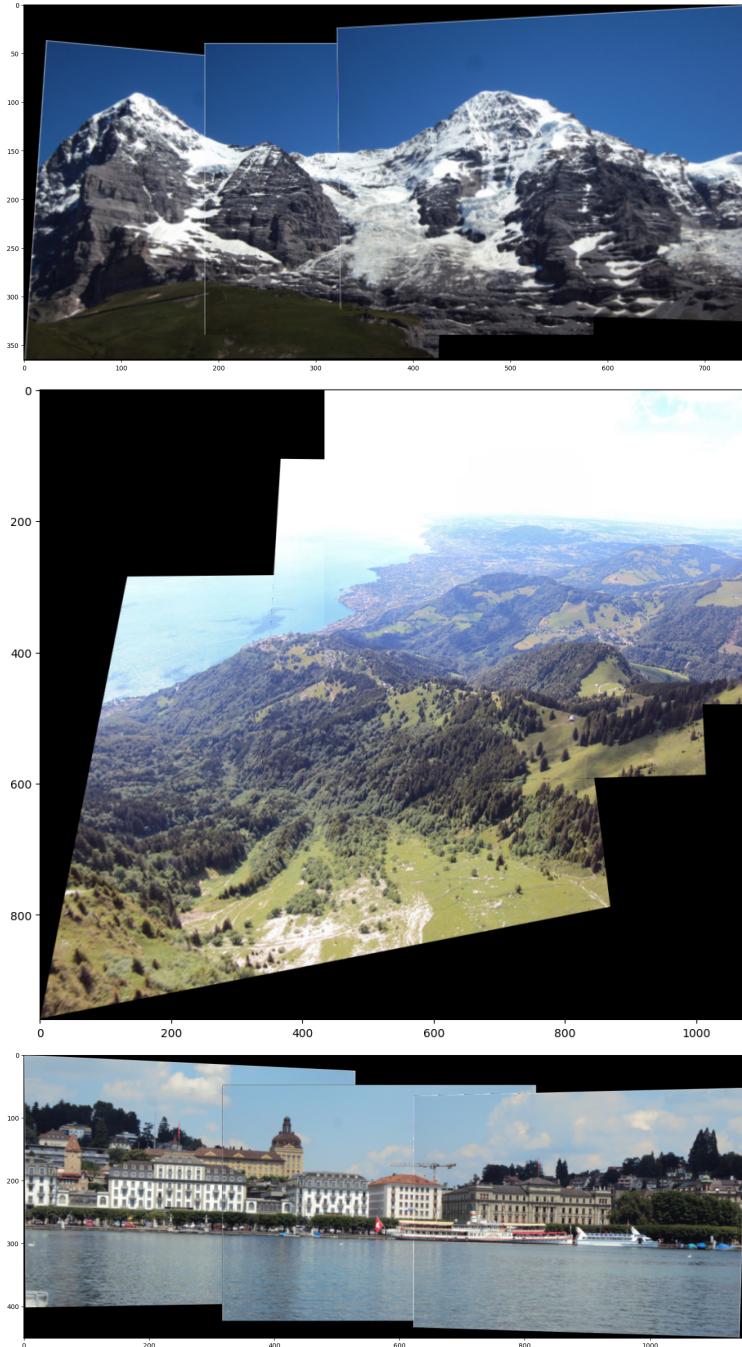


Figure 5: Demonstration of stitched panoramas and quality comparisons.

of λ demonstrated in Table 2 are moderate and reasonable because the highest selected λ (1.2) is not even high enough to make the penalty term comparable to the data fidelity term in Eq. (3).

9 Discussion and Conclusion

9.1 Conclusion

This work tries to calculate global homography matrices based on pairwise local homography estimation. We propose an optimization problem that utilized inlier feature points to correct or rectify the local homography estimations on the stitched image plane. We also design an alternating direction method that solves the optimization problem and converges in at most $\mathcal{O}(\frac{1}{\varepsilon})$ iterations. We evaluate the resulting global homography estimations using outlier-invariant distance metrics, i.e., Geman-McClure function. Our proposed method yields solutions that outperform the classical RANSAC algorithm in panorama-construction applications, but only achieves a few improvements in image-stitching applications.

9.2 Discussion

There are also limitations to our proposed methods. The following list showcases a few of those.

- The estimated global homography matrices only perform perspective transformations on the images. Thus, the resulting panorama is only of acceptable quality when the lighting conditions associated with the stitched images are the same, e.g., see the second scene *ledge* in Figure 5. However, when the lighting conditions are different, the image boundaries are easily recognized such as the first and the third images in Figure 5. Future extensions should absorb the pixel intensity transformations into the objective function.
- We showed every stationary point of the objective function is a local minimum. However, it is difficult to characterize the number of local minima. As a result, the actual solution obtained is subject to the initialization of the algorithms. For plenty of time in our simulations, we need to run the same code with random initialization multiple times to obtain an improved result. One solution to it, as is tested to be efficient in our experiments, is to initialize the matrix M in algorithms 1 and 2 as a random perturbation of the local homography matrix $H_q^{\Theta(q)}$. This initialization is reasonable since we are seeking an inlier-corrected version of $H_q^{\Theta(q)}$.
- Another limitation is the dependence of the method on the value of λ . One may need to fine-tune this hyper-parameter for optimal improvement. This could hinder its application to tasks that require fast and accurate online computations, such as photo auto-processing on smartphones and street view construction algorithms on autonomous driving cars. One direction of extension is to exploit the ADM algorithm since it can be adapted to distributed systems. Another way is to estimate the optimal choice of λ . Based on our conjectures, the more the matched feature points and the brighter the pixel intensities, the higher λ should be. Based on these heuristics, one may develop a statistical estimation of the optimal value of λ , saving tons of time compared to typical hyper-parameter tuning techniques such as grid search.

9.3 Future Works

In addition to the model extensions discussed above, additional work is required in this area to broaden the applicability of global homography estimations. One important direction is to incorporate the stitching quality into the objective function such as photometric accuracy and geometric smoothness. All these metrics guarantee a smooth and natural view of the stitched images or constructed panoramas. Saha [?] has already taken this attempt by conducting bi-level optimization with the global homography estimation as the first-level optimization and the stitching quality as the second-level optimization. Although they achieved incremental success compared to pure RANSAC estimations, additional attention should be paid to the efficiency and convergence properties of this kind of combination.

References

- [1] D. Milgram, "Computer Methods for Creating Photomosaics," *IEEE Transactions on Computers*, vol. C-24, no. 11, pp. 1113–1119, Nov. 1975, conference Name: IEEE Transactions on Computers.
- [2] Yaseen, O.-J. Kwon, J. Lee, F. Ullah, S. Jamil, and J. S. Kim, "Automatic sequential stitching of high-resolution panorama for android devices using precapture feature detection and the orientation sensor," *Sensors*, vol. 23, no. 2, 2023. [Online]. Available: <https://www.mdpi.com/1424-8220/23/2/879>
- [3] S. Li, L. Yuan, J. Sun, and L. Quan, "Dual-Feature Warping-Based Motion Model Estimation," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 4283–4291, iSSN: 2380-7504.
- [4] Brown and Lowe, "Recognising panoramas," in *Proceedings Ninth IEEE International Conference on Computer Vision*, Oct. 2003, pp. 1218–1225 vol.2.
- [5] H.-Y. Shum and R. Szeliski, "Construction and refinement of panoramic mosaics with global and local alignment," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, Jan. 1998, pp. 953–956.
- [6] F. Zhang and F. Liu, "Parallax-Tolerant Image Stitching," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 3262–3269, iSSN: 1063-6919.
- [7] K. Chen, J. Tu, B. Xiang, L. Li, and J. Yao, "Multiple Combined Constraints for Image Stitching," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Oct. 2018, pp. 1253–1257, iSSN: 2381-8549.
- [8] A. Möbius and J. A. Barth, *Der barycentrische Calcul ein neues Hülfsmittel zur analytischen Behandlung der Geometrie dargestellt und insbesondere auf die Bildung neuer Classen von Aufgaben und die Entwicklung mehrerer Eigenschaften der Kegelschnitte angewendet von August Ferdinand Möbius Professor der Astronomie zu Leipzig.* verlag von Johann Ambrosius Barth, 1827. [Online]. Available: <https://books.google.com/books?id=ahsoar-z7Z8C>
- [9] A. Saha, S. Maity, and B. Bhowmick, "Multi-modal Image Stitching with Nonlinear Optimization," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 1987–1991, iSSN: 2379-190X.
- [10] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving dlt," in *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 2339–2346.