33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Creating a Virtual Slide Map of Sputum Smears by Auto-Stitching

Bhavin Patel, Tania S. Douglas

Abstract—**Under a microscope, only a small field of a slide can be seen through the eye piece. A digital replication of a large section of the smear on a slide is desirable for several applications, including image-based auto-positioning for automated microscopy and telepathology. To achieve this, images of the various fields on the slide can be captured and stitched automatically to form a virtual slide map. We compared two auto-stitching methods, namely the geometric hashing and the scale invariant feature transform (SIFT), applied to images of tuberculosis-infected sputum smears. Experimental results show good agreement between the two methods.**

I. Introduction

Automation of microscopy for tuberculosis (TB) screening aims to speed up the screening process, to improve its sensitivity and to reduce its reliance on technicians. Automatic localisation of previously defined regions of interest on a slide may be used in automated microscopy to manually verify results, and to compare the image quality and bacillus detection accuracy in the same field on a slide at different settings of a microscope or across different microscopes. Automatic localisation can be achieved with the aid of a virtual slide map created by image stitching; a point of reference on the virtual slide map can be used to automatically bring desired fields on the slide to the field of view of the microscope.

Image stitching is the process of merging two or more overlapping images to produce a panorama or a composite image [1]. It is a powerful approach to creating a single composite image at high resolution of a large or whole section of a slide since a large section of the slide cannot be captured at high resolution at once even with a high resolution camera and at a low power objective. Overlapping images of the various fields on a slide can be captured manually using a manual *xy* stage or automatically using a motorise stage. Image stitching comprises two stages; image registration, which involves identifying the common overlap between adjacent images and the appropriate transformation relating the overlapping regions, followed by image merging using the resulting transformation [1].

We compare two auto-stitching schemes, namely the geometric hashing scheme (GHS) and the scale invariant feature transform (SIFT) for stitching Ziehl-Neelsen (ZN) sputum smear images.

Manuscript received March 24, 2011. This work was supported by the NIH/NIAID under Grant R21 AI067659-01A2.

Corresponding author: T.S. Douglas (Phone: +27 21 4066541; fax: +27 21 4487226; email: tania@ieee.org)

B. Patel and T. S. Douglas are with the MRC/UCT Medical Imaging Research Unit, Department of Human Biology, University of Cape Town, South Africa.

II. Methods

A. Scanning and Image Acquisition

We used a conventional bright field microscope, Zeiss Axioskop 2 with a 40x objective lens at 0.75 numerical aperture for image acquisition. A reasonable rectangular section of the smear on a slide was marked and the slide placed in the slide holder so that it is aligned to the x-axis of the *xy* stage. Both stage movement and focusing were manually performed. The accompanying Axiovision 4.7 software allows real time display of the field-of-view which was captured at 1030x1300 pixels using the attached Axiocam HR digital camera. Images were acquired sequentially with a 30-50% overlap between adjacent images to ensure good stitching results.

B. Automatic Image Stitching

Image stitching was performed in a non-continuous sequence as shown in Fig. 1 where the black outline represents the rectangular region on the slide and the red arrows illustrate the direction of image stitching. I_{rk} is the last image of row k and the total number of rows of images $= R = rk + 1$.

Let the captured images be labeled I_i, $1 \le i \le n$, where n is the total number of images captured for that slide. A reference frame was created which was large enough to fully contain the complete virtual slide map, V_n, which was constructed by sequentially stitching the next image, I_{j+1}, to the current partial virtual slide map, V_j.

$$V_{j+1} = V_j + I_{j+1}\ , \quad 0 \le j < n \tag{1}$$

The symbol + in this context refers to stitching of images. The first image, I_1, i.e. when $j = 0$, was placed and fixed near the top-left corner of the reference frame and this forms the first partial virtual slide map V_1.

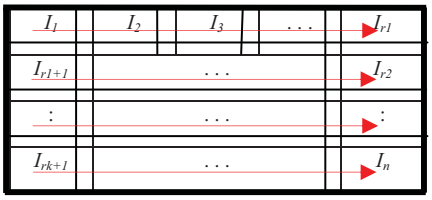


Fig. 1: Image stitching sequence.

Automatic construction of the virtual slide map requires automatic identification of the overlap region between V_j and I_{j+1}. An automatic matching scheme that detects common points between the two images was used to perform this task. However, for larger values of j, V_j would be very large relative to the image I_{i+1} and therefore the matching scheme would be prone to errors.

C. Image Stitching using a Small Portion of V_j

The matching process would be more accurate and faster if a smaller portion of V_j were considered in the matching process. Since prior knowledge was available as to which image needs to be stitched to which image, the portion-of-interest, POI_j, of V_j that shares an overlap region with the image, I_{j+1}, to be stitched was known and could be extracted from V_j. Using the sequence shown in Fig. 1 the position of the POI_j varies depending on the current configuration of V_j.

At a given time, V_j can be at one of three possible configurations shown in Fig. 2. Shown also is the POI_j that needs to be extracted as it shares a common overlap region with the next image to be stitched.

Common feature points between the POI_j and I_{j+1} were detected and these were used to compute a transformation linking the two. Prior to the scanning process, the camera was well aligned to the xy stage i.e. objects moved parallel to the x and y axis of the field-of-view. This guaranteed the correct movement of the stage, and also prevented the individual adjacent images from being rotated in relation to each other during image acquisition. Furthermore all images were taken at 40x magnification and therefore the scale factor between adjacent images was 1.

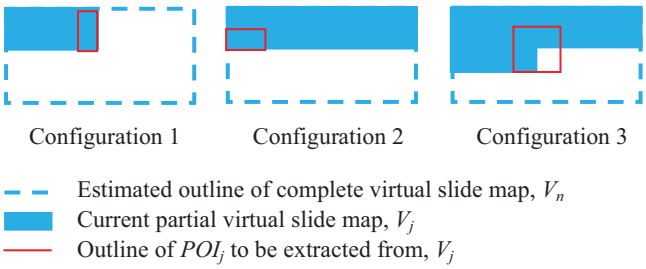


Fig. 2: Configurations of the V_j.

Therefore, the orientation and the scaling could be assumed to stay the same throughout the scanning and processing of the images, and hence the computed transformation relating POI_j and I_{j+1} consisted only of translation (parameters t_x and t_y).

Automatic identification of the overlap region requires a matching scheme (object recognition scheme) to detect the common feature points between images. A limited number of shared feature points in addition to feature point positional inaccuracies (due to feature detection errors caused by illumination changes, focus level changes and noise between adjacent images) affects the accuracy of the computed transformation and hence the performance of the stitching process. Using an overlap of 30-50% between adjacent images helps in reducing the impact of these factors.

The suitability of two auto-stitching methods, namely the scale invariant feature transform (SIFT) and the geometric hashing scheme (GHS) to stitch numerous ZN-stained sputum smear images was evaluated by comparing the resulting composite images.

D. Scale Invariant Feature Transform

Feature point extraction

SIFT keypoints were extracted directly from POI_j. Firstly, potential interest points invariant to scale were detected using the difference-of-Gaussians [2] which is calculated for two nearby scales of an image, I, by:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}$$

where k is a constant multiplicative factor used for changing the scale and x and y are the coordinates of a pixel in the image I. $L(x,y,\sigma)$ is computed as:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

where $G(x,y,\sigma)$ is the Gaussian filter for smoothing the image and σ is the width of the filter.

The keypoints were then detected by comparing each pixel in the image to eight neighbours in the same scale and nine neighbours in the neighbouring scales and if it was a local maximum or minimum, it was declared a candidate keypoint. Keypoints with low contrast, and hence high sensitivity to noise, were rejected using information that was derived by performing a detailed fit to the nearby data for location, scale, and ratio of principal curvatures [2].

Feature description

To achieve image rotation invariance, an orientation was assigned to each keypoint by computing a gradient orientation histogram in the neighbourhood of the keypoint. The scale of the keypoint was used to select the Gaussian smoothed image, L. For each image sample $L(x,y)$, the magnitude $m(x,y)$ and orientation $\theta(x,y)$ were computed as:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \tag{4}$$

Each keypoint got assigned a descriptor which is a 128 dimensional feature vector. These descriptors are invariant to similarity transformation and also partially to illumination [2]. Both the keypoint location and the descriptor of every keypoint in POI_j were stored in a database.

Image registration and merging

Similarly, SIFT keypoints were extracted from I_{J+1} followed by matching each keypoint independently to the database of keypoints. The matching keypoints formed a putative set of correspondences between POI_j and I_{j+1}. The random consensus sample (RANSAC, [3]) was executed to eliminate highly mismatched keypoints (outliers). The inliers obtained from RANSAC were used to re-compute a better estimation of the translation parameters of the transformation in the least-squares sense. Using the transformation mapping, image I_{j+1} was stitched by replacing the common overlapping region between them by only that of Image I_{j+1}. The process was repeated until the entire virtual slide map, V_n, was constructed.

E. Geometric Hashing Scheme (GHS)

Feature point extraction

Feature points were extracted from POI_j via a segmentation process, which separates tuberculosis bacilli from the background. In ZN-stained sputum smear images, bacilli stain red against a blue background. A quadratic pixel classifier was employed in which training image pixels are used to establish a quadratic mapping between image pixels and labels. The discrimination is drawn using the class mean and covariance matrices [4]. Owing to the long and thin shape of TB bacilli, area and eccentricity filters were used to remove non-bacillus objects from the segmented images. The medial axis transform [5] was employed to extract the topology skeletons of the objects in the filtered segmented images. The branch points of the branched skeletons and the mid-point of branchless skeletons were extracted as the feature points (red dots in Fig. 3).

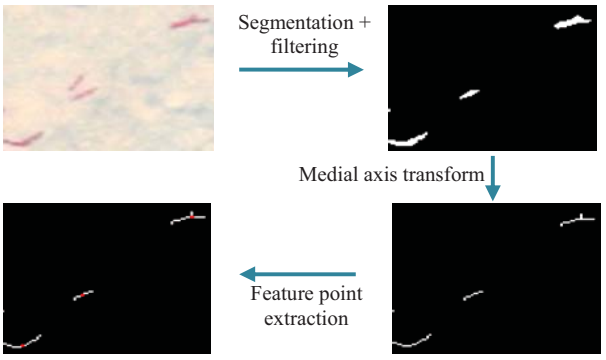


Fig. 3: Feature extraction.

Image representation

The extracted feature points were collectively used to represent the image invariantly to similarity transformation using the geometric hashing technique [6]. Let $\{m_1, m_2, \ldots, m_k\}$ be the feature points in POI_j. To form a co-ordinate frame, a pair of ordered points was selected. This pair of points is termed the basis of the 2-D co-ordinate frame. If the basis is formed by the ordered pair points (m_1, m_2), then a vector $(m_2 - m_1)$ and another vector at the midpoint of m_1 and m_2 and perpendicular to it form a co-ordinate frame as shown in Fig. 4. The perpendicular vector is formed by rotating vector $(m_2 - m_1)$ by 90°. The origin of the coordinate frame will be at the midpoint of m_1 and m_2.

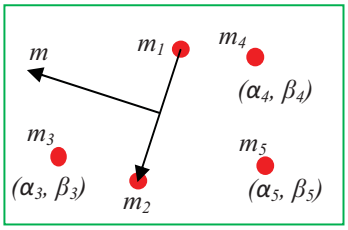


Fig. 4: Image representation using GHS.

In this frame, the coordinates (α_j, β_j) of each of the other feature points m_j in the POI_j for $3 \leq j \leq k$, where k is the number of feature points present, satisfy the equation:

$$m_j = \frac{m_1 + m_2}{2} + \alpha_j (m_2 - m_1) + \beta_j (m - m_1) \qquad (5)$$

where m is the end point of the vector obtained by rotating vector $(m_2 - m_1)$ by 90°. This essentially re-scales the image so that the magnitude of the vector $(m_2 - m_1)$ is equal to 1 [6]. The corresponding entry (m_1, m_2) was then stored in a hash table bin at indices (α_j, β_j) , $3 \leq j \leq k$. The process was repeated for every possible pair of ordered points i.e. basis. The entries of the form (m_u, m_v) in the hash table are an invariant representation of POI_j.

Image registration and merging

To execute the registration process, I_{j+1} was segmented, filtered, and feature points extracted similarly to POI_j. An arbitrary basis, B_q, in I_{j+1} was selected based on the overlap region shared between it and the extracted POI_j which is dependent on the configuration of V_j. This ensured that the overlapping regions were correctly matched. For configuration 1, both the chosen points making up B_q were located in the left half of I_{j+1} since only the left half of I_{j+1} overlaps with POI_j. Similarly, for configuration 2, both the chosen points were located in the upper half of I_{j+1} and for configuration 3, one of the chosen points was located in the left half while one in the upper half of I_{j+1}.

The invariant coordinates (α_j, β_j) of the other feature points in I_{j+1} in the coordinate frame formed by B_q were computed and used as indices to assess the respective hash table bins. Every entry in the bins was given a vote and entries that received at least 40% of the maximum vote were candidate basis pairs (*CBs*) matching B_q and hence were carried forward for verification. For a given *CB*, an approximate transformation was computed using the *CB* and the B_q which are essentially two point-to-point correspondences. Using this transformation along with

Voronoi tessellation [7], a putative set of corresponding points between POI_j and I_{j+1} were determined. RANSAC [3] was executed to eliminate outliers and the inliers were used to re-compute a better estimation of the translation parameters of the transformation in the least-squares sense. The process was repeated for every *CB* and the *CB* with the highest number of inliers was the best match to B_q with the corresponding least-squares-fit transformation. Using this transformation mapping, image I_{j+1} was stitched by replacing the common overlapping region between it and POI_j by only that of Image I_{j+1}. The process was repeated until the entire virtual slide map, V_n, was constructed.

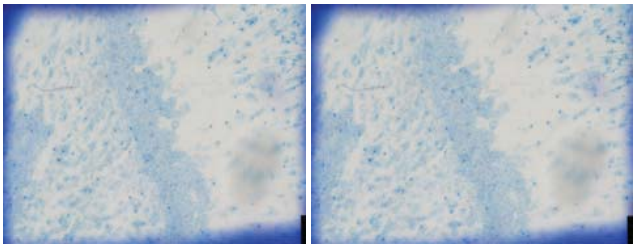
III. RESULTS AND DISCUSSION

A rectangular region measuring 1.6 x 1.1 mm on a ZN-stained sputum smear slide was scanned which produced 62 images. Two virtual slide maps were constructed, one with SIFT, named SIFT VS, and the other using GHS, named GHS VS. The resulting SIFT VS measured 4049 x 5976 pixels while the GHS VS measured 4044 x 5979 pixels. The almost identical sizes indicate very little discrepancies exist between the two VSs. Furthermore, there were no detectable differences at and near the visible seam lines upon visual comparison of the two VSs. Fig. 5 shows the two VSs.

If the stitching quality of both the methods is same then corresponding triangles - defined by 3 pairs of corresponding points - in the two VSs are expected to be congruent. A triplet of points was selected in the SIFT VS and the corresponding points in the GHS VS were found. The three points were joined to generate a triangle and the ratio between the lengths of the three sides were computed and compared [1]. Congruent triangles have the same size and shape and hence they also have the same perimeter. Therefore, perimeters of the corresponding triangles were also compared.

A set of 1140 corresponding triangles in the SIFT VS and the GHS VS were used for this analysis. The triangles varied in size and covered various regions in the two VSs.

Table I shows the results obtained for 5 randomly chosen pairs of corresponding triangles.



(a) SIFT (b) GHS

Fig. 5: Virtual slide maps using (a) SIFT and (b) GHS

TABLE I: QUANTITATIVE COMPARISON OF SIFT AND GHS AUTO-STITCH.

Triangle	Ratio of sides in triangle	
	In SIFT VS	In GHS VS
1	1 : 2.799 : 1.799	1 : 2.800 : 1.801
2	1 : 2.119 : 1.193	1 : 2.119 : 1.193
3	1 : 4.399 : 3.791	1 : 4.404 : 3.798
4	6.667 : 6.985 : 1	6.665 : 6.984 : 1
5	2.050 : 1.395 : 1	2.049 : 1.394 : 1

As seen in Table I the discrepancies between the ratios of the triangles from the two VSs are extremely small. The same was observed for all the other triangles. The average difference between the perimeters of corresponding triangles was 4.08 pixels with a standard deviation of 3.52 pixels, which is visually undetectable.

IV. CONCLUSION

A digital replication of a large section of a ZN-stained sputum smear slide can be created by auto-stitching individual images of the different fields on the slide. The GHS and SIFT auto stitching schemes were compared and the results showed negligible differences, indicating high similarity of stitching quality between the two.

REFERENCES

[1] B. Ma, T. Zimmermann, M. Rohde, S. Winkelbach, F. He, W. Lindenmaier and K.E.J. Dittmar, "Use of autostitch for automatic stitching of microscope images," *Micron*, vol. 38, no. 5, pp. 492-499, July 2007.

[2] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, Nov. 2004.

[3] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381-395, June 1981.

[4] R. Khutlang, S. Krishnan, R. Dendere, A. Whitelaw, K. Veropoulos, G. Learmonth and T. Douglas, "Classification of mycobacterium tuberculosis in images of ZN-stained sputum smears," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, pp. 949-957, July 2010.

[5] H. Blum, "A transformation for extracting new descriptors of shape," *Models for the Perception of Speech and Visual Form*, vol. 19, no. 5, pp. 362-380, 1967.

[6] H.J. Wolfson and I. Rigoutsos, "Geometric hashing: An overview," *IEEE Computational Science & Engineering*, vol. 4, no. 4, pp. 10-21, Oct. 1997.

[7] De Berg, M., O. Cheong, M. Van Kreveld and M. Overmars, *Computational Geometry: Algorithms and Applications.* New York: Springer-Verlag, 2003.