

# Mosaicing of Microscope Images based on SURF

Wen Rong, Hui Chen, Jiaju Liu, Yanyan Xu  
 School of Information Science and Engineering  
 Shandong University  
 Jinan, Shandong 250100, China  
 Email: huichen@sdu.edu.cn

Ralf Haeusler  
 Department of Computer Science  
 The University of Auckland  
 Auckland, New Zealand  
 rhae001@aucklanduni.ac.nz

**Abstract**— In this paper, we describe the design of a mosaicing technique for images from a microscope system with automatically controlled object stage and image capture unit. Due to the limited field of view in microscope imagery, larger objects are split up into many adjacent, but slightly overlapping frames. In many fields, such as medicine or biology, it is vastly beneficial that these image patches are recomposed to a single (panoramic) image.

We propose a feature matching and registration method based on SURF (Speeded-Up Robust Features). This method is most accurate for microscopy images, which usually have repetitive, blob-like structures.

Further steps in our algorithm are estimation of transformation parameters for image warping and blending for elimination of color and luminance differences between images.

For feature matching, we propose a new method of dividing descriptor windows. This increases matching speed considerably.

The experimental results provided demonstrate the performance of our method.

## I. INTRODUCTION

Image mosaicing has been widely applied in many areas such as robotics, virtual reality, intelligent surveillance and so forth [1], [2], [3]. In microscopy, image mosaicing is the only way of creating large-scale images. For example, our microscope, using a 40X Olympus UIS2 objective lens, has a field of view that allows to capture only object sections of  $0.3\text{mm}$  in size with a mounted digital camera. Using weaker magnifications provides a larger field of view, but is decreasing the spatial resolution accordingly. Image mosaicing allows to create images with a wide (theoretically unlimited) field of view at full resolution. The main component of image mosaicing techniques is generally referred to as image registration. Existing image registration methods can be categorized into area-based and feature-based ones. Area-based methods are sensitive to intensity changes such as noise and varying illumination. This limits accuracy and robustness of respective methods [4]. However, feature-based methods based on SIFT (Scale-invariant Feature Transform) [5], [6] for example, have superior discriminative properties and are relatively efficient. SURF (Speeded-Up Robust Features) [9], is another scale- and rotation-invariant feature extraction method with improved performance compared to SIFT. Specifically, SURF is advantageous in detecting repetitive blob-like structures as seen in microscopy images, such as cells or cell organelles. For prospective on-line applications of our system, speed is as crucial as accuracy and robustness. Therefore, we propose an improvement of SURF in our mosaicing scheme.

The paper is organized as follows: In Section 2, we review matching with SURF and propose a new strategy for separating descriptor windows, resulting in an increased speed. Mismatches are rejected using RANSAC [10]. Section 3 is revising the estimation of projective transformation parameters from point correspondences found with SURF.

A simple method for color balancing and blending image boundaries is illustrated in Sections 4 and 5. Section 6 describes the results obtained with our algorithm on selected microscopy images. Section 7 concludes.

## II. MATCHING ALGORITHM BASED ON SURF

### A. Integral Images

A key concept used in SURF are integral images[7]. With these, values in a rectangular subimage can be added quickly and independent of the window size. From an input image  $I$  and a point  $(x, y)$ , the integral image  $I_{\Sigma}(x, y)$  is computed as the sum of intensity values between the point  $(x, y)$  and the origin  $(0, 0)$ .

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y) \quad (1)$$

The sum of pixel intensities over any upright rectangular area now can be computed with only three additions. For a rectangle  $S$  bounded by vertices  $A$ ,  $B$ ,  $C$  and  $D$ , the sum of pixel intensities  $\Sigma_S$  is:

$$\Sigma_S = I_A + I_D - (I_B + I_C) \quad (2)$$

### B. Selecting Key Points

SURF is based on the determinant of the Hessian matrix. With the Hessian matrix, locations of the blob-like structures can be detected accurately. For a point  $\mathbf{x} = (x, y)$  in an image  $I$ , the Hessian matrix  $H$  is defined as follows [9]:

$$H = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{yx}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix} \quad (3)$$

where  $L_{xx}(\mathbf{x}, \sigma)$ ,  $L_{yy}(\mathbf{x}, \sigma)$ ,  $L_{xy}(\mathbf{x}, \sigma)$  and  $L_{yx}(\mathbf{x}, \sigma)$  are the convolution of the Gaussian second order derivatives in  $x-$ ,  $y-$ ,  $xy-$  and  $yx-$  direction with the image  $I$  centered at point  $\mathbf{x}$ .  $\sigma$  denotes the scale in scale-space. Approximation of these derivatives with box filters [9] (see Figure 1) reduces computational costs significantly.

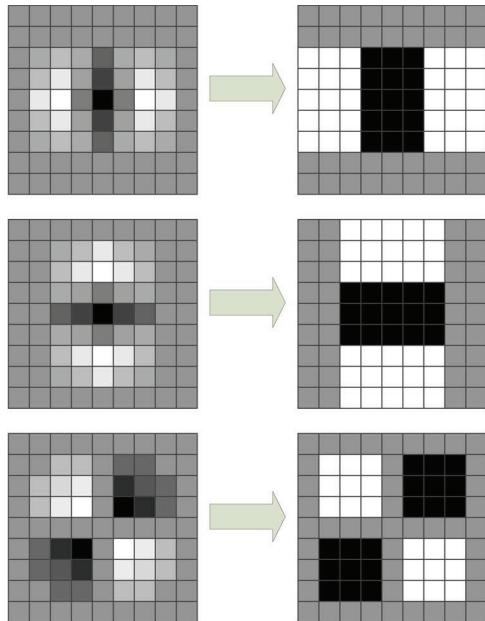


Fig. 1: Top to bottom: Approximation for the Gaussian second order partial derivative in x-, y- and xy- direction.

With the box filters  $D_{xx}$ ,  $D_{xy}$  and  $D_{yy}$  and assuming that the mixed derivatives are continuous, the determinant of the Hessian can be approximated by:

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (4)$$

Here,  $\det(H_{approx})$  is the blob response of a point. Key points  $\mathbf{x} = (x, y, \sigma)$  are defined by local maxima of this response for different scales.

To construct scale space, we can apply box filters of any size at exactly the same speed directly on the original image rather than on a scaled down layer [9]. The computing cost is constant for the filtering time in this method is independent of mask size.

Scale space is divided into several octaves which are constructed by a constant number of scale levels. To locate key points, first we remove all responses that are below a predetermined threshold. Then we perform a non-maximum suppression to determine a set of candidate points. Here, each pixel in the scale-space is compared to its 26 neighbors, comprised of the 8 points in the native scale and the 9 ones in each of the scales above and below (see Figure 2). Finally we interpolate nearby data to determine the point location in both, space and scale, to sub-pixel accuracy. Figure 3 shows detected key points on a microscopy image. This illustrates characteristics of features obtained using Hessian-based detectors.

### C. Key Point Description

To construct key points descriptors, Haar wavelets are used to compute gradients in the  $x$  and  $y$  directions. To determine the dominant orientation, a sliding orientation window of  $\pi/3$  is applied to obtain the sum of all responses within it. The

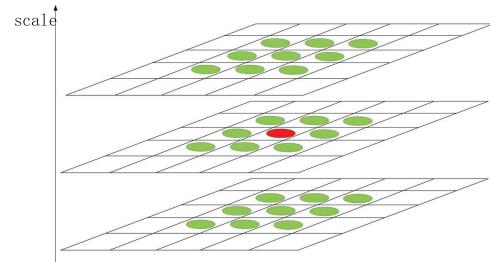


Fig. 2: The 26 neighbors in scale-space.

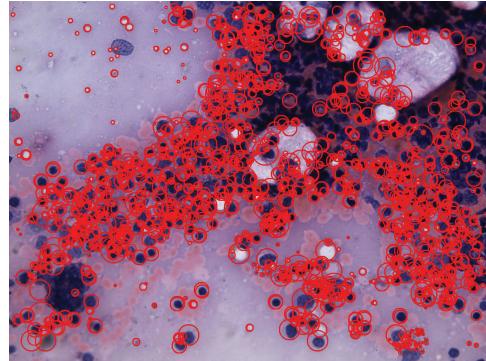


Fig. 3: Detected key points for red bone marrow.

longest vector over all windows defines the orientation of the key point (see Figure 4).

At the beginning, a  $20\sigma$  square window around the key point is constructed for extracting the SURF descriptor, where  $\sigma$  refers to the detected scale. In the algorithm of SURF, the descriptor window is divided into  $4 \times 4$  regular subregions. Each response of the subregions is weighted with a Gaussian centered at the key point. This defines a vector of length 128.

In order to reduce the computational cost, we propose a new method for dividing the descriptor window (see Figure 5). Here, Haar wavelets of size  $2\sigma$  are calculated for all points of each subregion. Each subregion contributes eight values (see Figure 6) to the descriptor vector resulting in a combined vector of length 64. The eight vectors of every point are the sums of  $d_x$  and  $|d_x|$  for both  $y < 0$  and  $y \geq 0$  respectively, and that of  $d_y$  and  $|d_y|$  for both  $x < 0$  and  $x \geq 0$  respectively. As the neighboring subregions of the key point contribute more to this vector, our region dividing method can approximate the Gaussian weighting measure. Thus computational cost is saved in generating Gaussian mask and convolution. This also increases the matching speed considerably as the length of vectors is reduced by half. However, the matching result does not degenerate significantly (see Table I and Figure 7). Table I compares our method of window division with SURF-128. Values compared are computing time, number of matching points and mismatch rates on the microscopy images of red bone marrow of size  $1600 \times 1200$ .

In Figure 7, there are some mismatching points on the upper right of the image, due to our method. These mismatches did not occur in SURF-128. However, this does not seriously

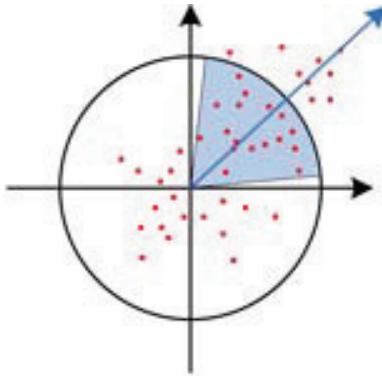


Fig. 4: Orientation assignment.

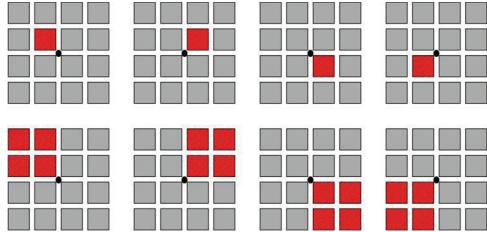


Fig. 5: Subdivision of the descriptor window into eight regions.

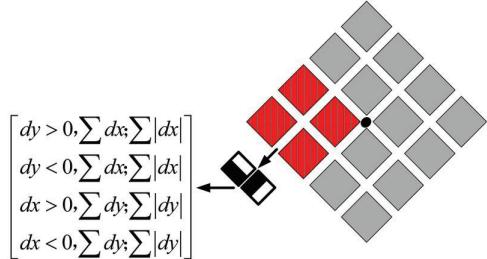


Fig. 6: Construction of the feature vector with length 64.

TABLE I: Comparison of matching results of our method with SURF-128.

Method	Computing time [ms]	Number of points(mismatches)	Mismatch rate
SURF-128	14937	300(2)	0.67%
Our method	11865	392(12)	3.06%

affect the estimation of transformation parameters since most mismatching points are removed by RANSAC in a subsequent refining step.

#### D. Matching and Refining

In the process of matching, we compute the nearest and second nearest Euclidean distance for each key point to its counterpart in the related microscopy image. If the ratio of the nearest Euclidean distance to the second nearest one is less than 60%, we choose the nearest pair as matching points. After this step, we use RANSAC to refine the matching points. We connect every pair of matching points with a line, so each line

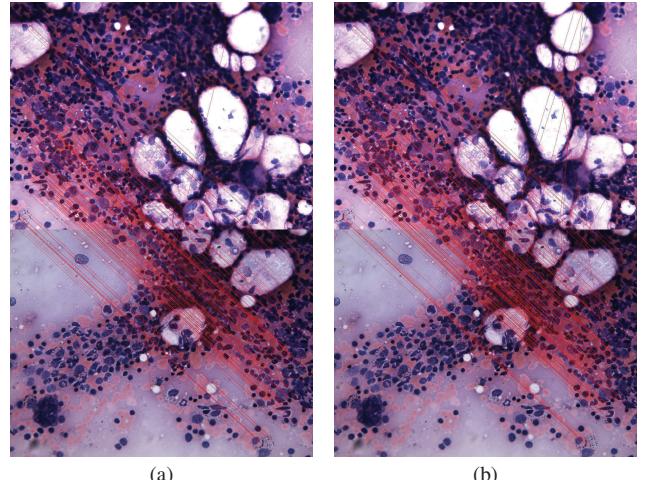


Fig. 7: Matching comparison: (a) SURF-128; (b) our algorithm.

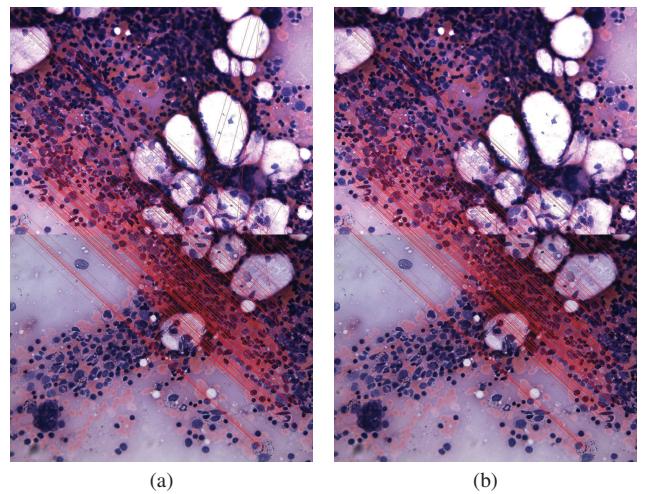


Fig. 8: Refinement comparison: (a) before refining; (b) after refining.

represents a pair of matching points. The average of all line directions is called dominant motion. Point correspondences that are in line with the dominant motion are regarded as good matches. Points matches inconsistent with the dominant motion are rejected. Figure 8 shows point correspondences after refinement.

### III. TRANSFORMATION PARAMETER ESTIMATION

Spatially adjacent microscopy images can be related to each other geometrically with a planar perspective transform  $H$  (also called homography):

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (5)$$

Here,  $(x_i, y_i)$  and  $(x'_i, y'_i)$  is a pair of matching points.  $H$

is a  $3 \times 3$  matrix, which can be written as:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \quad (6)$$

As  $h_{33}$  is normalized to one, the DOF (Degree of Freedom) of this matrix is eight.  $h_{13}$  and  $h_{23}$  are translations in  $x$  and  $y$  direction respectively. Rotation and scaling factors are  $h_{11}$ ,  $h_{12}$ ,  $h_{21}$  and  $h_{22}$ , while  $h_{31}$ ,  $h_{32}$  are perspective transformation factors. We use the following objective function to estimate the geometric transformation parameters:

$$E = \sum_{i=1}^M [(x'_i - \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}})^2 + (y'_i - \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}})^2] \quad (7)$$

where  $M$  is the number of matched pairs after refinement. We estimate the parameters by optimizing the above function. These defines the homography matrix  $H$ .

#### IV. COLOR AND LUMINANCE COMPENSATION

Due to changes in illumination and object density at each viewpoint, color and brightness differ between images. Let  $L_c$  denote the radiance of a color channel. Assuming that the objects in the scene have Lambertian surfaces, the pixel color is proportional to the light radiance [11]:

$$P_c = \alpha \times L_c \quad (8)$$

Here,  $\alpha$  is a proportional coefficient. Let  $P_{i,c}(x)$  denote the pixel color at  $x$  in the overlapping area of the  $i^{th}$  image. Since the overlapping area contains the same scene, it can be described as follows:

$$\begin{aligned} P_{0,c}(x) &= \alpha_0 \times L_c \\ P_{1,c}(x) &= \alpha_1 \times L_c \end{aligned} \quad (9)$$

This suggests that two pixel values at the same position are related linearly. Considering noise and errors in local registration, this linear relation may be approximated by the average value of overlapping areas as follows:

$$\frac{\alpha_0}{\alpha_1} = \frac{\bar{P}_{0,c}}{\bar{P}_{1,c}} \quad (10)$$

Here,  $\bar{P}_{i,c}$  is the mean value of the  $i^{th}$  overlapping image. With the  $0^{th}$  image as the reference, the  $1^{st}$  one is:

$$P'_{1,c}(x) = \frac{\alpha_0}{\alpha_1} \times P_{1,c}(x) \quad (11)$$

This compensation is applied to the entire image. Colors and light densities of images are then globally normalized.

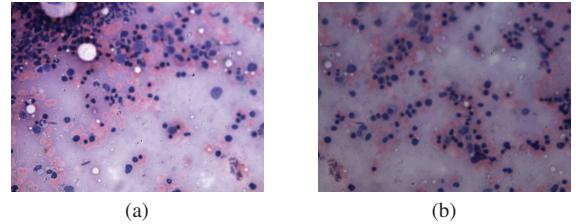


Fig. 9: Test images of red bone marrow.

#### V. STITCHED IMAGE SMOOTHING

Finally we smooth the image to blend the stitched boundaries. This is now straight forward after color and luminance compensation as described above. A weighted average method as in [12] is applied to all stitched boundaries. With the weighted coefficient  $\alpha$ , the pixel value in the overlapped area is:

$$I(i,j) = \alpha I_1(i,j) + (1 - \alpha) I_2(i,j) \quad (12)$$

with  $0 \leq \alpha \leq 1$ . When  $\alpha$  decreases gradually from 1 to 0, the pixel value in the overlapped area will change from  $I_1(i,j)$  to  $I_2(i,j)$ .

#### VI. EXPERIMENTS

The images in this paper were acquired using an Olympus BX41 microscope which is equipped with a motorized precision stage. The positioning accuracy of the stage is  $3\mu m$ . All images were acquired using a 20X Olympus UIS2 objective lens. The resolution of every image is  $1600 \times 1200$ . Figure 9 shows two test images of red bone marrow with obvious differences in color and brightness.

As shown in Figure 10a, we can see that the images are not well aligned by SAD (sum of absolute differences), where some blur is apparent in the right part of the image. SURF matching is invariant to variations in brightness. We can see that the alignment in Figure 10b is more accurate. Figure 11a shows an image obtained by stitching the image according to the transformation parameters. We can see the boundaries and differences in color and brightness. The processed image after color compensation and image smoothing is shown in Figure 11b. The boundaries are almost invisible. The brightness difference between these two images is small. We also tested our algorithm on another set of 6 images for mosaicing: Figure 12 shows a high-resolution ( $4331 \times 2302$ ) mosaic of microscopy images obtained with our method.

#### VII. CONCLUSIONS

We presented a method for mosaicing of microscopy images based on SURF. A new method of dividing descriptor window is proposed to increase matching speed. This results in a few more mismatched point correspondences. However, in general, this does not affect our final estimation, since mismatched are discarded in a subsequent refining step. Experiments on real microscopy datasets show that our algorithm produces mosaics much faster than preexisting methods, while a satisfactory quality is maintained.

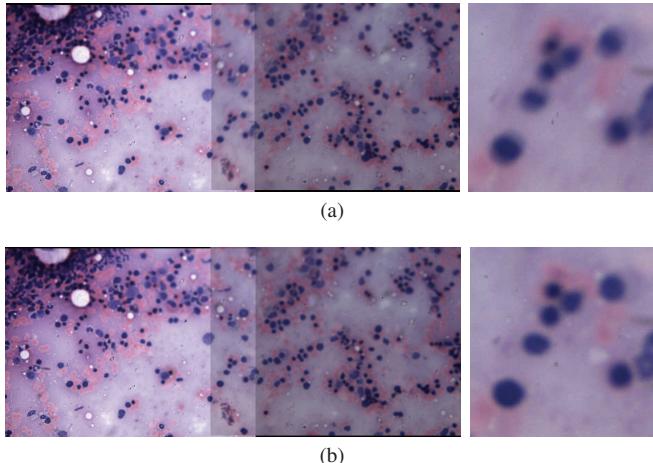


Fig. 10: Mosaic results: (a) SAD; (b) mosaic based on SURF.

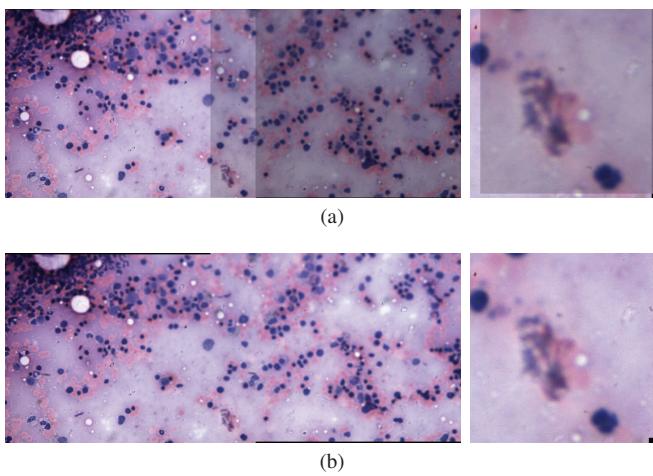


Fig. 11: Results of blending: (a) direct; (b) with color compensation and smoothing.

Future work includes building an automated microscope system incorporating microscope controls, image mosaicing and image browsing. The main contributions of this project are where improvements on robustness for fast matching and mosaicing and software development for microscopy systems.

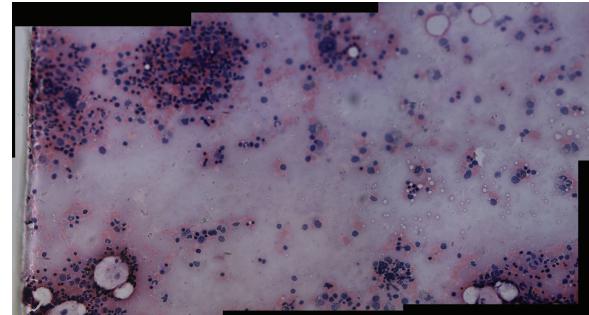


Fig. 12: Mosaic of six images.

#### ACKNOWLEDGEMENT

This work is partly supported by the Natural Science Foundation of China under No.60872119 and the NSF of Shandong China under No. 2009ZRB01675.

#### REFERENCES

- [1] D. Gledhill, G. Y. Tian, D. Taylor, and D. Clarke, *Panoramic imaging—a review*, Computers and Graphics, 2003, vol. 27, pages 435–445.
- [2] L. G. Brown, *A survey of image registration techniques*, ACM Computing Surveys, 1992, vol. 24, pages 325–376.
- [3] F. Huang, R. Klette, and K. Scheibe, *Panoramic Imaging*, Wiley, Chichester, 2008.
- [4] B. Zitova and J. Flusser, *Image registration methods: a survey*, Image and Vision Computing, 2003, vol. 21, pages 977–1000.
- [5] M. Brown and D. Lowe, *Recognising panoramas*, Int. Conf. Computer Vision, 2003, vol. 2, pages 1218–1225.
- [6] D. Lowe, *Distinctive image features from scale-invariant keypoints*, Int. Journal of Computer Vision, 2004, vol. 60, pages 91–110.
- [7] F. Crow, *Summed-area tables for texture mapping*, 1984, In Proc. SIGGRAPH, pages 207–212.
- [8] T. Lindeberg, *Feature detection with automatic scale selection*, Int. Journal of Computer Vision, 1998, vol. 30, pages 79–116.
- [9] H. Bay, A. Ess, T.uytelaars, and L. Van Gool, *SURF: Speeded up robust features*, Computer Vision and Image Understanding, 2008, vol. 110, pages 346–359.
- [10] M. A. Fischler, *Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography*, Journal of Communication Association Machine, 1981, vol. 24, pages 381–395.
- [11] S. J. Ha, H. I. Koo, S. H. Lee, N. I. Cho, and S. K. Kim, *Panorama mosaic optimization for mobile camera systems*, IEEE Transactions on Consumer Electronics, 2007, vol. 53, pages 1217–1225.
- [12] Y. Xiong and K. Turkowski, *Registration, calibration and blending in creating high quality panoramas*, Applications of Computer Vision, 1998, pages 69–74.
- [13] J. Zhang, B. Song, X. Feng, M. Li, and J. Kong, *A novel cartridge image mosaic approach*, Intelligent Information Technology Application, 2008, pages 472–476.
- [14] C. Sun, R. Beare, V. Hilsenstein, and P. Jackway, *Mosaicing of microscope images with global geometric and radiometric corrections*, Journal of Microscopy, 2006, vol. 224, pages 158–165.