

# Automatic Registration of Serial Mammary Gland Sections

I. Arganda-Carreras<sup>1,2</sup>, R. Fernandez-Gonzalez<sup>2,3</sup>, C. Ortiz-de-Solorzano<sup>2</sup>

<sup>1</sup>Biocomputing Unit, National Centre of Biotechnology, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Madrid, Spain

<sup>2</sup>Bioimaging Group, Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

<sup>3</sup>UC Berkeley / UC San Francisco Joint Graduate Group in Bioengineering, Berkeley, CA, USA

**Abstract**—We present two new methods for automatic registration of microscope images of consecutive tissue sections. They represent two possibilities for the first step in the 3-D reconstruction of histological structures from serially sectioned tissue blocks. The goal is to accurately align the sections in order to place every relevant shape contained in each image in front of its corresponding shape in the following section before detecting the structures of interest and rendering them in 3D. This is accomplished by finding the best rigid body transformation (translation and rotation) of the image being registered by maximizing a matching function based on the image content correlation. The first method makes use of the entire image information, whereas the second one uses only the information located at specific sites, as determined by the segmentation of the most relevant tissue structures. To reduce computing time, we use a multiresolution pyramidal approach that reaches the best registration transformation in increasing resolution steps. In each step, a subsampled version of the images is used. Both methods rely on a binary image which is a thresholded version of the Sobel gradients of the image (first method) or a set of boundaries manually or automatically obtained that define important histological structures of the sections. Then distance-transform of the binary image is computed. A proximity function is then calculated between the distance image of the image being registered and that of the reference image. The transformation providing a maximum of the proximity function is then used as the starting point of the following step. This is iterated until the registration error lies below a minimum value.

**Keywords**—Automatic registration, Image processing, Biotechnology

## I. INTRODUCTION

A correct visualization and quantification of the morphology of the mammary gland, in a state as similar as possible to the specimen living conditions, is basic in the study of normal mammary gland biology and its neoplastic variants (i.e. breast cancer). With this aim, we reconstruct mammary gland epithelial structures from fully sectioned paraffin tissue blocks, stained with histological markers. Our lab has developed a computerized microscopy system [1] that can scan entire sections of mammary gland tissue

---

This work has been supported by the US Department of Defense and California Breast Cancer Research Programs under grants DAMD17-00-1-306 and 8WB-0150 (respectively), by the LBNL Laboratory Directed Research and Development Program, under contract CSLD12 and by the National Institute of Health under grant 1R01HL67465-01. I. Arganda-Carreras is being supported by a predoctoral FPU-CAM fellowship since October 2003.

blocks. After scanning the images, the system provides manual and automatic tools to detect structures of interest and reconstruct them in 3D, under a unified in-house developed JAVA user interface. To achieve accurate 3-D reconstructions, proper alignment of the images is needed, in order to place every relevant 2-D structure in one section in front of its correspondent 2-D structure in the following section.

The challenge we face is to perform this image alignment process on sections obtained using a predominantly manual sectioning protocol (Manual sectioning can cause non-linear distorting effects, such as tissue stretching, folding and/or tearing). To further difficult the registration, we have to account for significant differences between the same structures in consecutive sections, caused by the distance between the sections (i.e. the resolution), which is normally larger than in other 3D image modalities.

This first step of our registration algorithm aims at the best rigid body-transform for each pair of images of consecutive sections. Although this method does not ensure perfect registration, it is usually enough for consolidating the tri-dimensional reconstruction, or as a starting point to be refined locally using non-linear registration techniques.

A rigid body registration can be done manually, by marking alternatively three points in every image and calculating the rigid body-transformation (i.e. rotation plus translation) that minimizes the lineal quadratic error between both sets of points. However, manual registration of hundred of sections, as required when working with complete cases of fully sectioned tissue blocks, could become very slow and tedious.

In this paper we present an automated method for the rigid body registration of tissue sections. The examples used to illustrate the algorithm use mammary gland tissues sections; nevertheless, the same algorithm can be successfully used on other tissue sources and imaging modalities.

## II. METHODOLOGY

We used paraffin embedded tissue blocks containing either normal mouse mammary glands or human tissue biopsies of patients with ductal carcinoma in-situ (DCIS) of the breast. The blocks were fully sectioned at 5 microns thickness and stained with Hematoxylin and Eosin (H&E), which highlights histological structures such as mammary ducts or tumors. All the sections were imaged at low resolution (2.5X) using a monochrome CCD camera without

using any interference or absorption color filtering. The images were stored in either TIFF or ICS [2] format in sets of related images that we refer to as ‘cases’.

To register each pair of images we used a *Hierarchical Chamfer Matching Algorithm* [3] to determine the similarity between the image to be aligned (or target image) and the reference image.

Briefly, the reference image was transformed in a distance image, meaning that the edges of the structures of the image receive a value of white (255) and the rest of the pixels present gradually smaller values depending on their increasing distance to the edges. Once the distance image is calculated, the system matches it with a binary image obtained by applying an automatic thresholding function to the gradient –using the Sobel transform- of the image to be aligned. This contour image contains the most relevant edges and contours of the images. Before applying the matching function, the contour image is rigid-body transformed with three degrees of freedom: translations in the x- and y- directions and rotation around the x-axis. Given  $p_c$  set of pixels of the thresholded contour image and  $p_d$  set of pixels of the distance image, the matching function value is defined as:

$$m(d, x, y) = \sum_{p_c[i] \neq 0} p_d[j], \forall i \in p_c \quad (1)$$

Where  $d$ ,  $x$  and  $y$  are respectively the angle of rotation and the length of x- and y- translations, and  $j$  is the correspondent pixel index in the contour image after applying the transformation over the index in the distance image ( $i$ ).

Since the proximity –matching- function is calculated as the sum of the pixel values of the distance image that match with white pixels in the transformed contour image, perfect matching correlates with a function maximum. Therefore, when this sum is maximized, the best transformation –and consequently alignment - is achieved.

Considering the large size of the images being processed –usually around 6000x6000 pixels in average- and the number of images to be registered –in the hundreds- it would be inefficient to apply the matching algorithm to full resolution images. Therefore, we approach the functional maximum using decreasingly subsampled versions of both the target and reference images, which will produce an important reduction in the dimension of the images, and consequently in the computational time. The search for the local maximum starts on versions of the images subsampled by a factor of  $2^n$ . We calculate the best match between the images after applying all possible rotations and translations to the target image. Namely, the system uses all possible rotations from  $0^\circ$  to  $350^\circ$  in steps of  $10^\circ$  and all possible translations in steps that depend on the image size and guarantee at least 50% of image overlap. The alignment obtained using this brute-force search is then used as the starting point for the second step, where the range of translations and rotations is reduced to a neighborhood of the previous result, and the rotation and translation steps are decreased in order to better refine the registration, which is

applied to images subsampled by a factor of  $2^{n-1}$ . When necessary, additional steps can be introduced with increasingly small neighborhoods and decreasingly subsampled images. This way, by using a multiscale pyramidal algorithm, the maximum is approximated step by step approaching it using low-resolution, low cost calculations and fine-tuning it using high-resolution, high cost calculations. Since the subsampling factor decreases in every level (i.e. step), lower levels involve more image information, providing more accurate results. A complete description of one level process is illustrated in Fig. 1.

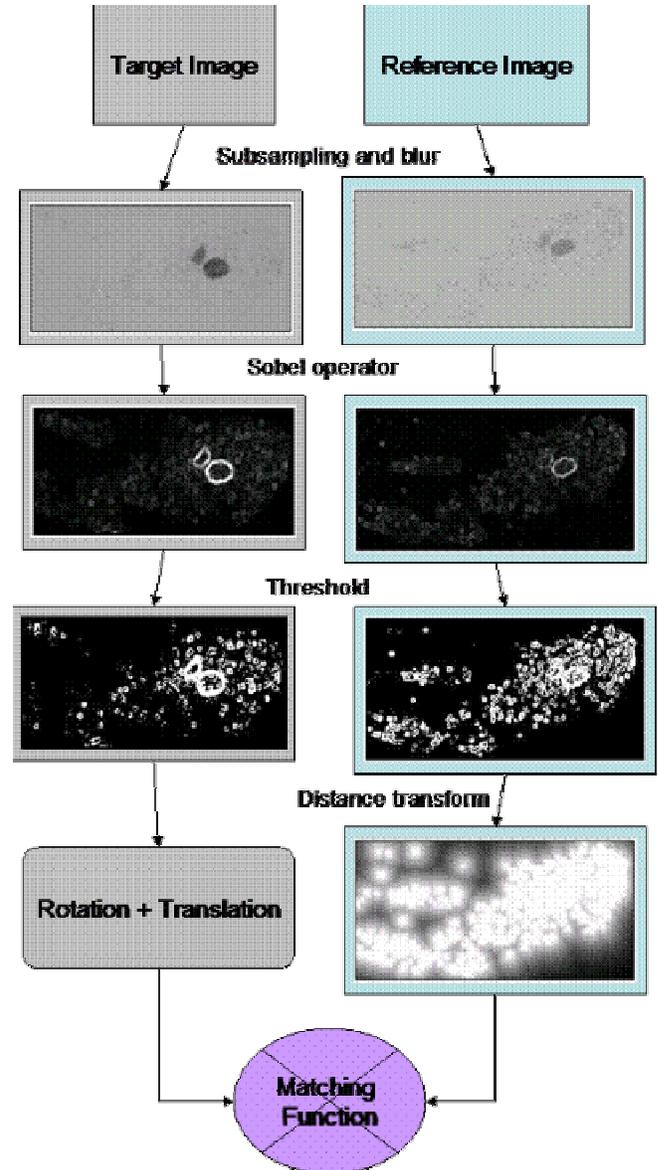


Fig. 1: Flow chart describing the registration process. The three first steps (subsampling with gaussian blur, gradient and threshold) are identical for both the reference and the image being registered (target image). Then the target image is rigid body transformed and the reference image is converted into a distance image. The same described process is applied several times after modifying –decreasing- the sampling factor of the image and the ranges of rotations and translations used in the affine transformation.

The second registration method proposed in this paper is a natural extension of the method already described. We take advantage of an automatic segmentation tool developed in our lab [4], to present a new algorithm built from the segmentation results (Fig. 2). The segmentation produces a list of contours that can be used to create a new binary image (black –0– background and white –255– contours), which is then used as the binary version of the reference image used in the original algorithm. This produces more accurate results, since the contours obtained used in our segmentation algorithm are precisely the relevant contours of the images.

Notice here that the frequent user of the segmentation tool is a mammary gland biologist, who will select the structures more relevant for his/her work, an experience that does not exist in the previous automatic registration method. Another advantage of this second method (let us call it shape registration) is the reduction in the complexity of the image: only the two contour images need to be created along with the distance transform of the contour image corresponding to the reference image. Fig. 3 shows graphically the algorithm. Another advantage of the shape registration is its speed, which is higher than the one on the standard method. Since the binary contour images present long areas (more than 95% of the image surface) with black –0– pixels, most of the rigid transformations in (1) do not need to be calculated, producing an important reduction in the computational time.

### III. RESULTS

Once the correct operation of the system was demonstrated experimentally for artificial cases, that is, cases made of one section artificially rotated and translated, the method was applied to real mammary gland sections. One way of displaying the accuracy of the registration algorithm consists of creating a color image combining both target and reference images, representing the original reference image in red, and the target image, after being transformed according to the parameters obtained, in green. Overlapped structures should then be yellow, corresponding to equal aggregates of green and red intensities. Therefore, a perfectly aligned pair of images will present most of the areas in yellow and an incorrectly aligned pair will show

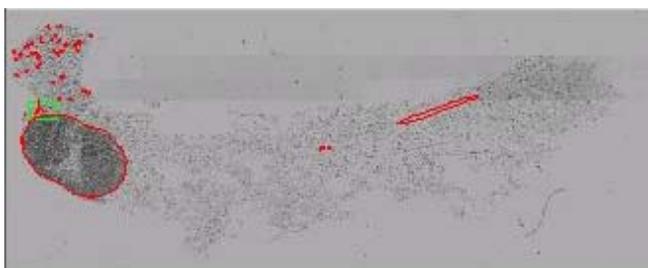


Fig. 2: Example of mice breast mammary gland section segmented automatically by our lab tool [3]. The contours in red represent the limits of relevant shapes, such as ducts, tumors or lymph nodes.

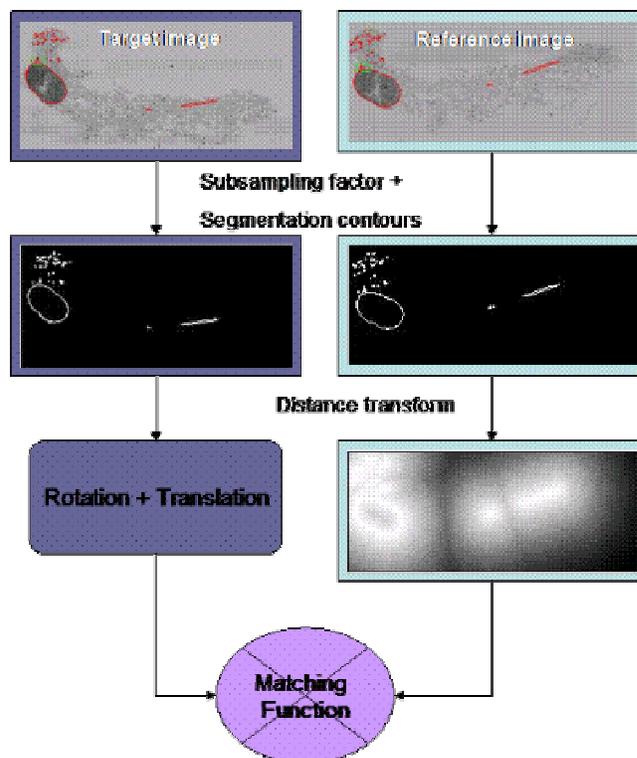


Fig. 3: Flow chart describing one level of the registration process for the “shape registration” method. Extracting the segmentation information from the original images and rescaling the contours depending on the subsampling factor, from which the binary images are built. As before, the reference is transformed in a distance image and the target is rotated and translated before applying the matching function. For this example, same image as used in Fig. 2 was processed as reference image.

areas with intermediate colors (green towards red). Fig. 4 shows the difference in accuracy between two consecutive levels of the algorithm, applied to two consecutive sections of a mouse mammary gland.

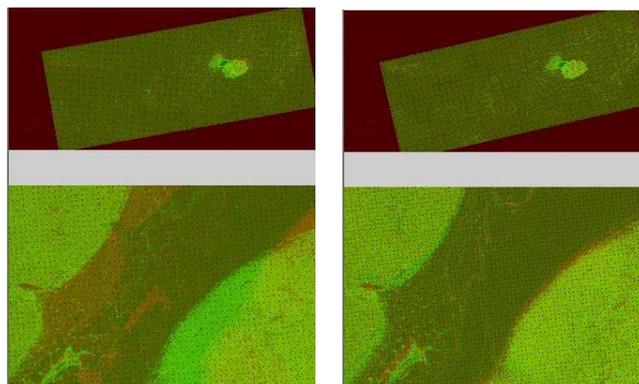


Fig. 4: Complete result and detailed zoom of the first and second level (respectively) in the standard automatic registration algorithm. Yellow areas mean structure overlapping and red and green areas represent incorrect alignment. These color pictures allow us to visualize the error reduction evolution.

Running the application on a PC Pentium IV (2.66 GHz, with 512 Megabytes of RAM memory) under Linux, the standard method of automatic registration on two typical sections (around 35 Megabytes each) takes between 4 and 5 minutes using 3 different levels of resolution. The shape registration is faster than the standard method taking approximately 2 minutes for the same pair of sections. Usually two pyramid levels are sufficient to obtain a proper registration. For more complex images, three or more levels are recommended.

One way of numerically representing the accuracy of the registration result consists of normalizing the final value of the matching function. Let  $m(d, x, y)$  be the final matching value of the images being registered (where  $d$ ,  $x$  and  $y$  represent the final rotation angle and translation in x- and y-axis), and let  $w_c$  be the number of white pixels in the thresholded contour image, then the normalized value is given by:

$$\overline{m(d, x, y)} = \frac{m(d, x, y)}{w_c \times 255} \quad (2)$$

This number, always between 0 and 1, is a pseudo-percentage of image overlap and allows quantifying the evolution of the error in every level of the algorithm, and also comparing both methods.

TABLE I  
SHAPE AND STANDARD AUTOMATIC REGISTRATION COMPARISON

	Automatic Registration Method	Normalized Matching Function Value		
		Level 1	Level 2	Level 3
Fig. 1	Standard	0.92494629	0.94978216	0.949903
	Shape	0.94635782	0.98543804	0.987002
Fig. 2	Standard	0.96788139	0.98384227	0.984542
	Shape	0.97498177	0.99652354	0.997498

Table I shows the results obtained with both registration methods on the sections used for fig. 1 and 2. We can observe that shape registration provides more accurate results than the standard method, and that the differences between level 2 and 3 values are usually quite small.

In some of the mice mammary gland cases, different types of staining were alternatively applied to the consecutive sections, increasing the complexity of the problem due to the different gradients obtained with every staining. An easy way to solve this inconvenience would have been to separately and automatically register the sections with same staining, then manually register two different stained sections, and last to propagate the correction to the rest of sections. However, a little adjustment in the blur filter (reduction in the level of blurring) allowed us to generalize our automatic method to all kind of stained sections.

## IV. DISCUSSION

As stated at the beginning of the paper, consecutive sections could present non-rigid deformations due to human interaction in the sectioning and tissue processing. In those cases, an affine transformation would be insufficient for a perfect alignment. Local corrections of the rigid registration would then provide a more accurate approximation to the problem. Regardless, the rigid registration provides a correct result in most of our cases, and can be used as the first step towards a fully non-linear elastic registration in the cases where linear registration does not provide an accurate registration.

The systematic search of the functional maximum used in our method, even if it was improved forcing a 50% image overlap between images and adjusting the number of rotations and translations to the image size and pyramid level, could be replaced by an optimization method, thus approximating the maximum faster and more efficiently.

## V. CONCLUSION

We described a fully automatic algorithm for the registration of microscopy images of consecutive tissue sections, and a variant of the same method based on the previous segmentation of the images. The system is based on a multiresolution and pyramidal approach that calculates the optimum rigid transformation between the sections. We have described the algorithm and show results using Hematoxylin and Eosin (H&E) stained pairs of sections of mouse mammary glands.

## ACKNOWLEDGMENT

I. Arganda-Carreras thanks ImageJ open source project for provided code (<http://rsb.info.nih.gov/ij/>).

## REFERENCES

- [1] Fernandez-Gonzalez R., Jones A., Garcia-Rodriguez E., Chen P.Y., Idica A., Barcellos-Hoff M.H., Ortiz-de-Solorzano C. "A System for Combined Three-Dimensional Morphological and Molecular Analysis of Thick Tissue Specimens", *Microscope Research and Technique* 59(6): 522-530, 2002.
- [2] Dean, P., Mascio, L., Ow, D., Sudar, D., Millikin, J. Proposed Standard for Image Cytometry Data Files. *Cytometry* 11:561-569; 1990.
- [3] R. Hult, *3-D reconstruction of insect ganglia*, Thesis work. Examensarbete Nr 22, Centre for Image Analysis, Uppsala, 1995.
- [4] Fernandez-Gonzalez R., T. Deschamps, Idica A., Malladi R., Ortiz de Solorzano C. *Automatic segmentation of histological structures in mammary gland tissue sections*. *Journal of Biomedical Optics* (accepted, scheduled for publication in May 2004).