A ROBUST METHOD FOR INTER-MARKER WHOLE SLIDE REGISTRATION OF DIGITAL PATHOLOGY IMAGES USING LINES BASED FEATURES

Anindya Sarkar, Quan Yuan, Chukka Srinivas

Ventana Medical Systems, Inc. a member of the Roche Group, Mountain View, CA

ABSTRACT

An automated image registration method is proposed for registration of digital whole slide scans of adjacent tissue sections. An example workflow is where a pathologist annotates a tumor region in a primary (H&E [1]) stained image and the corresponding regions, in nearby sections, stained with other biomarkers, are identified and analyzed for aiding in clinical diagnosis. The different markers convey different clinical information leading to a more comprehensive analysis. This registration process, if done manually, is tedious and time-consuming. We propose a two-pass algorithm. A lines-based approach, computed from tissue boundary regions, returns the global transformation. Then normalized correlation, based on gradient magnitude images, is used for more precise local matching in a multi-resolution approach. The algorithm is robust to flips, large rotation angles, tissue wear-and-tear, and Area of Interest ([2]) mismatch, and outperforms current state-of-the-art registration methods for digital pathology images.

Index Terms— registration, color deconvolution, line fitting, robust estimation, multi-resolution search

1. INTRODUCTION

In standard digital pathology applications, pathologists mark regions of interest (will be referred to as Fields of View, FOV, subsequently) in H&E images, and then the same regions are located in adjacent sections, stained with one/more IHC (immunohistochemistry: a form of secondary marker) markers. In terms of stain color, H&E and IHC are different, and the structural similarity between adjacent sections is relied upon to locate similar regions. Examples of H&E-IHC and H&E-fluorescent registration are shown in Fig. 1. Thus, the framework is useful for different stains (markers) and imaging modalities (brightfield and fluorescent)[3].

Generic registration approaches involve computing keypoints and matching the keypoint descriptors, when we consider different viewpoints/color transformed versions/ of the same image. Here, the adjacent sections are different physical entities, though being structurally similar. The tissue mask

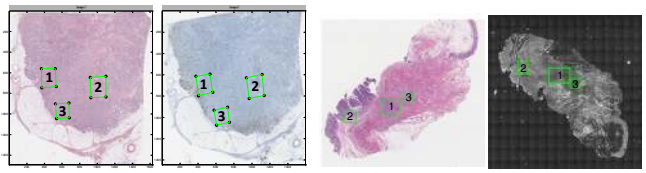

Fig. 1. left pair: (a) H&E-IHC registration; right pair: (b) H&E-fluorescent registration. The numbered regions are FOVs mapped from one image to the other.

obtained from nearby slices may partly differ and matching of entire binary tissue masks may not properly align the adjacent slices. We model the boundary regions of the tissue with smaller line segments, and then use robust measures to match the two sets of line-segments to obtain the overall global transformation (coarse matching). Line-matching allows for partial shape matching and the orientation information makes it more effective than point-set registration. We also use the lines-based matching to account for AOI mismatches. As the tissue deformation between adjacent sections can be non-rigid, a second module of finer sub-image registration (normalized correlation based block matching on gradient magnitude image) is used to compute local refinements between the globally aligned images. The inter-slice registration has been well studied in various medical imaging domains [1, 2]. In [2], correlation based block matching is used for registration but it accounts only for smaller rotation angles and robustness against flips and wear-and-tear have not been demonstrated.

Contributions: The key novelty in our proposed registration framework is in computing line-based features along tissue boundary regions, and then proposing robust estimators to obtain the global transformation. Our proposed method:
(i) handles large global transformation (180^o rotation is shown in Fig. 6(a)), translation and flips[4], as in Fig. 1(b), and internal tissue wear-and-tear (Fig. 6(b))
(ii) handles AOI mismatches - *line-based boundary matching works when more than 50% of extracted lines match (median based distance is used in matching)*
(iii) the boundary based lines-matching is insensitive to internal content; hence, when lines-based matching yields similar scores for multiple orientations (e.g. for shapes symmetrical

[1] H=hematoxylin stain (blue), E = eosin stain (red)

[2] AOI is the area on physical slide scanned in high resolution

[3] certain sub-cellular structures show up in fluorescent images and not in brightfield ones - thus helping in diagnostic process

[4] flipping of the tissue on a slide is possible as slide preparation, staining, and laying it on a scanner-bed can cause wear-and-tear or flips

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 1, 2013.

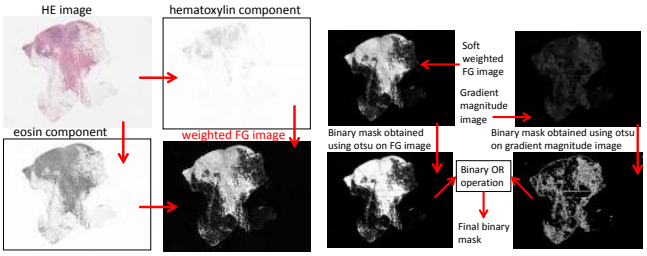

Fig. 2. left(a): soft weighted FG computed using color deconvolution, right(b): OR-ing of masks based on stain and gradient provides final tissue foreground mask

about origin), edge-map based matching, based on internal structure, is used to determine the optimal transformation
(iv) for the second pass, robust criteria are used to decide if high resolution internal structure provides more precise matching; hence, finer resolution matching does not deteriorate the registration even after internal wear-and-tear.

2. REGISTRATION WORKFLOW

The algorithm flow for global alignment is described below:
• A soft-weighted foreground (FG) image is computed from the input image. A region is assigned a higher weight in the FG image where the stain contribution is higher (considering the two dominant stains) or the gradient magnitude image G, computed on the grayscale image, is stronger.
• A binary mask is obtained from the FG image, and line-based features are computed along its boundary regions.
• Transformation parameters are computed between two sets of line-features - the transformation is expressed through rotation, reflection (horizontal/vertical flip) and translation.
• 1^{st} image is transformed to globally align it with 2^{nd} image.

2.1. Foreground Image mask computation

We need an accurate estimate of the FG image so that line-based features can be computed along its boundary. The effective binary mask image for the FG is obtained by OR-ing of 2 binary masks - (i) one mask is obtained after OTSU [3] thresholding of a stain-based soft-weighted FG image, and (ii) another mask is obtained after OTSU thresholding of a gradient magnitude image. The soft weighted FG image is computed based on the stain components in the image (Fig. 2(a)), using the assumption that the foreground regions have higher contribution of the stain components as compared to the background. We also use the gradient magnitude image to detect the foreground - there can be foreground regions where the stain contribution is very faint and the gradient magnitude can help distinguish between fainter foreground regions and smoother background regions (see Fig. 2(b)).

Computation of soft weighted foreground image: Let us consider the soft weighted FG computation (Fig. 2(a)) for a primary stained HE image, where the primary components are hematoxylin (H channel) and eosin (E channel). The first step is color deconvolution [4] to separate the image into H and E stain components. For each of the H and E channels, we computed a weighted H (W_H) and weighted E (W_E) channel image, respectively, and the final weighted HE image is expressed as max(weighted H, weighted E), per pixel. In the H (or E) image, darker regions (lower intensity values) correspond to regions where the H contribution is higher ((Fig. 2(a))). Hence, W_H is higher for lower values of H and vice versa. We empirically use a linear mapping to compute W_H(or W_E) from H(or E), whose steps are:
(1) Let δ_H = OTSU [3] threshold computed on H,
(2) let $\delta_L = f.\delta_H$ (empirically f = 0.8) where for $H < \delta_L$, W_H is set to 1
(3) let $\delta_M = \max(H)$ denote maximum value in H channel.

$$W_H(i,j) = \begin{cases} 1 & \text{if } H(i,j) < \delta_L \\ 1 - \left(\frac{H(i,j)-\delta_L}{\delta_M - \delta_L}\right) & \text{otherwise} \end{cases} \quad (1)$$

THis is just a linear rescaling with intensity truncation. Why does this help and why not truncate the higher intensity?

For IHC images, we use the same principles of color deconvolution (the two dominant stains are known) and obtain a weighted FG image using similar linear weighting on the deconvolved stain components.

THis is just a linear rescaling with windowing.

2.2. Computing Line based features

To find correspondence between images from adjacent tissue slices, we need a feature to represent the structural similarity between them. It has been empirically observed that tissue wear-and-tear results in more significant changes in the internal parts of the tissue as compared to the boundary regions. The line feature estimation is motivated from [5], where the focus was on wavelet-based filtering to obtain dominant orientations - our focus is on using high gradient points as candidates for line fitting and then random sampling is used for robust matching between two sets of lines.

The image is broken into 80x80 sized windows (a line segment is computed per window), with window shifts of 40 pixels along x and y axes. Let us consider the local line segment computation for a single window. Let there be N boundary points $\{x_i, y_i\}_{i=1}^{N}$. For every two points, a line segment model can be computed. We perform a random sampling (as in RANSAC [6]) of the point pairs. On a pair of random sampled points (x_m, y_m) and (x_n, y_n), we fit a line $(x_m, y_m)$to (x_n, y_n) and measure the total weights of gradient points within a distance (± 2) to the line. Then among all randomly sampled pairs, the line with greatest weight is retained. For windows where the best fitting line is of low weight, no line is fitted. The line features are shown in Fig. 3.

2.3. Computing Transformation Using Line Features

Assume that there are N_1 lines for image 1 with line centers at $(x_{i1}, y_{i1})_{i=1}^{N_1}$, and line angles (θ_{i1}) and strength (M_{i1}); the line segments are sorted in descending order of (M_{i1}). Similarly, for the 2nd image, we assume N_2 lines with centers

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 1, 2013.

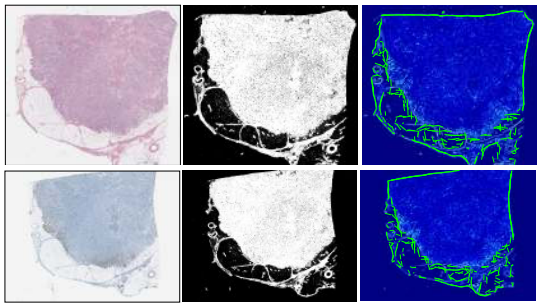

Fig. 3. left to right: stained image, binary FG mask, fitted lines superimposed on the gradient magnitude image (plotted in colormap jet domain in MATLAB); top example - H&E image; bottom example - IHC image

at $(x_{i2}, y_{i2})_{i=1}^{N_2}$, angles (θ_{i2}), and strengths (M_{i2}); where the line segments are sorted in descending order of (M_{i2}). The aim is to obtain the translation $(\Delta x_{opt}, \Delta y_{opt})$ and rotation angle (θ_{opt}); and also find the sub-part of image 1 which best matches to a sub-part of image 2.

AOI Mismatch problem: While scanning, the regions of a physical slide which get picked up for high resolution scanning constitute the Area of Interest (AOI). When two adjacent slices are scanned, the AOI picked up may differ between two slides. When we are registering two images, our first aim is global registration where one image is transformed so as to align it completely with the second image - this alignment requires that the proper subsets are selected in both images.

in other words, rigid with and without flipping.

In the global registration model, we allow two cases, for each of which we compute rotation (θ), and shifts $(\Delta x, \Delta y)$
(1) rotation + x-y translation (no flipping),
(2) flipping (horizontal/vertical) + rotation + x-y translation.

Algorithm 1 explains the lines-based matching process.

Robustness: We compute top 10 minima in S (see Algo. 1) and if they are all similar (e.g. all θ values are within $\pm 5^o$ of each other), we compute edge-map for the competing cases, exploit internal structure, and select the transformation condition which results in minimum Chamfer distance [7]: For 2 edge-maps E_1 and E_2, Chamfer distance $\mathcal{C}(E_1, E_2)$ is given by: $\mathcal{C}(E_1, E_2) = E_1.*(\mathcal{B}(E_2)).+E_2.*(\mathcal{B}(E_1))$, where (.* denotes inner product), and $\mathcal{B}(E)$ denotes the distance transformed image computed from the edge-map E.

Finding Matching Subsets in both Images:
(a) For point (x_{i1}, y_{i1}) in image 1, suppose it gets mapped to (x'_{i1}, y'_{i1}) after transformation. We select line-centers in image 1 which are close enough (empirically assumed to be 10 pixels) to line-centers in image 2 after transformation - these are "representative" points and their enclosing rectangles provide the relevant subset per image.
(b) The set of representative points in image 1:
$P = \{(x_{k1}, y_{k1}) : \min_i d(\{x'_{k1}, y'_{k1}\}, \{x_{i2}, y_{i2}\}) \leq 10\}$.
(c) The relevant bounding box in image 1 is given by the points in P; while the relevant bounding box in image 2 is given by the corresponding matching line-centers in image 2.

Algorithm 1 Computation of the optimal transformation based on line matching by exhaustive search

Input: Let L_1 and L_2 be set of lines computed from images 1 and 2, respectively. Consider the top T_1 lines in L_1 and the top T_2 lines in L_2 (based on gradient magnitude)
Output: determine the condition (flip or not) which ensures best lines-matching along with transformation parameters $(\Delta\theta_{opt}, \Delta x_{opt}, \Delta y_{opt})$

1: Let $S(i, j, k)$ denote the lines-based matching cost between i^{th} line in L_1, and j^{th} line in L_2, and we consider 2 possible angles (k = 0,1) between 2 lines (considering 180^o shifted version of the 2nd line)
2: **for** every triplet (i, j, k), consider i^{th} line in L_1, j^{th} line in L_2, k = 0, 1 depending on angle considered **do**
3: $k = 0 \Rightarrow \delta_r = \theta_{i1} - \theta_{j2}$ (in radians),
4: $k = 1 \Rightarrow \delta_r = \theta_{i1} - (-sign(\theta_{j2}).(\pi - |\theta_{j2}|))$
5: For each triplet (i, j, k), rotate lines in L_1 by δ_r, and the line centers (x_{i1}, y_{i1}) get transformed to (x'_{i1}, y'_{i1})
6: translations $dx = x'_{i1} - x_{j2}$, $dy = y'_{i1} - y_{j2}$
7: $S(i, j, k)$ = median of (minimum distance from every line center in set 1 to nearest transformed line center in set 2), where distance = (L_2 distance between centers) + α.(angle difference in radians); (we use $\alpha = 1/3$)
8: consider flipped line-centers to compute S term for horizontal/vertical flips
9: **end for**
10: **return** for the transformation case which results in minimum cost S_{min}. return the rotation($\Delta\theta_{opt} = \delta_r$), x-y shifts($\Delta x_{opt} = dx$, $\Delta y_{opt} = dy$) and flips (if any)

Why restrict to 0 and 180? Why not look at increased granularity, i.e. 0, 90, 180, 270?

2.4. Multi-Resolution Matching for Finer Registration

Fig. 4 shows how image 1 can be transformed so as to globally align it with image 2. For registering two 20x whole slide images, their low resolution versions (say 1x images) are considered sufficient - they are used for foreground estimation, line feature computation, and subsequent line-matching based transformation estimation.

For the 3 FOVs marked in Fig. 4(a), we map them to the transformed image grid (Fig. 4(b)), and then consider a certain window around it in image 2 (denoted by black rectangles in Fig. 4(c)). Normalized correlation is used in the gradient magnitude domain to match the FOVs in Fig. 4(b) to the shifted windows obtained in the search regions, as in Fig. 4(c). The maximum of the correlation matching based search is chosen unambiguously - *once top 5 maxima points are obtained, the standard deviation of their x-y locations is considered and if it is less than a certain threshold (e.g. 3 units) in both dimensions, it indicates that all 5 maxima are placed closely enough - then only the maximum location is used.* There are situations involving tissue wear-and-tear (FOV-1 in Fig. 6(a), FOV-3 in Fig. 6(b)) where local matching will not be useful - this criterion helps avoid such cases. The

There are several heuristically chosen parameters which raises concerns about

Table 1. Concordance of registration methods with ground truth annotations with scores expressed as statistics of difference in pixels: *performance wise, adjusted* > *fine* > *coarse*: our coarse method provides the most fair comparison to other cited methods as only the global alignment is considered

Method	Variants	median	mean±std
our proposed method	**coarse**	**14**	**17±13**
	fine	**7**	**9±10**
	adjusted	**6**	**8±7**
point-set registration	TPS-L2[9]	16	86±404
	TPS-KC[10]	16	85±400
Elastix[8] using default configurations	Affine	114	167±408
	Rigid	105	163±410
	Bspline	181	243±417

search for the best matching window takes place in a multi-resolution approach and with each resolution, the search window is halved to reduce the computational complexity. Usefulness of the second-pass matching is shown in Fig. 5.

Let σ_x and σ_y be the std. deviation between (coarse to fine) annotation differences along x-y axes. "Adjusted" registration limits annotations to within $\{\pm(2 * \sigma_x), \pm(2 * \sigma_y)\}$ of coarse matching based annotations along x-y axes, to limit errors due to incorrect/ambiguous second-pass matching.

2.5. Results Comparison and Analysis

elastix is an image registration algorithm dealing with physical spacing on the order of brains, lungs, etc. However, these images are tiny. How did the authors ensure that the algorithm was used correctly?

For validation, we have compared our results (Table 1) with Elastix[8](`http://elastix.isi.uu.nl/`), robust point-set matching methods, using mixture of Gaussians (TPS-L2)[9] or using correlation (TPS-KC)[10]. Our results are markedly superior to [8] and slightly superior to [9, 10]. For [9, 10], the input point-sets for the registration are our line-centers. Thus, the comparable results of our method with [9, 10] show the utility of our boundary extraction and line-features. For each (rectangular) annotation region, we have manually marked the top-left and bottom-right corners for the retreived annotation, and we compare them with the top-left and bottom-right of the annotations retreived for the various methods. We compare the (median, std. deviation and mean) for the L_2 distance between manual and retreived annotation region corners in Table 1 - the statistics are computed over 40 HE-IHC image pairs, where each image is about 2500x2500 pixels in size, and there are 5-10 annotations per HE image.

3. CONCLUSION

The various modules of our two-pass algorithm are parallelizable - computing line features in different windows, matching of different pairs of line-segments, searching for best window using correlation in finer registration mode. For a 8-core 1.87 GHz Intel processor with 16GB RAM, the global registration for 2 20x images takes about 1.2 sec (considering 1x images for matching), and assuming 4 FOVs, a further 1 sec is needed for the 2nd pass. In the future, we propose to discard noisy lines, pen-marks and artifacts, which distort the line-feature computation which may corrupt the global alignment.

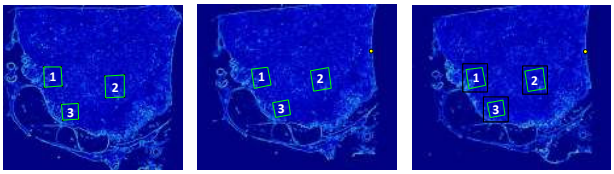

Fig. 4. gradient magnitude images shown with marked FOVs; from left-to-right: (a) image 1; (b) image 1 aligned with image 2; (c) image 2: the black window shows the search window around annotations retrieved after global alignment

What search windows are these?

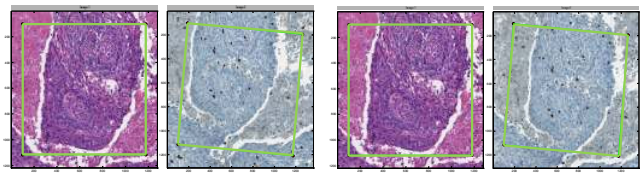

Fig. 5. (left-to-right): FOV matching (a) after global registration; (b) refinements after finer multi-resolution matching

4. REFERENCES

[1] D. Mueller, D. Vossen, and B. Hulsken, "Real-time deformable registration of multi-modal whole slides for digital pathology," *Computerized Medical Imaging and Graphics*, vol. 35, no. 7, pp. 542–556, 2011.

[2] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," in *Proc. of MICCAI*. Springer, 2000, pp. 557–566.

[3] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[4] A. Ruifrok and D. Johnston, "Quantification of histochemical staining by color deconvolution.," *the Intl. Academy of Cytology*, vol. 23, no. 4, pp. 291–299, 2001.

[5] Y. N. Wu, Z. Si, H. Gong, and S.C. Zhu, "Learning active basis model for object detection and recognition," *Intl. Journal on Computer Vision*, vol. 90, no. 2, pp. 198–235, 2010.

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

[7] G. Borgefors, "Distance transformations in digital images," *Comp. vision, graphics, and image proc.*, vol. 34, no. 3, pp. 344–371, 1986.

[8] S. Klein, M. Staring, K. Murphy, M.A Viergever, J.PW Pluim, et al., "Elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. on Med. Imag.*, vol. 29, no. 1, pp. 196–205, 2010.

[9] B. Jian and B. C Vemuri, "Robust point set registration using gaussian mixture models," *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, vol. 33, no. 8, pp. 1633–1645, 2011.

[10] Y. Tsin and T. Kanade, "A correlation-based approach to robust point set registration," in *Proc. of ECCV*, pp. 558–569. Springer, 2004.

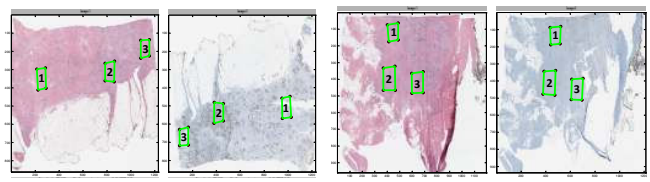

Fig. 6. robust under (a) 180^o rotation, (b) wear-and-tear

Preprint submitted to 2014 IEEE International Symposium on Biomedical Imaging.
Received October 1, 2013.