Master's Thesis

Modeling of universal stitching parameters for feature-based image stitching algorithm

You-Jin Ha

Department of Biomedical Engineering

Graduate School of UNIST

2019

Modeling of universal stitching parameters for feature-based image stitching algorithm

You-Jin Ha

Department of Biomedical Engineering

Graduate School of UNIST

Modeling of universal stitching parameters for feature-based image stitching algorithm

A thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

You-Jin Ha

06. 12. 2019

Approved by

____

Advisor

Hyun-Wook Kang

Modeling of universal stitching parameters for feature-based image stitching algorithm

You-Jin Ha

This certifies that the thesis/dissertation of You-Jin Ha is approved.

06/12/2019

signature

Advisor: Hyun-Wook Kang

signature

Thesis Committee: Hyun-Deok Kang

signature

Thesis Committee: Semin Lee

ABSTRACT

Image stitching is a well-known method to make panoramic image which has a wide field-of-view and high resolution. It has been used in various fields such as digital map, gigapixel imaging, and 360-degree camera. However, commercial stitching tools often fail, require a lot of processing time, and only work on certain images. The problems of existing tools are mainly caused by trying to stitch the wrong image pair. To overcome these problems, it is important to select suitable image pair for stitching in advance. Nevertheless, there are no universal standards to judge the good image pairs. Moreover, the derived stitching algorithms cannot be compatible with each other because they conform to their own available criteria.

Here, we present universal stitching parameters and their conditions for selecting good image pairs. The proposed stitching parameters can be easily calculated through analysis of corresponding features and homography, which are basic elements in feature-based image stitching algorithm. In order to specify the conditions of the stitching parameters, we devised a new method to calculate stitching accuracy for qualifying stitching results into 3 classes; good, bad, and fail. With the classed stitching results, the values of the stitching parameters could be checked how they differ in each class. Through experiments with large datasets, the most valid parameter for each class is identified as filtering level which is calculated in corresponding feature analysis. In addition, supplemental experiments were conducted with various datasets to demonstrate the validity of the filtering level. As a result of our study, universal stitching parameters can judge the success of stitching, so that it is possible to prevent stitching errors through parameter verification test in advance. This paper can greatly contribute to guide for creating high performance and high efficiency stitching software by applying the proposed stitching conditions.

Keywords: Image Stitching, Feature-based, Panorama, Stitching Parameter, Image Pair

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

With the development of various imaging devices, we can obtain a variety of images that cannot be seen by the human eye or acquired by conventional pinhole cameras. In addition, the user can create desired images through the post-processing. Among many image processing technologies, image stitching can produce a wide field-of-view and high resolution image. Image stitching is a powerful technique for creating panoramic images while maintaining the quality of the original images.

Figure 1.1 Panoramic image generation by image stitching

There are two methods for image stitching, Direct (pixel-based) method and Feature-based method [1, 2]. The goal of both methods is to align the overlapping areas well and combine the given images, but the principle and application field are different. In **Table 1**, there is a comparison of the two methods, and note how the applications vary according to the advantages of each method.

Depending on the strengths and weaknesses of the two methods, the application fields are quite different. Our research has adopted a feature-based method to deal with images acquired with more general photography. In feature-based image stitching, it is most important to determine the correct image pair among the given images. Selecting the correct image pairs and aligning them in the right order in one coordinate system is also known as image registration. Image registration directly affects stitching results, and its feature-based approach has several constraints. The constraints are that the image pairs have enough common areas to each other, and the parallax is small. We defined these constraints as stitching parameters and presented their conditions. The proposed stitching parameters can be computed through the basic elements of feature-based image stitching so that they can be used universally. Therefore, the universal stitching parameters are widely used for judging whether a given image pair can be stitched. The use of universal stitching parameters makes it possible to reduce stitching errors by eliminating image pairs that were judged to fail stitching in advance. Moreover, the conditions of the stitching parameters are expected to be utilized as basic conditional statements in other stitching challenges.

Table 1 Comparison of image stitching methods: Direct method vs Feature-based method

	Direct (pixel-based) method	Feature-based method
Principle	Minimizing the pixel-to-pixel dissimilarity	Extracting corresponding features and deriving the relationship between two images
Advantage	• Make the best use of available information (e.g. camera pose when acquiring images) • Not affected by image features	• Works with images without any information • Amount of computation is relatively small
Drawback	• Limited range of convergence • Not invariant to image scale and rotation	• Must have enough features • Confused in too textural region • Corresponding features should be evenly distributed, and enough region must be shared between image pairs • Relies on robust feature extraction and matching schemes
Application field	• Position fixed multi-camera device • Microscope (but you need to know the camera's environment such as camera pose, magnification, and motion model)	• General photography (but images must have enough overlapping areas with sufficient features) • Microscope

(a) Direct method (b) Multi-camera device (c) Feature-based method

Figure 1.2 Schematic diagram of image stitching method and application field

1.1. Purpose of research

Our research covers the creation of panoramas from multiple images captured in various cameras (or shooting environments) and examines feature-based image stitching techniques. Today, feature-based image stitching algorithms have evolved to produce good results in challenging environments. However, stitching technologies are still difficult to apply to commercial products because of their slow processing speed and huge amount of computation. To overcome the problem, commercially available stitching techniques used to limit camera movement or fix the camera placement. For imposing these constraints, we focused on what criteria could determine the proper movement and position of camera. As an example of the criteria (or guide) for proper movement for stitching, the camera should rotate within 40 degrees to the left. Thus, in order to specify restrictions on the camera, the criteria should be defined based on universal parameters since different types of cameras have to be calibrated.

Meanwhile, in the case of panorama software, it is important to determine whether given image pairs are suitable for stitching. If we can know in advance that a given image pair is not appropriate for stitching, it can reduce stitching errors and increase stitching efficiency. Therefore, we need new criteria to indicate whether stitching is possible, and the criteria should be based on universal parameters that can be used as useful conditional statements in other software. Thus, the goal of our study is to define the universal stitching parameters and present their conditions.

1.2. Overview: Feature-based image stitching

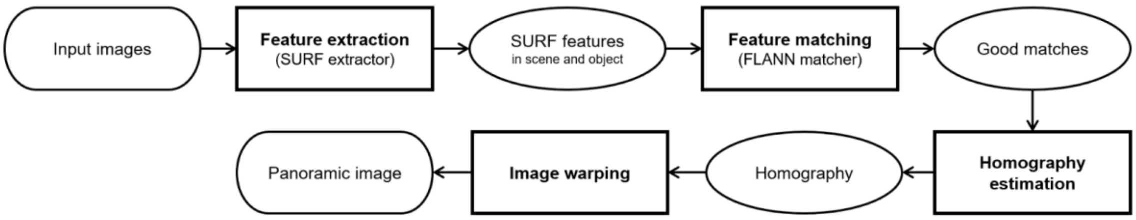
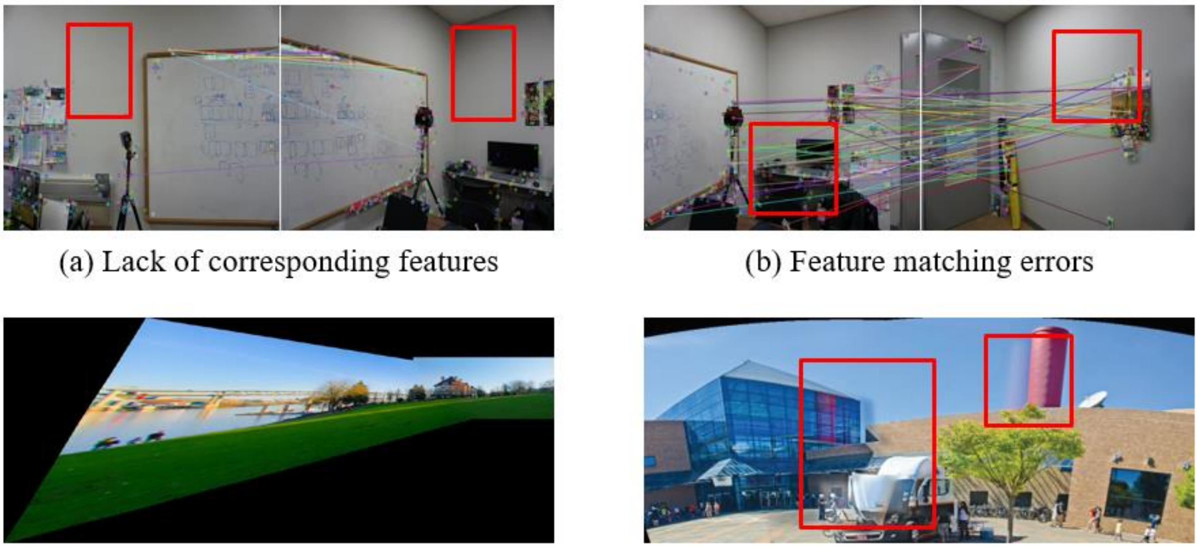


Figure 1.3 Feature-based image stitching algorithm

Feature-based image stitching includes various technologies of computer vision such as feature extraction, matching, 2D transformation [3], and image warping [4]. Each technique has a direct impact on feature-based image stitching results. In other words, in order to obtain high-quality results, appropriate feature extraction, feature matching, and image warping through 2D transformation should be performed. However, there is no information or guidance on the elements for stitching, and many issues arise due to the nature of the feature-based approach. The challenges currently reported are:

- Lack of features (e.g. plain wall)
- Lack of local features (e.g. textural region)
- Insufficient corresponding features
- Feature matching errors
- Optimal warping problems
- Parallax handing



(a) Lack of corresponding features

(b) Feature matching errors

(c) Optimal warping problems

(d) Parallax handing

Figure 1.4 Challenges of feature-based image stitching

Due to the inherent characteristics of feature-based image stitching, it often fails if the available features are not sufficient. Feature-based image stitching failures also occur for the reasons of matching errors and warping problems, even if the features are sufficient. Recent studies on image stitching [5-18] focus on warping problems or 2D transformations. [14] explains that global transformation (or homography) is not suitable for real image acquisition conditions (e.g. casual camera motions, taken from various perspectives, large depth change, etc.). They proposed perspective-preserving warping by combining local projective transformations and similarity transformations which avoid perspective distortions. Optimal warping is directly related to parallax handing. In [15], they also use two types of warping that combined homography and content-preserving warping. These advanced stitching techniques are great achievements in solving some of the challenges. However, each technique applied their own cost function to determine the appropriate warping method. The application of the own criteria has the limitation that it cannot be easily implemented or applied by other researchers.

2. Theory of Feature-based Image Stitching

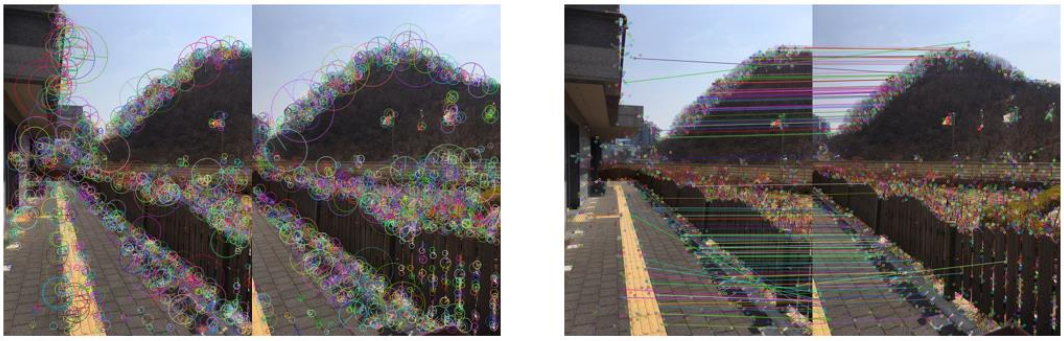
Given a sequence of consecutive images, multiple image stitching combines pairs of images with the most overlapping regions. Thus, it is very important to select the correct image pair in a given sequence of images, which is the key to improving the quality of the result and the efficiency of the work. Section 2 describes the basic form of the feature-based image stitching algorithm (**Section 2.1**) and proposes stitching parameters to determine the correct image pair (**Section 2.2**).

2.1. Feature-based image stitching algorithm

Feature-based image stitching algorithm is divided into three steps as follows [19].

1) Feature extraction and matching
2) Homography estimation
3) Image warping

The first step is to fine the overlapping area (or common scene between two images). This step begins with extracting features from each image. The feature (or keypoint) is the image pattern which differs from its immediate neighborhood. There are many feature detectors or extractors, which are well documented in "Local Invariant Feature Detectors: A survey" [20]. For image stitching, SIFT (Scale Invariant Feature Transform) [21], SURF (Speeded Up Robust Features) [22], ORB (Oriented FAST and Rotated BRIEF) [23] are mostly used as feature extractors because they can robustly extract features which are invariant to translation, rotation, scale, and illumination. After feature extraction is complete, all features should be matched each other to find common features. The matching process popularly done by FLANN (Fast Library for Approximate Nearest Neighbors), open source library for nearest neighbor matching [24]. Through this matching process, we get many matched features. In order to get more reliable matched features, the user must define "Good matches".



(a) Feature extraction (b) Feature matching

Figure 2.1 Feature extraction and matching (step 1)

The next step is estimating the homography through the good matches. The planar homography is a non-singular linear relationship between points on planes and it plays an important role in the geometry of multiple views [3]. The homography determines how to overlap images together for stitching. It is a 3×3 2D transformation matrix expressed in a homogeneous form in Eq. (1). The first 2×2 submatrix of the Eq. (1) represents the rotation, scale, shearing, and reflection, $[h_3,\ h_6]$ represent translation, $[h_7,\ h_8]$ represent a perspective change. At least four corresponding points are required to estimate the homography, and the algorithm RANSAC (random sample consensus) is popularly used for good homography estimation [25].

$$\mathrm{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \tag{1}$$

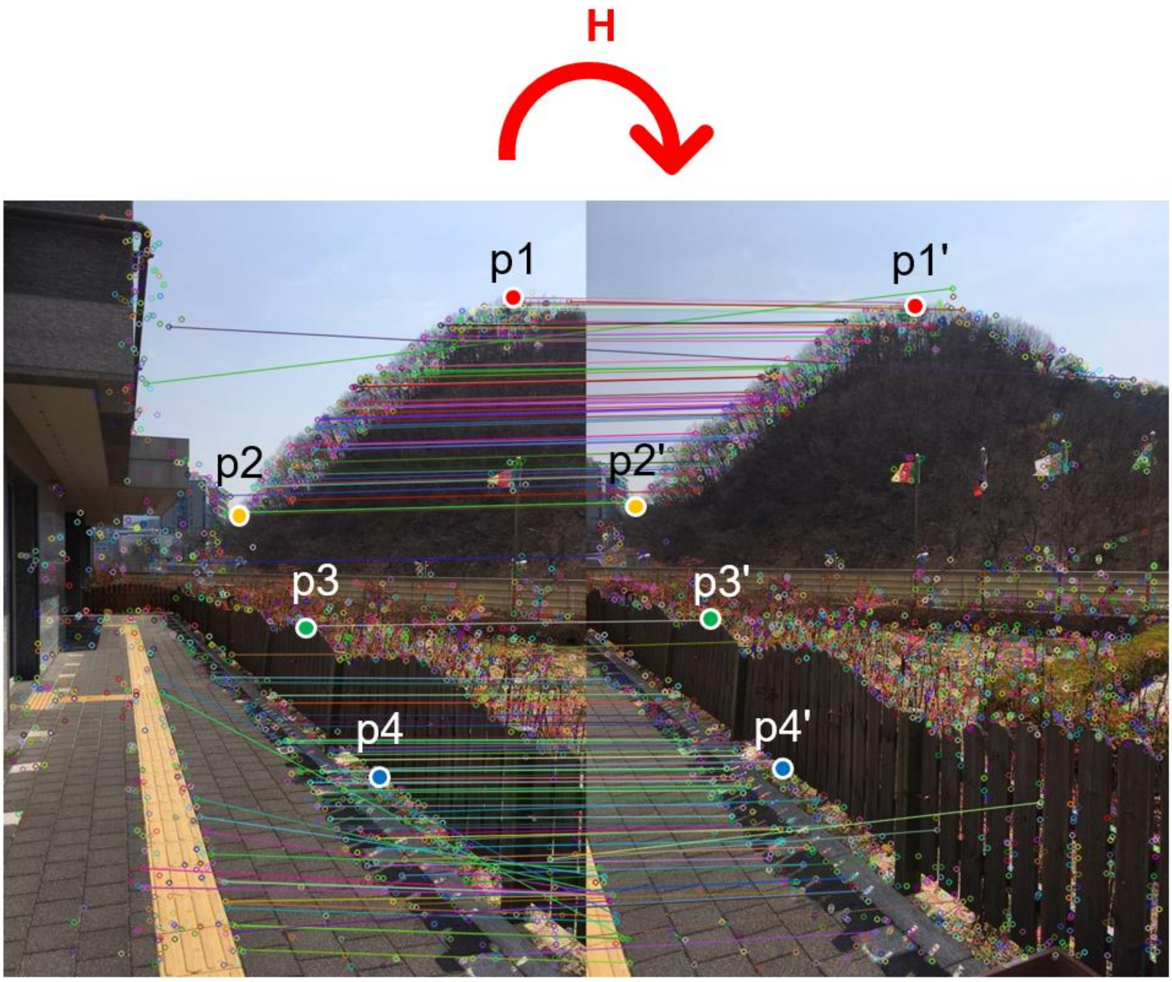


Figure 2.2 Homography estimation (step 2)

The final step is image warping through the calculated homography. Image warping refers to the process of repositioning pixels in the original image [4]. Therefore, when the target image (or object) is warped with the homography matrix, the two images (reference and target image) can be stitched together based on the corresponding features. In addition, for better results, a process called blending which is used to produce seamless panoramas is included in the algorithm. Multi-band blending has been widely introduced for image stitching [2, 19].

(a) Image warping (b) Display overlapping (c) Result

Figure 2.3 Image warping (step 3)

2.2. Stitching parameters

We have focused on a basic algorithm of feature-based image stitching to define universal stitching parameters. We analyzed good matches and homography, which are key intermediates of feature-based image stitching, to understand their physical meaning. Stitching parameters are constructed through two analysis methods, which are parameterized physical properties that can directly affect stitching result. First, two stitching parameters, number of good matches and filtering level are defined which are related with good matches through corresponding features analysis. Second, homograhy determinant, X and Y-axis scaling factor, and perspective distortion are defined as the stitching parameters that can be calculated from the homography matrix through homography analysis.

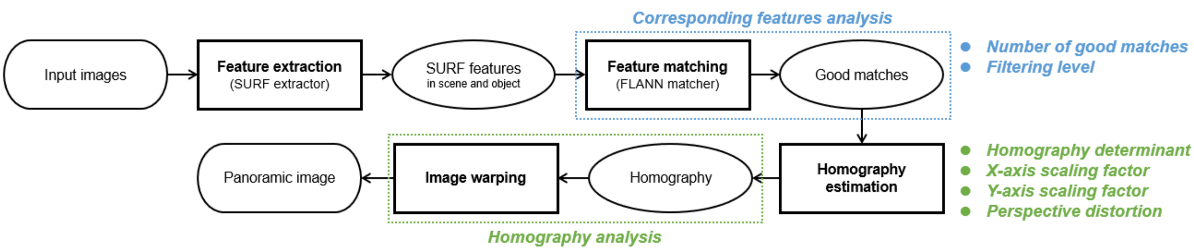


Figure 2.4 Stitching parameters and their analysis in feature-based image stitching algorithm

1) Number of good matches

Good matches (or corresponding features) are the most basic elements in feature-based image stitching [2, 19]. They are pairs of features which are extracted from two images, scene and object (or reference image and be stitched image). The features extracted from the two images can be corresponded to each other through various matching methods [26-28]. Image stitching uses flannMatcher [24] which can efficiently match high-dimensional vectorized features.

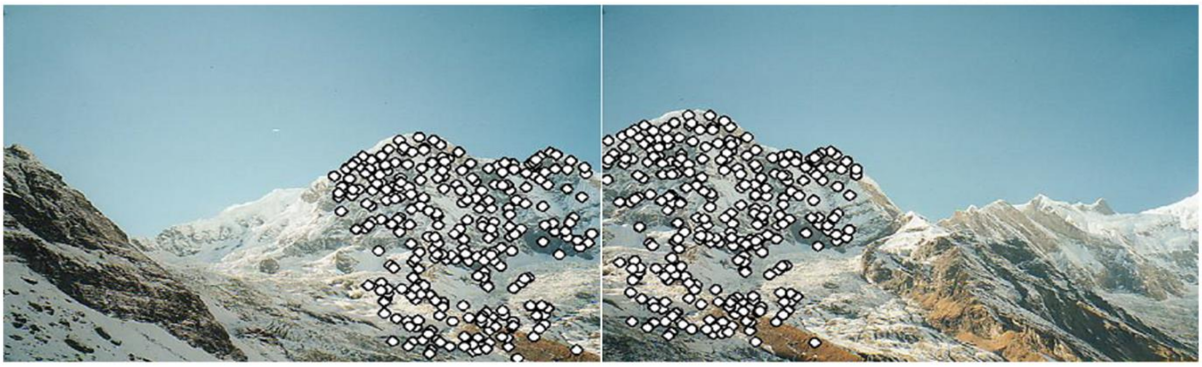
Computation

First, the features are extracted from the two images using the SURF (Speeded Up Robust Features) extractor. We used SURF because it can robustly extract features which are invariant to image scale, rotation, and illumination. The next step is matching, and we used the flannMatcher. The FLANN matcher maps all feature in the scene to the features of the object that have the minimum distance. Then, based on min_dist, which is the value of the most relevant matches, the final good matches are determined as the matches having the min_dist×3 relationship.

Figure 2.5 Examples of good matches

2) Filtering level

The filtering level is a parameter that intuitively modified the probabilistic model proposed by [1]. It reflects the geometrical relationship of good matches. Most feature-based image stitching only determines the number of good matches to obtain a warping matrix, that is, suitable homography for stitching. In [29], if the number of good matches is more than 50% of the total image, stitching is considered possible. However, there are many cases where stitching failed even though the conditions were satisfied in our experimental results. In addition, if the corresponding features are highly scattered (**Figure 2.6(b)**), proper homography cannot be obtained. Therefore, a parameter was needed to quantify that the corresponding features are closely related. What we mean by the filtering level is how well filtered out the good matches from matches, and whether they are closely related to the appropriate domain. Check the characteristics of filtering level through **Table 2**.



(a) Ideal good matches

(b) Bad good matches

Figure 2.6 Examples of good matches distribution

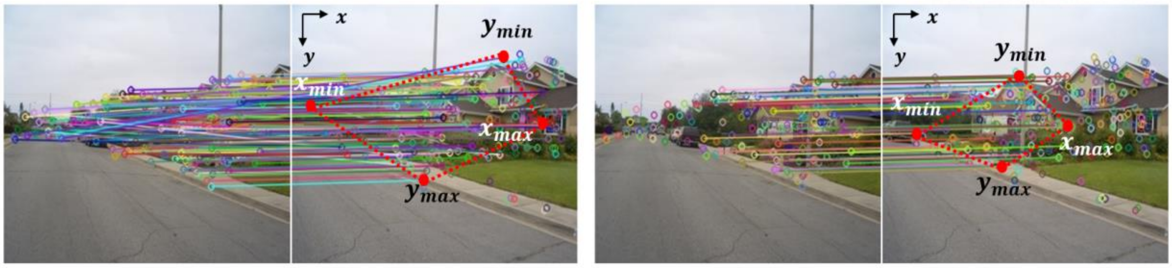
<u>Computation</u>

To calculate the filtering level, we need to find the matches area and the good matches area (**Figure 2.7**). We obtain the area ($(X_{max} - X_{min}) \times (Y_{max} - Y_{min})$) by finding the values of the outermost matches (X_{min}, Y_{min}, X_{max}, Y_{max}) among all matches. Good matches area is also obtained in the same way.

$$\text{Filtering level } (\%) = \frac{\text{Good matches area}}{\text{Matches area}} \times 100 \tag{2}$$

Table 2 Characteristics of filtering level

Filtering level	Description
Low (Excess filtering)	• Too few good matches are filtered out of matches. • The overlapping area is very closely related.
Middle (Proper filtering)	• Sufficient number of good matches are filtered out of matches. • The overlapping area is closely related.
High (Failure filtering)	• Good matches are not filtered out of matches. • The overlapping area is not related.



(a) Matches area (b) Good matches area

Figure 2.7 Components of filtering level

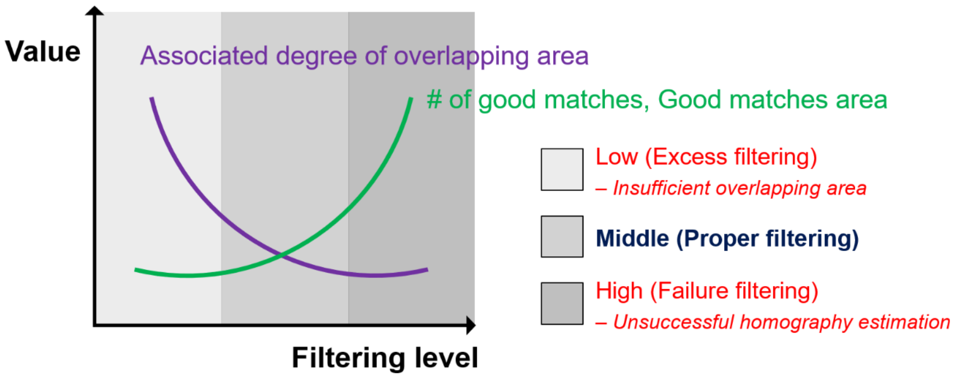


Figure 2.8 Graph of filtering level property

3) Homography determinant (D)

Homography is one of the 2D transformations and can represent various geometric relationships between two images. We estimate homography through the good matches. However, not all homographies estimated here are reliable. For example, good matches for homography estimation are wrong, or RANSAC fails to estimate proper homography.

$$\mathrm{H} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \tag{1}$$

The first 2×2 submatrix in H (Eq. (1)), a 3×3 2D transformation matrix in homogeneous form, contains information of rotation, scale, shearing, and reflection. The determinant obtained from this submatrix induces various rotation sequences or position relationships (reflections, twist, concave) between the points in the result of the 2D transformation (**Figure 2.9**). Especially, in the case of the twist, since it is a change that cannot occur in the real three-dimensional space, we can judge that the abnormal homography has been obtained. The homography determinant that induces the twist has a negative value. Since D is an element that can cause extreme errors in the warping process for stitching, we must be able to determine the appropriate D value for stitching.

$$\mathrm{D} = h_1 h_5 - h_2 h_4 \tag{3}$$

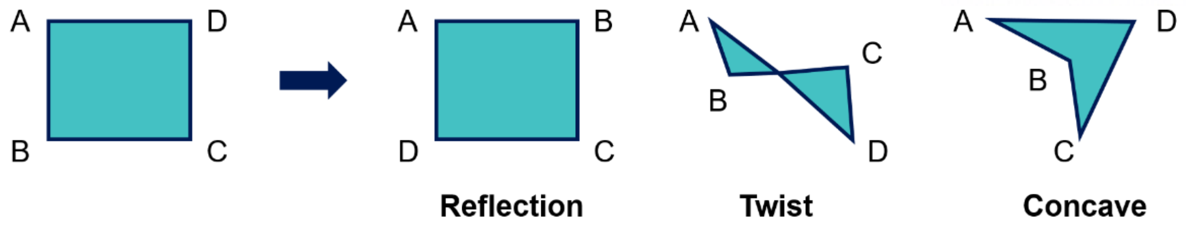


Figure 2.9 Homography determinant and morphology

4) X-axis scaling factor (S_x)

S_x is a parameter indicating the length of the X-axis unit vector by the first 2×2 submatrix in the homography matrix. If the scaling factor is too large or too small, the stitching results will be strange.

$$S_x = \sqrt{h_1^2 + h_4^2} \tag{4}$$

5) Y-axis scaling factor (S_y)

S_y is a parameter that indicates the length of the y-axis unit vector by the first 2×2 submatrix in the homography matrix.

$$S_y = \sqrt{h_2^2 + h_5^2} \tag{5}$$

6) Perspective distortion (PD)

The characteristic of "perspective" is that objects appear smaller as their distance increases the observer. So, the presence of perspective distortion means that the image plane of the rectangle gradually changes to a trapezoid. Similar to the scaling factor, if the PD value is too large, the distortion is severe.

$$\mathrm{PD} = \sqrt{h_7^2 + h_8^2} \tag{6}$$

3. Experimental Method & Results

We present the conditions of stitching parameters through four experiments.
Experiment 1: Efficient multiple image stitching method by grouping
Experiment 2: Analysis of similarity between images by histogram comparison
Experiment 3: Definition of universal stitching parameter condition
Experiment 4: Validation of stitching parameter conditions
In **Section 3.5**, we give examples of applications in term of industry and biology. The experiments use openCV, and the version is 2.4.13. The computer specification is Intel(R) Core(TM) i7-7700 CPU, 16.00GB (RAM), 64bit, Windows10.

3.1. Experiment 1: Efficient multiple image stitching method by grouping

Multiple image stitching requires an enormous amount of computation and time. In particular, the reference image should match all other images in order to align the images to be combined with each other. If we can avoid unnecessary image matching processes through grouping, efficient multiple image stitching will be possible while reducing the risk of errors. **Figure 3.1** is a panorama created by stitching 57 images. Here, the reference image (Red) is matched with a good image pair (Green) through a feature-based image stitching technique. In fact, the reference image is matched with the remaining 56 images to determine the best good image pair. This takes a lot of computation and time, and sometimes it can increase the probability of matching errors. If we can classify the wrong image pairs (Orange) that are not related to the reference image (Red) in advance, it will be a great advantage for multiple image stitching.

Figure 3.1 Multiple image stitching

Dataset

We prepared 15 image sequences obtained by rotating the camera horizontally. At this time, the 15 images can be classified according to the shooting time, so grouping becomes easy. The size of all images in the dataset is 3000×2000 pixel.

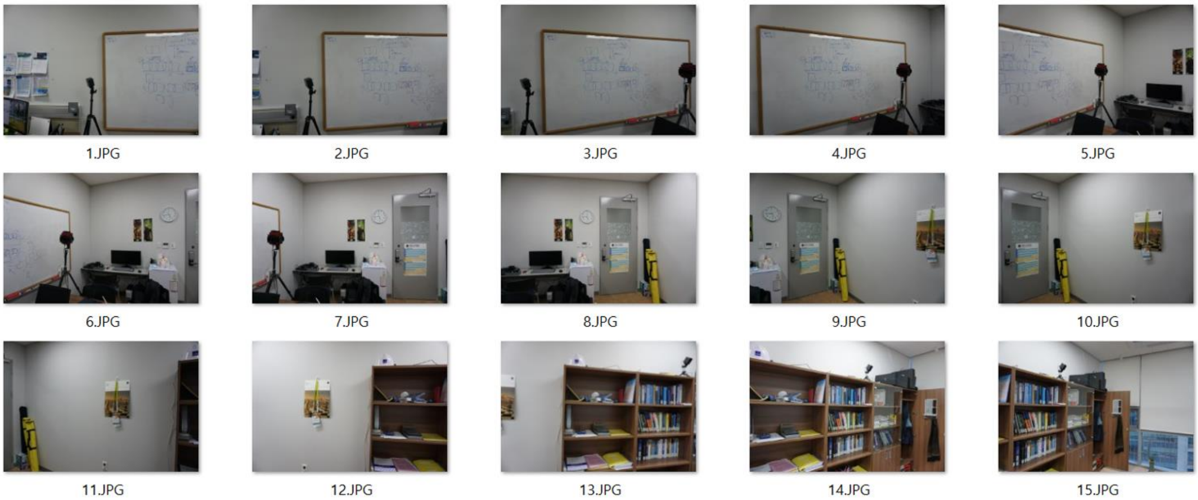


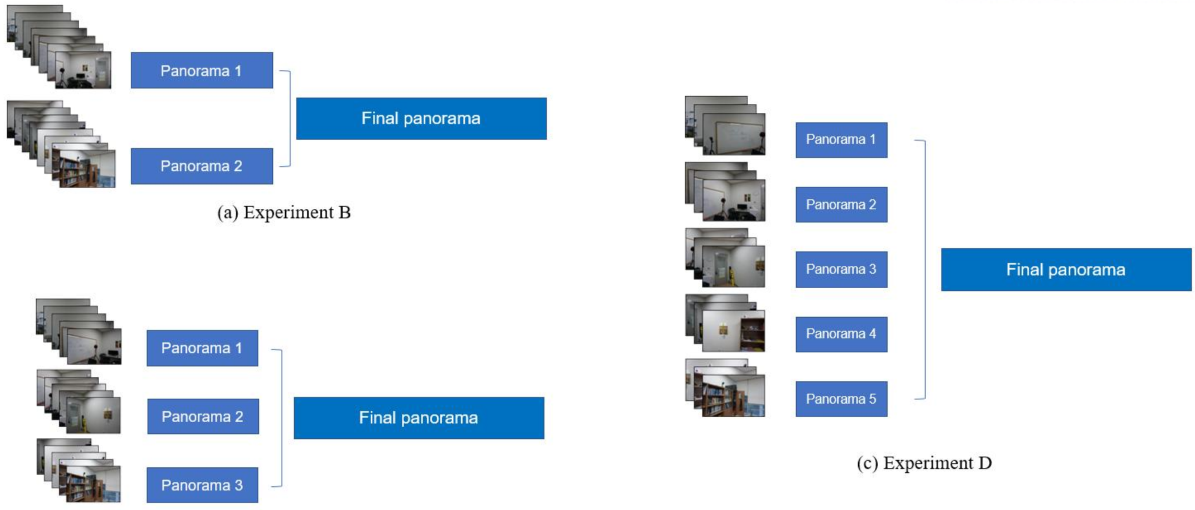
Figure 3.2 Image sequence as dataset of Experiment 1

Experimental method

We examine how much time is taken by stitching multiple images under different conditions, and consider the time-efficient stitching method. The four different conditions are as follows.

A. Stitching 15 images at a time
B. Dividing 15 images into two groups (7, 8 images) and stitching them respectively / Final stitching
C. Dividing 15 images into three groups (5 images 3) and stitching them respectively / Final stitching
D. Dividing 15 images into five groups (3 images 5) and stitching them respectively / Final stitching

In the case of A, the stitching process is performed once to make the entire panorama. In order to stitch 15 images at once, all images must be loaded simultaneously in a vector of Mat format, and they should be matched to each other to form correct image pairs and determine the overall stitching order. On the other hand, in the case of B to D, 15 images are stitched in several groups instead of stitching at once. B divides 15 images into two groups and stitches them individually. The intermediate panoramas, which are the results of stitching each of the two groups, require another stitching to create the final panoramic image. Likewise, C stitches each of the three grouped images and performs final stitching with three intermediate panoramas. D divides the images into five groups and stitching them respectively, then performs final stitching with five intermediate panoramas. In short, B, C, and D are stitched in two steps, the total number of stitching operations is 3, 4, and 6.



(a) Experiment B

(b) Experiment C

(c) Experiment D

Figure 3.3 Schematic diagram of the Experiment 1 (B, C, D conditions)

We also stitched 15 images with Hugin, a commercial image stitching software [30]. Hugin creates a panoramic image through two steps. First, as an "align" step, the features of all given images are extracted and matched to determine the correct relationship between the images. Then, perform the "composite" step, which is to remapping the aligned images in one plane and blending them to make them seamless. Finally, Hugin crops properly to get a plain result.

<u>Result</u>

Figure 3.4 shows the results of each experiment. (a) to (d) have an average size of 11430×1910 pixel, and the error range is 10 pixels in width and height. (e) is the result of Hugin. Since Hugin crops itself when creating the final panoramic image, the average size of (e) is 7505×1257 pixel and the error range is 10 pixels in width. The results of each experiment are not visually different.

Table 3.1 Processing time of stitching in each experimental condition (A~D, Hugin)

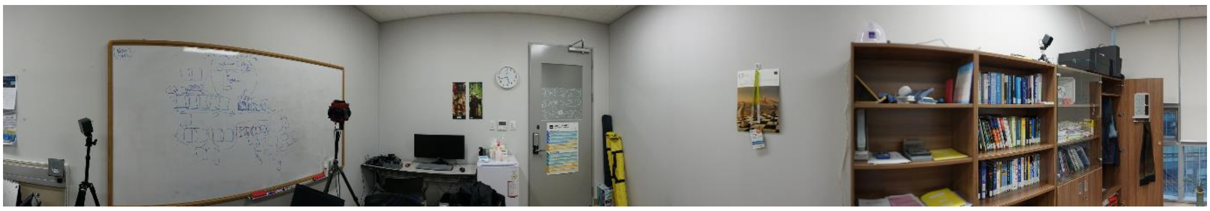
	A	B (1st / Final stage)	C (1st / Final stage)	D (1st / Final stage)	Hugin (Align / Composite)
#1	45.66	42.81 / 11.90	41.59 / 14.29	40.64 / 21.72	63.91 / 55.02
#2	46.52	43.62 / 12.43	40.53 / 14.21	39.43 / 20.93	65.22 / 54.67
#3	48.23	43.49 / 12.10	39.09 / 14.33	38.33 / 20.63	62.87 / 54.32
#4	45.89	44.09 / 11.89	40.02 / 14.56	39.29 / 21.20	68.04 / 56.34
#5	46.31	43.81 / 11.87	39.73 / 14.63	38.65 / 21.33	65.76 / 55.23
Average	46.52	43.56 / 11.88	40.19 / 14.40	39.27 / 21.16	65.16 / 55.12
Total	46.52	55.44	54.59	60.43	120.28

(a) Experiment A result

(b) Experiment B result

(c) Experiment C result

(d) Experiment D result



(e) Hugin result

Figure 3.4 Results of stitching in each experimental condition (A~D, Hugin)

Table 3.1 shows the stitching processing time measured in five repeated experiments under each condition. For the B to D experimental data, it specifies the time taken to create the intermediate panorama (1st stage) and the time taken to generate the final panorama image (Final stage).

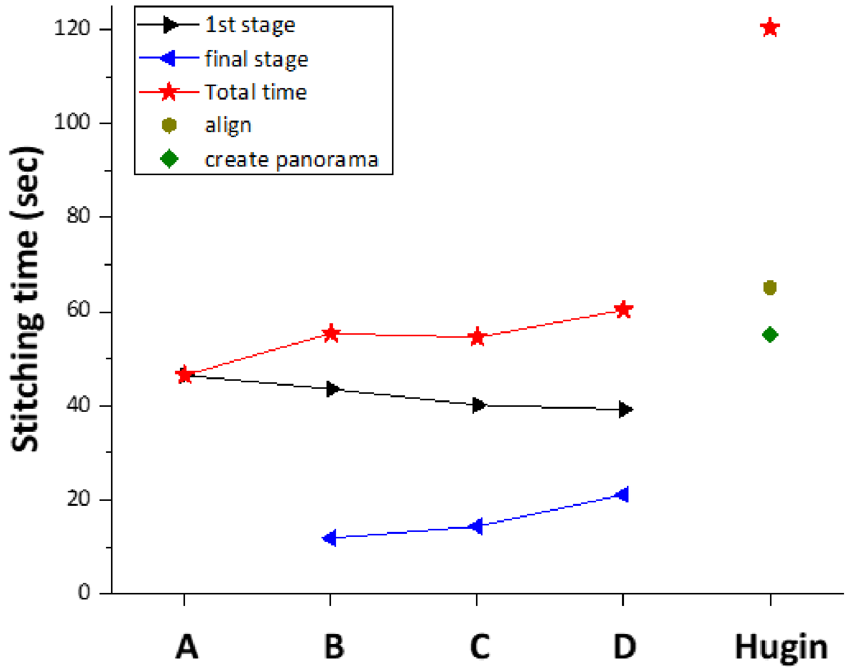


Figure 3.5 Stitching time graph under different conditions

Figure 3.5 is drawn based on the data in **Table 3.1**, and the points shown on the graph are the stitching time under each condition. In the legend to the left of the graph, the 1st stage means the processing time to stitch the grouped images to create the intermediate panoramas. The final stage is the time taken to stitch the intermediate panoramas to make the final panoramic image. In the case of A, since all 15 stitches are stitched at once, there is only value for 1st stage. On the other hand, Hugin has two types of time data that each takes to perform “align” and “composite”. The total stitching time is marked as star.

Before analyzing the experimental data in **Figure 3.5**, we summarized the expected impact of grouping on stitching. Firstly, stitching without grouping has an advantage in that the stitching process is performed only once. However, it is expected that huge amounts of memory and time will be consumed to match many images at once. Secondly, grouping when stitching is expected to reduce the number of matching cases because it does not stitch large amounts of images at once. Therefore, memory and time can be consumed efficiently when stitching. However, if you do grouping, stitching should be done as many as the number of groups to get the final panoramic image.

In **Figure 3.5**, the stitching time of 1st stage shows different trend from the final stage and total time. To analyze the stitching time in the 1st stage in detail, we prepared the following graph (**Figure 3.6**).

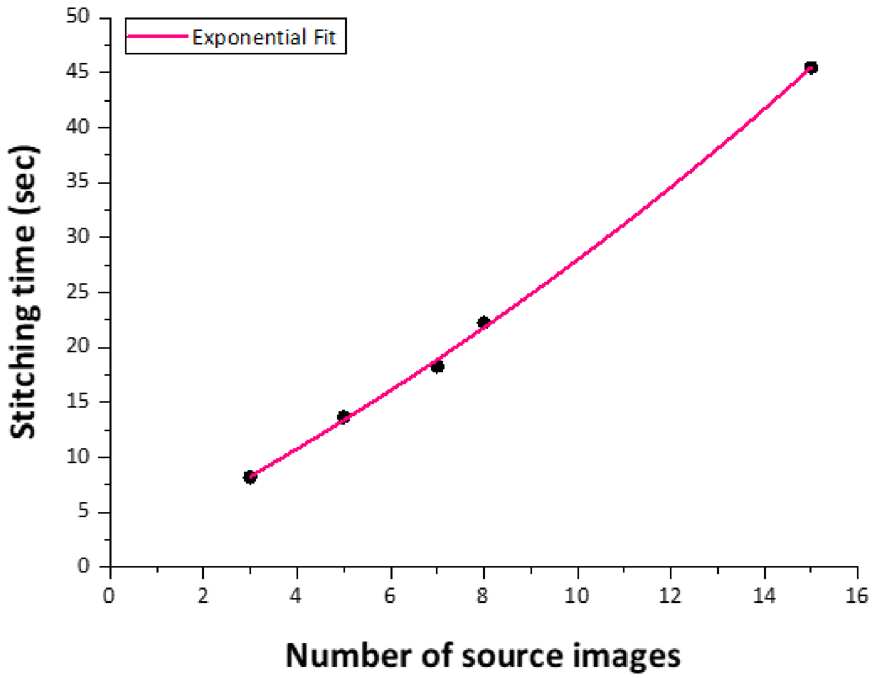


Figure 3.6 Relationship between number of source images and stitching time

The graph in **Figure 3.6** is the relationship between the number of source images and the stitching time, it shows that as the number of source images increases, the stitching time increases exponentially. This result is due to the exponential increase in the number of matching cases that take the longest time to process stitching. In other words, extracting and matching image features in feature-based image stitching has a significant effect on the speed as well as the quality of the stitching results. This can also be demonstrated in Hugin's results, where the "alignment" process takes more time than the "compositing" process. Therefore, stitching many images at once can be disadvantageous in terms of memory and speed.

Nevertheless, in **Figure 3.5**, the total time taken to produce a final panoramic image, A is the shortest. We thought that the reason for this result is because we experimented with only 15 images. Also, our experiments were done with computer with sufficient memory and performance, so we had no difficulty stitching 15 images at once. However, if stitching more than 100 images, stitching at once without grouping is expected to limit in memory and speed. It may also be difficult to perform high speed stitching on devices with low memory and poor performance. Therefore, when a large amount of stitching is carried out, grouping will help in efficient memory allocation and processing time.

In fact, many groupings for stitching is not always good. This is because the more grouping is done, the more stitching is required. This is a trade-off between the amount and the number of stitching at once (**Table 3.2**). However, it is a good idea to group source images properly in some situations. For example, when there are huge amounts of source images to produce a very large panoramic image, grouping makes it possible to avoid unnecessary matching process between images which are not relevant at all.

Table 3.2 Characteristics and effects of grouping in stitching

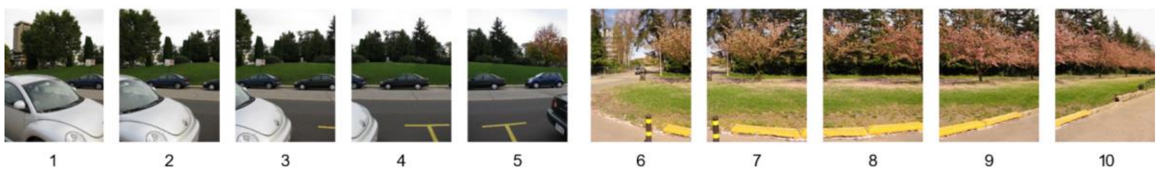
	Size of intermediate panoramas	**Total number of stitching**
Divide into few groups (grouping less)	Big →Take more time	Small →Take less time
Divide into many groups (grouping more)	Small →Take less time	Big →Take more time

3.2. Experiment 2: Analysis of similarity between images by histogram comparison

In image processing, a color histogram [31] is a representation of the color distribution in an image. A color histogram focuses only on the proportion of the number of different colors, regardless of the spatial location of the colors. So, they can indicate the essential tone of an image and the statistical distribution of colors. These histograms are often used to find similar images. The method consists of creating a histogram for each image, and then get a numerical parameter which express how well two histograms match with each other. To compare the histograms, first we have to choose a metric to express how well both histograms match. There are 4 different metrics to compute the matching, Correlation, Chi-Square, Intersection, Bhattacharyya distance [32].

Dataset

We prepared two sets of image sequences. The size of images in image set 1 is 768×1024 pixel (**Figure 3.7(a)**). The size of images in image set 2 is 480×640 pixel (**Figure 3.7(b)**). In the two image sets, the reference images are 1 and 6, respectively.



(a) Image set 1 (b) Image set 2

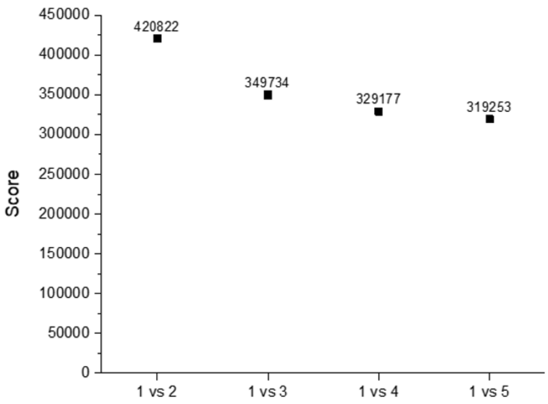
Figure 3.7 Two types of image sequences as dataset of Experiment 2

Experimental method

We used the "Intersection method" in openCV's compareHist to scoring the similarity. For the intersection method, the higher the metric (score), the more accurate the match.

$$d(H_1, H_2) = \sum_I \min(H_1(I),\ H_2(I)) \tag{7}$$

Result



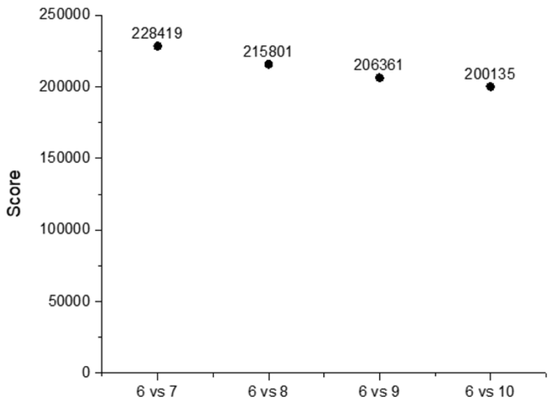


Figure 3.8 Results of histogram comparison

The main problem with histogram comparison for image classifications is that it ignores the appearance and texture of the object and depends only on the color of the object. Therefore, the histogram can be the same in other images sharing color information. In other words, there is no way to distinguish red apple from a red ball [33]. In fact, even in the result of Experiment 2, the histogram comparison cannot clearly find the most similar image. We should not be sure that the score means a value for the overlapping area. Therefore, we must find the correct image pairs using the local features of the image, not the histogram.

3.3. Experiment 3: Definition of universal stitching parameter condition

We should calculate the stitching accuracy to establish the conditions of the universal stitching parameters. Accuracy is usually calculated by comparing the experimental value with the correct value (or ground truth). In the case of image stitching, it is difficult to have ground truth. Because the main function of the stitching is to create an extended view of the image that cannot be acquired with a single camera, so the ground truth cannot be captured. Due to the absence of ground truth, the accuracy of image stitching is ambiguous to calculate. Nevertheless, previous stitching studies have been performed in their own way to evaluate improved stitching performance [34-37]. They presented their evaluation of the new algorithm with the actual stitching results or with the values of their own evaluation parameters. However, it is impossible to apply them to other stitching studies, so we cannot use them in a way to calculate stitching accuracy in our experiments.

We have devised a new experimental approach to calculate the accuracy of stitching and to establish reliable conditions for universal stitching parameters. To ensure the ground truth, we built a special dataset in a way that crops a single high-quality image. Then, we perform the feature-based image stitching with corresponding feature analysis and homography analysis to the cropped image. For the evaluation of each stitching result, an absolute difference method (opencv-absdiff) is applied to calculate the stitching accuracy.

Dataset

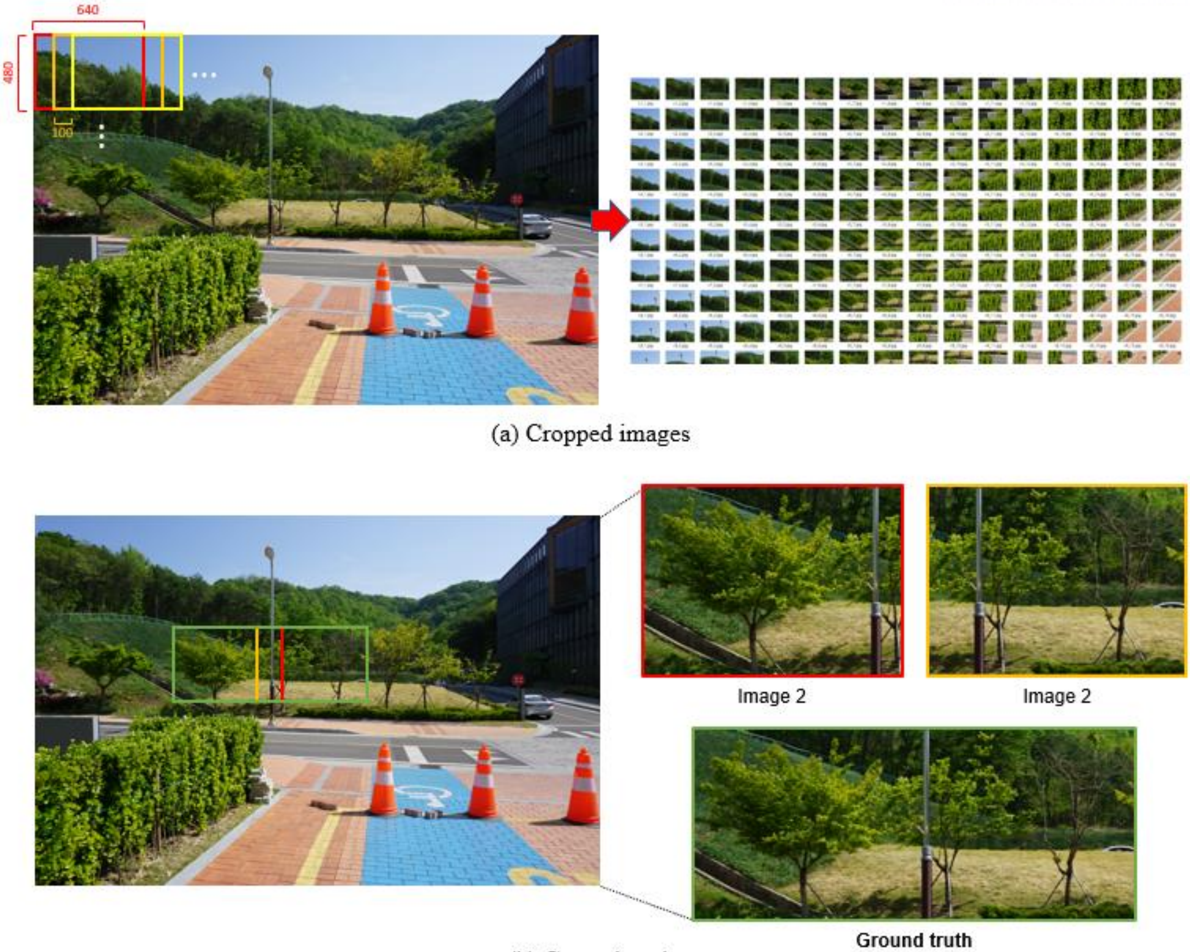
We prepared a high-quality image, including trees, mountains, buildings, objects, sidewalk blocks, etc., to create datasets with various information. Then, cut it into a 640×480 sized images with a 100-pixel step. Each image can be stitched at various overlapping levels. The image pairs to be stitched are divided into case 1 (horizontal) and case 2 (vertical), and the overlapping degree is different. For example, in the case of overlapping level 1 in case 1, it means red-orange image pairs shown in the **Figure 3.9(a)**, and red-yellow image pairs means overlapping level 2 in case 1.

Experimental method

The goal of our experiment is to calculate the stitching accuracy using the ground truth and to derive the values of the universal stitching parameters through the two analysis methods (Corresponding features analysis, Homography analysis). Stitching accuracy can be obtained using the difference between stitching result and ground truth (Eq. (8)).

$$\text{Accuracy}(\%) = (1 - \frac{difference}{Image\ size\ of\ Ground\ truth}) \times 100 \tag{8}$$

We used a normalization technique to calculate geometric errors (or differences) in stitching results. We normalized the reference image so that the object (or target image) can be stitched in any direction with respect to the reference image (**Figure 3.10**). Normalization is performed by centering the reference image and allocating as much as the object size for the up, down, left, and right directions. This normalization applies equally to the ground truth, which makes 1: 1 comparison possible with stitching results. Here, we compute the absolute difference between the normalized result and the normalized ground truth. By visualizing the absolute difference as an image (**Figure 3.11(c)**), we can count the number of pixels whose geometric positions are different from the ground truth.



(a) Cropped images

(b) Ground truth

Figure 3.9 Ground truth and cropped images as dataset of Experiment 3

Figure 3.10 Conceptual diagram for normalization of reference image

(a) Our result (b) Ground truth (c) Difference (absdiff)

Figure 3.11 Visualization of absolute difference between stitching result and ground truth

Result

We present the conditions of the stitching parameters based on the accuracy with 2258 stitching results. The stitching parameters are number of good matches and filtering level through the corresponding features analysis, and H, S_x, S_y, PD through the homography analysis. **Figure 3.12** shows the result of number of good matches and filtering level. Divide into three classes based on the point where a large change occurs between the data. Each of the three classes represents the stitching result as good, bad, and failure, and indicated in green, orange, and red. We present the result of corresponding features analysis based on accuracy (**Figure 3.12** and **Table 3.3**). In fact, we also wanted to provide the conditions of good homography that directly affect the stitching results. Because the dataset we used was a cropped image pair based on constant motion, we could not have a variety of homography values. Nevertheless, we give the abnormal homography condition in **Table 3.3** based on the stitching failure cases.

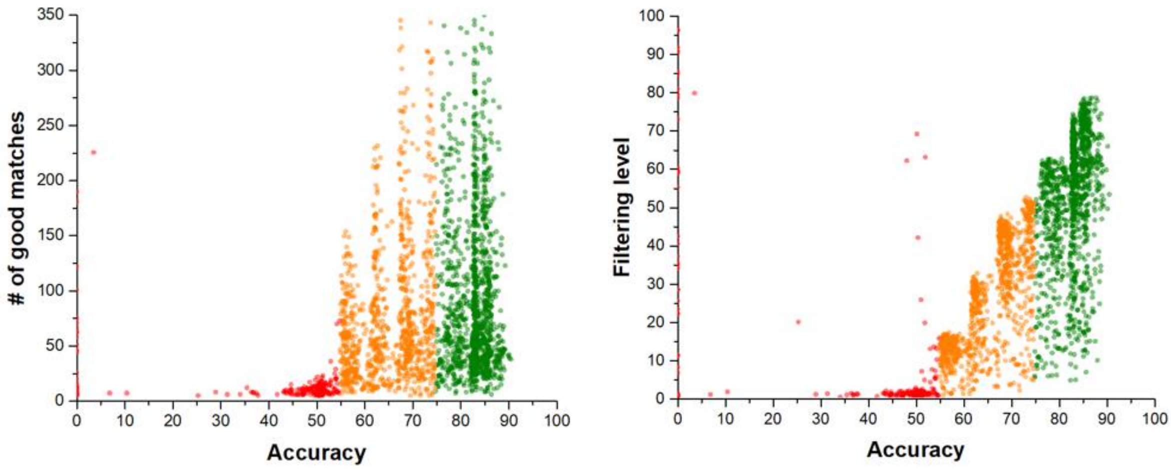


Figure 3.12 Number of good matches and filtering level based on stitching accuracy

Table 3.3 Conditions of number of good matches and filtering level

Accuracy	# of good matches	Filtering level
75-100% (Good)	$50 \leq N < 200$	$50 \leq L < 78$
55-75% (Bad)	$30 \leq N < 100$	$25 \leq N < 47$
0-55% (Fail)	$N < 40$	$L < 16\ or\ L \geq 78$

Table 3.4 Tendency of homography components

Accuracy	D	S_x	S_y	PD
75-100% (Good)	$D{\sim}1$	$S_x{\sim}1$	$S_y{\sim}1$	PD~E-07
55-75% (Bad)	$D{\sim}1$	$S_x{\sim}1$	$S_y{\sim}1$	E-05≤ PD <E-07
0-55% (Fail)	$D < 0\ or\ D > 50$	$S_x > 3$	$S_y > 3$	PD >E-03

3.4. Experiment 4: Validation of stitching parameter conditions

Experiment 4 was conducted to verify the conditions of the universal stitching parameters. We perform stitching in similar conditions to the actual stitching use and compute the values of the proposed stitching parameters. However, in such situation, the stitching accuracy cannot be calculated because the ground truth cannot be obtained. However, we classify the quality of the results into three classes and verify how the conditions of the proposed stitching parameters are valid for the real situation. Especially, we tried to verify the filtering level that presents the most obvious conditional statements according to the stitching result class.

Dataset

We prepared 155 image pairs which are acquired randomly. The datasets are obtained from various camera motion such as pan, tilt, translation (up, down, left, right), and zoom.

Figure 3.13 Various images as dataset of Experiment 4

<u>Experimental method</u>

Experiment 4 shows the results of the verification of the filtering level in Experiment 3. We stitch the image pairs obtained from various perspectives and analyze the filtering level according to each stitching result. The evaluation of the stitching results was determined by subjective judgment because the existing accuracy calculation cannot be applied. Judgment of the stitching result can be divided into three classes (good, bad and fail) as shown in **Figure 3.14**. After that, the filtering level of each image pair is analyzed based on the criteria shown in **Table 3.3**, whether it is true or false.

(a) Good (b) Bad (c) Fail

Figure 3.14 Examples of three classes (a-good, b-bad, c-fail) of stitching result

<u>Result</u>

Table 3.5 Classification result of filtering level condition in good class

Good		Condition (Filtering level)	
		Positive ($50 \le L < 78$)	Negative
Experimental Result	Positive	TP = 29	FP = 37
	Negative	FN = 21	TN = 68

Table 3.6 Classification result of filtering level condition in bad class

Bad		Condition (Filtering level)	
		Positive ($25 \le N < 47$)	Negative
Experimental Result	Positive	TP = 13	FP = 23
	Negative	FN = 31	TN = 88

Table 3.7 Classification result of filtering level condition in fail class

Fail		Condition (Filtering level)	
		Positive ($L < 16 \ or \ L \ge 78$)	Negative
Experimental Result	Positive	TP = 14	FP = 39
	Negative	FN = 24	TN = 78

Table 3.8 Verification of filtering level condition

Stitching result class	**Filtering level**	**TNR (Specificity)**	**Accuracy**
Good	$50 \le L < 78$	0.6667	0.5935
Bad	$25 \le N < 47$	0.7928	0.6516
Fail	$L < 16 \ or \ L \ge 78$	0.6476	0.6258

TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative, $True\ Negative\ rate = \frac{TN}{TN+FP}$, $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

3.5. Applications

1) Panoramic image production based on high-quality bloodstain image search

The bloodstains at the crime scene play an important role in reconstructing the incident and determining the direction of the investigation. The detailed features of bloodstain such as shape, distribution, form, and the number are mainly obtained through close-up photography. Most scientific investigations use a high-resolution camera to obtain clear bloodstain information and use a landmark that consists of a specific pattern or scale to indicate the size and location of the bloodstain. In order to reconstruct the crime scene, it is important to understand the positional relationship and distribution of the bloodstains in the three-dimensional space. However, since high-resolution bloodstain images are close-up photographs that do not have a wide field-of-view, the spatial position of the bloodstain is unknown. To solve the problem, we produced a panorama to display the overall space of crime scene and link the corresponding high-resolution bloodstains pictures on the panorama to recognize the spatial distribution information of the bloodstains.

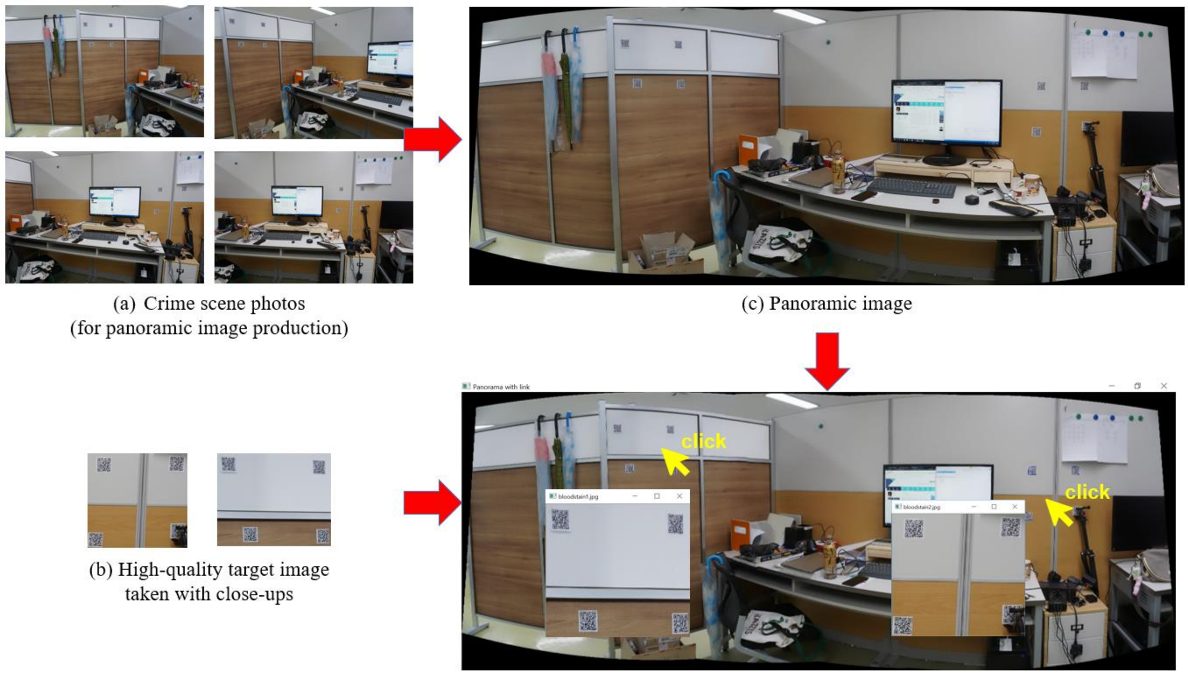
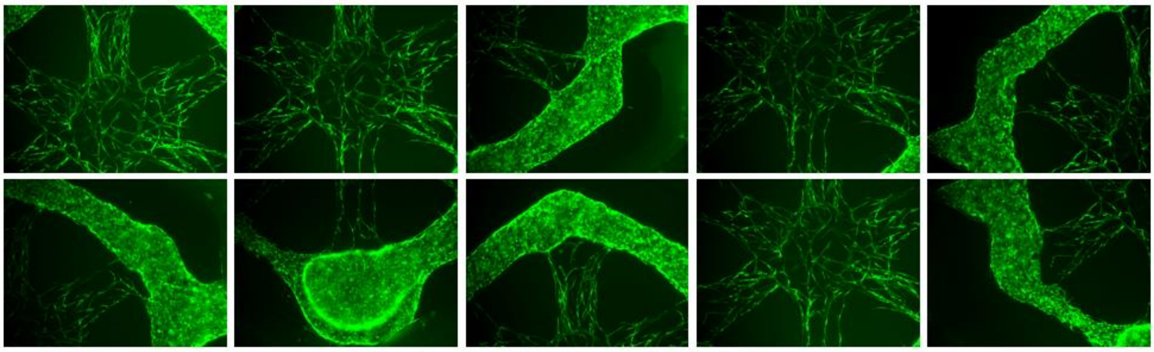


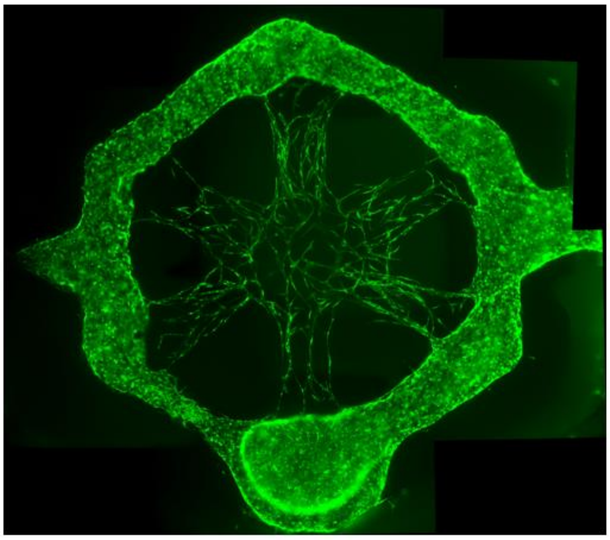
Figure 3.15 Panoramic image production based on high-quality target image search

2) Feature-based image stitching in a microscope

Microscope is an instrument used to observe objects that are too small to be seen by the naked eye. The field-of-view of the object image is determined from the objective lens. The higher the magnification of the microscope, the smaller the field-of-view of image. In order to have a wide field-of-view, a low-magnification objective lens must be selected, so that high-resolution imaging is impossible. With image stitching technology, wide field-of-view can be obtained even when the resolution is constant at high-magnification objective lens. This is significant to overcome the physical limitations set by the objective lens. Recently, image stitching has been used well in the microscopy [38]. We performed image stitching with various kinds of source images obtained by microscope and present the results. In particular, we confirmed that the feature-based image stitching works well with images of biological cell and tissue rather than general objects and landscape images. The source images in **Figure 3.16** and **Figure 3.17** are live captures with the stage moving freely. They have fine lines and complex textures as bio-images, and they support only a single color due to the dyeing characteristic for visualization.

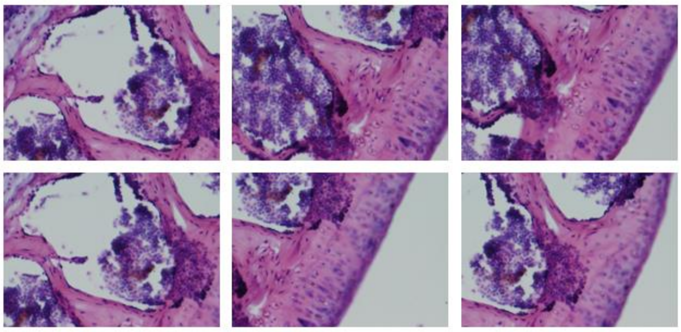


(a) Microscope photographs of f-lobule tissue

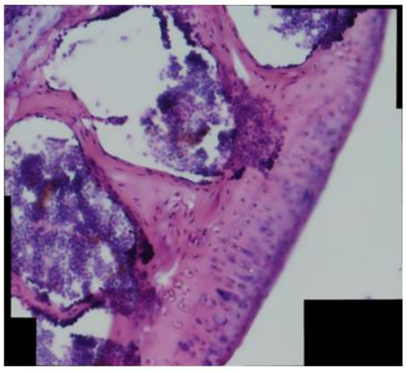


(b) Result

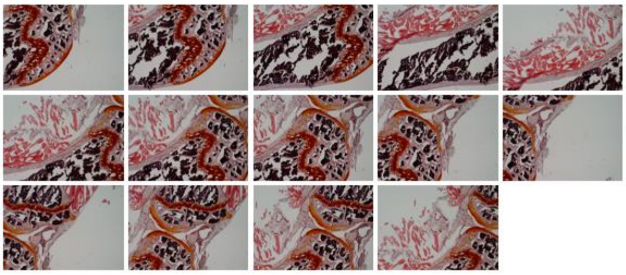
Figure 3.16 Microscope photographs of f-lobule tissue and their stitching result



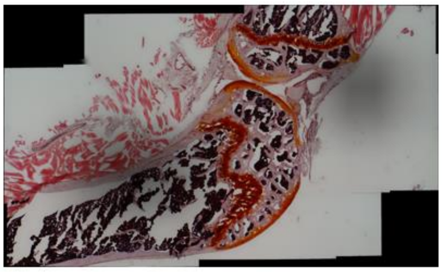
(a) Microscope photographs of cartilage cells of rats (case 1)



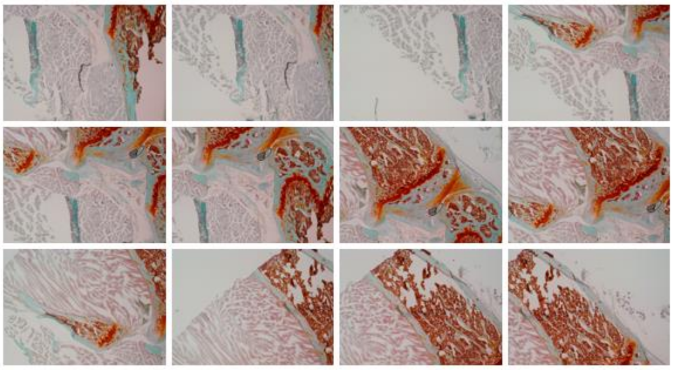
(b) Result (case 1)



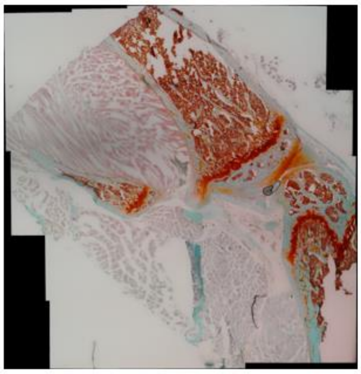
(c) Microscope photographs of cartilage cells of rats (case 2)



(d) Result (case 2)



(e) Microscope photographs of cartilage cells of rats (case 3)



(f) Result (case 3)

Figure 3.17 Microscope photographs of cartilage cells of rats and their stitching result

4. Discussion

This paper describes the effects of grouping (Experiment 1) in stitching, histogram comparison (Experiment 2) for finding similar images for grouping, and stitching conditions based on universal stitching parameters (Experiment 3). First, grouping has obvious advantages in multiple image stitching. It can reduce the amount of computation to be processed at once through grouping, so feature-based stitching can work well on devices with insufficient memory. In addition, unnecessary matching processes can be reduced in advance, and a stitching failure due to a matching error can be avoided through grouping.

We identified the benefits of grouping and were deeply concerned about the methodology of grouping. If we have a well-ordered image sequence, grouping is easy. However, given randomly mixed images, we should determine pairs of similar or related images. We have attempted to group images using histogram comparisons that are widely used when analyzing image similarity. As a result, the most similar image of the reference image can be presented through a specific score, but we have identified the limit that the histogram considers only the color distribution of the entire image.

For good stitching results, the geometric distribution of corresponding features in the given image pair must be shared without severe distortion, so we need to parameterize it. We defined the essential conditions for a good image pair through six parameters that are computed through corresponding features analysis and homography analysis. In Experiment 3, a special dataset was constructed to define conditions of the stitching parameters based on the stitching accuracy, thus providing clear stitching conditions.

The application of feature-based image stitching depends on the characteristics of the image. We covered the differences between the general industrial sector and the bio-imaging sector through the application examples in Section 3.5. In bioimaging, feature-based image stitching technology is still poorly developed. Because biomaterials are texturally duplicated with similar features, the feature extraction and matching errors occur frequently. Moreover, bioimaging reveals the structure of the sample through dyeing, so if the dyeing supports only one color, the diversity of color channels is also less than that of general industrial images. Therefore, in order to apply feature-based image stitching well in bio-imaging, other parameters such as camera acquisition path should be actively utilized.

5. Conclusion

This paper deals with the analysis and application of feature-based image stitching, and we have defined the universal stitching parameters for successful stitching. The values of the universal stitching parameters can be computed by analyzing the corresponding features and homography which are essential elements in feature-based image stitching algorithm. To evaluate the stitching results, we calculated the stitching accuracy through the ground truth obtained from the cropped images. Based on the stitching accuracy, conditions of the six stitching parameters (number of good matches, filtering level, homography determinant, X and Y-axis scaling factor, and perspective distortion) were established and validated through various datasets. We can apply the proposed stitching conditions to remove the wrongly selected image pair which is the main cause of stitching failure in advance. In addition, the advantage that the stitching parameters of the proposed condition are universal, will allow other researchers to easily use our stitching conditions.

REFERENCES

[1] M. Brown and D. G. Lowe, "Automatic Panoramic Image Stitching using Invariant Features," *International journal of computer vision*, vol. 74, no. 1, p. 59-73, Dec 2006.

[2] R. Szeliski, "Image Alignment and Stitching: A Tutorial," *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 1, p. 1-104, Jan 2006.

[3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge university press, 2003.

[4] P. S. Heckbert, "Fundamentals of texture mapping and image warping," University of California, Berkeley, 1989.

[5] J. Gao, S. J. Kim, and M. S. Brown, "Constructing image panoramas using dual-homography warping," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, p. 49–56.

[6] W. Y. Lin, S. Liu, Y. Matsushita, T. T. Ng, and L. F. Cheong, "Smoothly varying affine stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, p. 345–352.

[7] C. H. Chang, Y. Sato, and Y. Y. Chuang, "Shape-preserving halfprojective warps for image stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, p. 3254–3261.

[8] C. C. Lin, S. U. Pankanti, K. N. Ramamurthy, and A. Y. Aravkin, "Adaptive as-natural-as-possible image stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, p. 1155–1163.

[9] N. Li, Y. Xu, and C. Wang, "Quasi-homography warps in image stitching," *IEEE Transactions on Multimedia*, vol. 20, no. 6, p. 1365-1375, Jun 2018.

[10] Y. S. Chen and Y. Y. Chuang, "Natural image stitching with the global similarity prior," in *European Conference on Computer Vision (ECCV)*, 2016, p. 186–201.

[11] G. Zhang, Y. He, W. Chen, J. Jia, and H. Bao, "Multi-viewpoint panorama construction with wide-baseline images," *IEEE Transactions on Image Processing*, vol. 25, no. 7, p. 3099–3111, Jul 2016.

[12] S. Li, L. Yuan, J. Sun, and L. Quan, "Dual-feature warping-based motion model estimation," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, p. 4283–4291.

[13] N. Li and T. Liao, "Single-Perspective Warps in Natural Image Stitching," arXiv:1802.04645 [cs], Feb 2018.

[14] T. Xiang, G. S. Xia, and L. Zhang, "IMAGE STITCHING WITH PERSPECTIVE-PRESERVING WARPING," arXiv:1605.05019 [cs], May 2016.

[15] F. Zhang, F. Liu, "Parallax-tolerant image stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, p. 3262-3269.

[16] K. Lin, N. Jiang, L. F. Cheong, M. Do, and J. Lu, "Seagull: seam-guided local alignment for parallax-tolerant image stitching," in *European Conference on Computer Vision (ECCV)*, 2016, p. 370-385.

[17] F. Zhang, F. Liu, "Casual stereoscopic panorama stitching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, p. 2002-2010.

[18] C. Herrmann, C. Wang, R. S. Bowen, E. Keyder, M. Krainin, C. Liu, and R. Zabih, "Robust image stitching with multiple registrations," in *European Conference on Computer Vision (ECCV)*, 2018, p. 53-67.

[19] Y. J. Ha and H. D. Kang, "Evaluation of feature based image stitching algorithm using OpenCV," in *The 10th International Conference on Human System Interactions (HSI)*, 2017, p. 224-229.

[20] T. Tuytelaars and K. Mikolajczyk, "Local Invariant Feature Detectors: A Survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, p. 177-280, Jun 2008.

[21] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International journal of computer vision*, vol. 60, no. 2, p. 91-110, Nov 2004.

[22] H. Bay, A. Andreas, T. Tuytelaars, and L. V. Gool, "Speeded-Up Robust Features (SURF)," *Computer vision and image understanding*, vol. 110, no. 3, p. 346-359, Jun 2008.

[23] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *IEEE International Conference on Computer Vision (ICCV)*, 2011, p. 2564-2571.

[24] M. Muja and D. G. Lowe, "Scalable Nearest Neighbor Algorithms for High Dimensional Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, p. 2227-2240, Nov 2014.

[25] A. Agarwal, C. V. Jawahar, and P. J. Narayanan, "A survey of planar homography estimation techniques," Centre for Visual Information Technology, Hyderabad, 2005.

[26] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *VISAPP International Conference on Computer Vision Theory and Applications*, 2009, p. 331-340.

[27] T. S. Shaikh and A. B. Patankar. "Multiple feature extraction techniques in image stitching." *International Journal of Computer Applications*, vol. 123, no. 15, p. 29-33, Aug 2015.

[28] R. Tron, X. Zhou, C. Esteves, and K. Daniilidis, "Fast Multi-Image Matching via Density-Based Clustering," in *IEEE International Conference on Computer Vision (ICCV)*, 2017, p. 4057-4066.

[29] H. Chau and R. Karol, "Robust panoramic image stitching," Stanford University, CA.

[30] Hugin, "Hugin - Panorama photo stitcher," 2017. [Online]. Available: http://hugin.sourceforge.net. [Accessed: 1- Mar- 2017].

[31] M. J. Swain and D. H. Ballard, "Indexing Via Color Histograms," in *IEEE International Conference on Computer Vision (ICCV)*, 1990, p. 390-393.

[32] Opencv dev team, "Histogram comparison," 2015. [Online]. Available:

https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_comparison/histogram_comparison.html?highlight=histogram%20comparison. [Accessed: 10- Oct- 2017].

[33] X. Y. Wang, J. F. Wu, and H. Y. Yang, "Robust image retrieval based on color histogram of local feature regions," *Multimedia Tools and Applications*, vol. 49, no. 2, p. 323-345, Aug 2010.

[34] R. Szeliski, *Computer vision: algorithms and applications*. Berlin: Springer Science & Business Media, 2010.

[35] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, "Interactive digital photomontage," *ACM Transactions on Graphics (ToG)*, vol. 23, no. 3, p. 294-302.

[36] C. Herley, "Automatic occlusion removal from minimum number of images," in *IEEE International Conference on Image Processing (ICIP)*, 2005, p. 1046-1049.

[37] W. Y. Lin, L. Liu, Y. Matsushita, K. L. Low, S. Liu, "Aligning images in the wild," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012, p. 1-8.

[38] J. Chalfoun, M. Majurski, T. Blattner, K. Bhadriraju, W. Keyrouz, P. Bajcsy, and M. Brady, "MIST: Accurate and Scalable Microscopy Image Stitching Tool with Stage Modeling and Error Minimization," *Scientific reports*, vol. 7, no.1, p. 1-10.

ACKNOWLEDGEMENT

I would like to thank my thesis committee: Prof. Hyun-Wook Kang, Prof. Hyun-Deok Kang and Prof. Semin Lee, for their insightful comments and encouragement. The words from the professors will be a great force for my life in the future. I wish professors continued success and prosperity.
I would like to thank Prof. Hyun-Wook Kang for accepting me as a student during my master's degree, and I would like to express my gratitude for your kind assistance and support.

Prof. Hyun-Deok Kang,
In the winter of 2015, it is still vivid to meet you for the first time. Thank you very much for leading me who has no knowledge about image processing and computer vision so far. The research topic you have been teaching has always been interesting and has allowed me to experience various research worlds. Thanks to your encouragement and support, I could come up to this point. I look forward to continuing my research and social life with you.
My labmates: Si Hyeong, Tae Heon, and my lab senior Hyeongeun,
I was so grateful that you guys have been helping me with my task all the time and thinking about it together. At first I was studying alone, but after I met you guys, I could grow up faster. I hope that our relationship will continue in the future and let us all succeed!
My Start-up team colleagues met in spring 2018: members of PIRECO, DogQuiet, LOAD, DID,
I was delighted to work with you in researching, commercializing and drawing the future together. And I was so impressed that you were supported my graduation degree. I will now fully participate in our future.
My best friends: Sol Mi, Ha Eun,
Thank you for being my closest supporter while preparing for graduation. Although I graduate first, I encourage you to finish your doctorate well. It was a blessing to me that we could meet in the first year of 2013 and be with us so far.
My dearest family: parents, grandmother, and older brother,
Thank you so much for watching me and for understanding me in any fuss. Your youngest, I now have a perfect diploma and go to a real society. I am always thankful that you are with me and I will try to live hard for my goals with your support.

Thank you again to those who have contributed to my degree.