Faculty of Informatics
Masaryk University Brno

Mosaicking of High-Resolution Biomedical Images

RNDr. Thesis

Mgr. Vladimír Ulman

Brno, 2005

I declare this thesis to be my original work that I have written singly. All the sources and literature that I have used I cite properly and provide full link to its source.

I would like to express thanks to my supervisor Doc. Kozubek. Great deal of thanks also belongs to MUDr. Feit for testing the program and for all fruitful notes to it.

Abstract

Image mosaicking is a process in which small images are composed together into one single image. We will focus on this topic in this work. Especially, we will focus on a computer processed image mosaicking of regular-shaped input images regularly arranged into a grid. Displayed data is from the field of tissue pathology where large 2D high-resolution color images were acquired from wide-field optical microscope. The output image is obtained by registration and stitching adjacent images resulting into a large-scale 2D high-resolution color image.

The solution will be presented through-out this work as well as particular environment where our in-house developed program is used. Novel approaches to this problem will be described too and their aspects discussed. A special measure is developed characterizing the quality of image with respect to voxel-based image registration techniques. The program makes use of special order, derived from application of graph theory, for registering adjacent images. That enabled it to handle images mainly containing empty background well. Program also performs well by using a stream stitching approach. In particular, it is feasible to compose an image of size larger than the memory capacity of processing computer. The result of stitching is claimed to contain no harming artifacts from the point of view of pathological analysis.

The work is based on two accepted articles that are included in the appendix of this thesis.

Keywords:
rigid image registration, image stitching

Contents

1 Introduction **4**
- 1.1 Preface . . . 4
- 1.2 Problem introduction and specification . . . 5
- 1.3 Overview of the thesis . . . 7

2 Registration and stitching **8**
- 2.1 Material and Methods . . . 8
- 2.2 Registration order establishment . . . 9
- 2.3 Registration methods . . . 10
  - 2.3.1 Alignment evaluation methods . . . 11
  - 2.3.2 Registration optimization . . . 14
- 2.4 Mosaic composition and stitching . . . 15

3 Results and discussion **18**
- 3.1 Registration order . . . 18
- 3.2 Comparison of alignment evaluation methods . . . 19
  - 3.2.1 Robustness comparison . . . 19
  - 3.2.2 Time consumption comparison . . . 23
- 3.3 Establishment of final global coordinates . . . 23
- 3.4 Stitching and stitch camouflaging . . . 24

4 Conclusion **27**

5 Acknowledgments and notes **29**

A Sample images **32**

B Original papers **35**

Chapter 1

Introduction

1.1 Preface

We will deal with the computer processed image mosaicking problem in this thesis. The image mosaicking is a process in which large image is somehow composed from several smaller images according to given rule or methodology. We will specify this rather vague definition in more details in the next section to achieve a good insight into the problem that was solved and that is the subject to this thesis.

For this moment, we will just allow ourselves to classify this problem into the field of computer graphics which, in turn, belongs to the most often studied fields in computer science. Moreover, the problem can be considered as an image processing problem in our case since the image was further processed after the final mosaic was created. Composed large images were further annotated, i.e. specific features of displayed data are discovered, labeled and described, and images were stored into a database afterthat.

Since this work was solved in collaboration with the Department of pathological anatomy at Brno Faculty Hospital — Bohunice, the displayed data were images of tissues typically. The task was to acquire large color image of tissue at the best possible resolution and picture quality. The adopted approach was to acquire image of specimen part by part and then compose these parts together into final large image. Details will be given in the next section. The output of the method, which is to be described later in this document, is available in the digital collection in the Hypertext Atlas of Dermatopathology at `http://www.muni.cz/atlases`. The program, a software outcome of this work, is still in frequent use there at the department.

The work has been published in two papers [7, 19] already. Each focused on just a part of the whole problem and at different level of details. Since the obvious space constraints given by publishers of papers, we decided to write this thesis in which we will concentrate on several interesting aspects of the work and describe them properly. Of course, the whole methodology will be described too. But still, we included both papers into the Appendix B (on page 35) as the sources of comprehensive summary of the method.

1.2 Problem introduction and specification

Since the advent of digital microphotography the researchers can get more from their microscope device than just photographs of specimen. The transition into the digital world removed some barriers. As an example of the most limiting barriers we can take either the size of final image (a grayscale photo) while maintaining the best resolution achievable, or the non-reproducibility and further processing of data stored analogously without losing the quality of data. The migration into digital world enabled us to overcome such barriers allowing us for on-line remote consultation with other colleagues, accessing remote database which contains reference images of some disease, easier lecturing etc.

The acquisition of a *large* image at the highest resolution and picture quality possible with respect to the given optical setup can be considered as a specific feature of digital microphotography. The area of specimen, that can be captured at the single moment, is limited by the optical instrument. We will call this a *field* — a single image that can be captured at once. Larger area of specimen can be captured by acquiring several smaller images (fields) and composing them together — mosaicking. Each field can be acquired at the limits of optical setup, namely at high magnification and resolution possible. The overall quality of large image doesn't necessarily have to be worse than the quality of every field. The composition preserves the resolution, sharpness as well as all other picture quality descriptors.

The arrangement of fields in a mosaic resembled orthogonal grid, see Figure 1.1, which was supposed to be stitched together. The mounting stage, to which the specimen was mounted, was controlled to make the adjacent fields slightly overlapping. However, the movement resolution of the stage is still lower than the resolution of the CCD chip that actually acquires, digitizes, the light information which forms the sensed image. In fact, the overlap was the only source that provided us with the information how to align adjacent images so they could truthfully constitute a large image of the whole specimen. We were expecting the mounting stage to translate perfectly. That allowed us to expect adjacent images to be just somehow translated to one another, no rotation or rescaling was expected. The process of determining the best alignment of, possibly adjacent, images is generally called the image registration.

The content of displayed images was merely tiny parts of human skin or organs, lots of cells with different textures and plenty of intracellular space. The system was not influenced by the ambient light, furthermore the acquiring instrument was located in a dark room. The specimen was a flat structure. Even that, the sharpness of adjacent fields slightly varied from time to time. The cause was often the mounting of the specimen on the stage or the specimen itself. The normal axis of the specimen plane was not exactly parallel to the optical axis resulting in different sharpness at the sides of field. So the sharpness was slightly changing over the mosaic. In addition, the non-perfectness of optics resulted in minor changes in the overlap data. For instance, if the optics or CCD chip suffered from tiny distortion on the left-hand side of the view and the mounting stage was moving then the same region was pictured once on the left-hand side and once on the right-hand side of adjacent fields, i.e. once with tiny distortion and once without that. Examples of input images are included in the Appendix.

Another registration related problem is the semantical content of a field. Specimens contained holes or the specimen's shape was not convex from time to time. Either way,

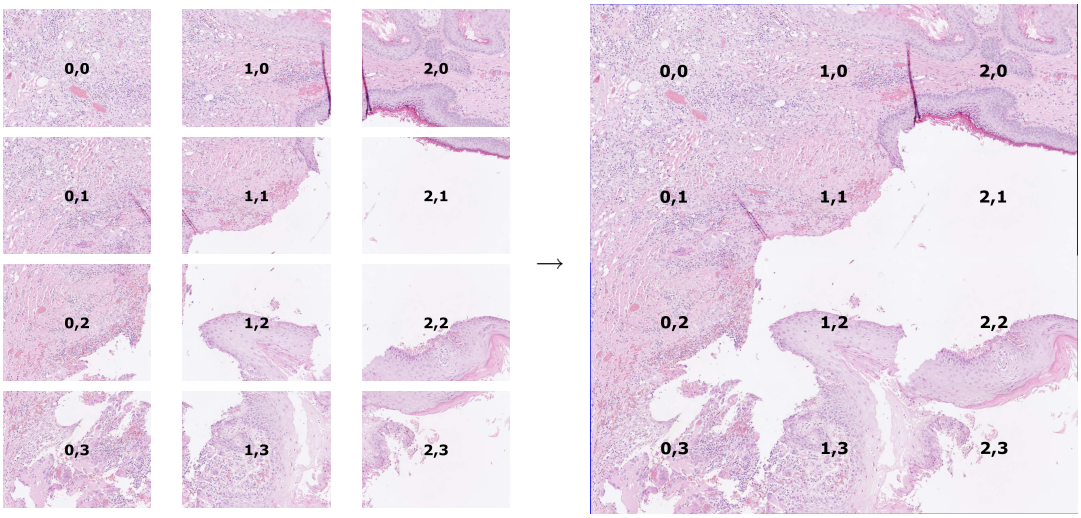


Figure 1.1: An illustration of mosaicking process and an arrangement of fields. Each image on the left-hand side from the arrow is termed as a field. Notice the overlap between adjacent fields $1,0$ and $2,0$. Also notice the nature of displayed data. Finally, note that there exist fields with background (white empty regions) in the overlap area that cause major problems with reliability of registration.

white uniform illumination formed a *background* in images. Background is those regions where no parts of specimen are present. Especially, background was often in the overlap area of adjacent images. In such situation it was quite difficult to align such images even for human observer.

Since the microscope was driven by an ordinary personal computer (PC), we planned to use that computer also for mosaicking process. Hence, three technical constraints arisen. The most limiting factors were the processor power and the amount of available memory. The processor power was important for registration establishment because registration methods can be rather demanding. The memory limitation as well as hard drive performance were critical in respect to the amount of processed data. The amount of data was usually not less than 200MB, although 1GB was nothing rare too. Thus, the design of the solution to image mosaicking had to be efficient with computer resources. In particular, the size of composed image could be considerably larger than the size of available memory capacity which was expected to be up to 512MB.

The stitching process had to deal with data that was not absolutely the same in their common overlap. Typically, if we aligned two adjacent images according to the upper part of the overlap, then the lower part was misaligned by few pixels (less than ten). No new artifacts should emerge after stitching, i.e. the shape of structures present in the overlap region must be preserved as well as tiny tissue lines should be continuous as much as possible. No data resampling was demanded in order to come over from one data to another smoothly. Decent data blurring was allowed in order to hide the stitch from human observer.

1.3 Overview of the thesis

The rest of the document is outlined as following.

In the beginning of the next chapter we will describe the entire methodology we had adopted. We will try to describe it briefly since the methodology is already described in papers included in the appendix. Furthermore, we will briefly describe the registration methods in order to provide at least some pieces of information required to understand the text in the subsequent chapter. Optimization techniques employed will be mentioned too.

The following chapter will summarize and discuss results we had achieved. We will mention the determination of special registration order, robustness of selected registration methods, optimization techniques employed, observed behavior of optical instrument and, finally, the way to compose large images with seamless stitches.

The fourth, final, chapter will conclude the thesis. This chapter is followed by appendix containing sample images (fields) and two original published papers.

Chapter 2

Registration and stitching

2.1 Material and Methods

A wide-field optical microscope Leica DMLB (Leica, Germany) with CCD device Nikon DXM 1200 (Nikon, USA) was used for fields acquisition of the specimen. We experimented, depending on the magnification desired, with several objectives ranging from 10x to 100x magnification (lens were HC PIApo 10/0.4, HC PIApo 20/0.7, HCX PIApo 40/0.85 CORR and HCX PIApo 100/1.35 Oil Imm). The instrument was driven by Lucia DI software (Laboratory Imaging, Czech Republic).

Lucia DI software was used for acquisition of every field. The system refocused before each acquisition in order to get sharp images. An automatic focusing failed from time to time — for example due to specimen mounting as described in previous chapter. The software provided a composing function for situations where even human operator driven focusing failed. The composing function could compose 2D image from several 2D images, each focused at different distance. Only sharp regions, selected from the sequence of images, were used in composition resulting in image that was sharp everywhere.

The output of Lucia software is a sequence of images that is passed to our in-house developed program. Relative positioning of two consecutive images is known due to meanderic scan used by the Lucia software. Thus, a grid position of each field is available prior to the mosaicking step (see Figure 1.1 and notice the coordinates printed inside each image).

Each image in the sequence is a 2D color image. Typical dimension was 1232×972 pixels (picture elements). The color was characterized via red, green and blue 8-bit color channels. The overlap between adjacent images was set typically from 5% to 10% of image dimension. There were usually from 30 to 1000 of images in one mosaic resulting in data through-put of 105MB to 3500MB.

We will describe the methodology of mosaicking. The method can be divided into three steps. The goal of the first and the second steps together is to establish global coordinates of top left corner of every image in the mosaic. Global coordinate unambiguously determines the final position of the field in pixel units with respect to the selected $(0, 0)$ point in the final large image coordinate system. In the first step we established a special order of fields according to which they will be processed in the second step of the method. In the second step each field is examined, the registration to adjacent images is conducted and

1. for $\forall i \in Im$ do $Pr[i] := 0$ and $Val[i] := -\infty$
2. $i' := \max_{i \in Im} \left(\sum_{j \in Sur[i]} Mes[i \leftrightarrow j] \right)$
3. while $\exists i \in Im : Pr[i] = 0$ do
4. for $\forall j \in Sur[i'] : Pr[j] = 0 \wedge Val[j] < Mes[i \leftrightarrow j]$ do $Val[j] := Mes[i \leftrightarrow j]$
5. $Pr[i'] := 1$, PRINT(i')
6. $i' := \max_{i \in Im : Pr[i]=0}(Val[i])$
7. end while

Figure 2.1: The modified Prim algorithm. The registration sequence is the order in which fields are printed in this algorithm (step (5)).

also the final global coordinates are computed. The final step stitches all fields together according to the supplied global coordinates. More details to each step will be described in the following sections.

2.2 Registration order establishment

Assessing of a special order was adopted. The purpose of the order is to register fields at first that can be probably well registered and their global coordinates can be decided well. In the end of the sequence should be images containing background in overlaps and these should be *embedded* into the mosaic in the least harming way.

We made use of graph theory for that purpose. We considered a mosaic to be an adjacency graph with fields as vertices and edges only between adjacent vertices (fields). Edges were weighted according to the measure described below. Registration order was established from modified Prim algorithm. The Prim algorithm computes usually the minimum spanning tree. We were printing out vertices in the order in which they are processed during the modified Prim algorithm. Let us denote a set Im to be the set of all fields in the given mosaic and $Sur[i]$ to be the set of all adjacent fields to the field $i \in Im$, $i \notin Sur[i]$. The algorithm is written in pseudocode in the Figure 2.1. Note the output sequence property that states: if $\forall i \in Im$ is s_i the index of i in the sequence and $s_i > 1$ then there always exist at least $j \in Sur[i]$ with the property $s_j < s_i$.

The weights associated to edges were computed from the measure given by equations (2.1) and (2.2). The measure should emphasize overlaps that provide more information with respect to registration methods,

$$Mes[i \leftrightarrow j] = Mes[i \rightarrow j] + Mes[j \rightarrow i], \tag{2.1}$$

$$\begin{aligned} Mes[i \rightarrow j] = & \sum_{(x,y) \in part_{i \rightarrow j}} |p_i(x,y) - p_i(x+1,y)| + \\ & \sum_{(x,y) \in part_{i \rightarrow j}} |p_i(x,y) - p_i(x,y+1)| \,. \end{aligned} \tag{2.2}$$

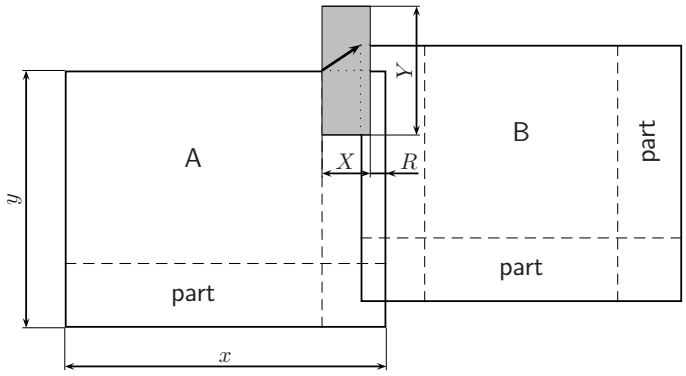


Figure 2.2: An illustration of terms demonstrated on an example of registration in horizontal direction. Notice *parts* of images as well as the area $X \times Y$ of all reasonable alignments. The shift vector is displayed as a thick arrow.

We define a $part_{i \to j}$ to be the set of all coordinates, within the coordinate system of image i, of pixels from the overlap except for the most right and the most bottom pixel lines. Value of pixel from the field i at coordinate (x, y) will be designated as $p_i(x, y)$. See Figure 2.2 for illustration of terms. Only the grayscaled (reduction from 24-bits to 8-bits per pixel) parts of images were stored in computer memory during the first two steps of the method. The measure was computed over the overlap data of both adjacent fields.

2.3 Registration methods

The information provided from the mounting stage about its movement was not accurate at the resolution of images. We used the information hidden in the overlapped area of adjacent images in order to recover exact translation between those two. We searched only for translation, no rotation, rescaling or whatsoever was expected.

We used registration methods for the optimal alignment recovery. There exists several methods for registration according to survey papers [3, 6, 5] and recently [26]. These are mainly grouped into the two categories: voxel-based and feature-based registration methods.

We excluded methods from the feature-based category due to the nature of displayed data. Extracting most of the classical image features [14, 11, 24] was expected to be less reliable for further processing, i.e. for registration.

The registration methods from the voxel-based category were used instead. The basic idea behind this methods is to evaluate all possible/reasonable alignments and select the most appropriate one. The evaluation of given alignment should therefore characterize the appropriateness, the quality of a match. The highest the evaluation was the more appropriate the alignment was. We will describe selected methods briefly. For a detailed description refer to [5, 18] or to the original sources that will be cited nearby. We had also make use of [4, 16]. The n-pass method is such method that needs for its computation to examine exactly n-times data from given overlap. All presented methods work directly on raw data.

The alignment of two adjacent fields was defined by a shift vector. Shift vectors were two dimensional (s_x, s_y) and ranged from the region all reasonable alignments, i.e. $s_x \in X \wedge s_y \in Y$ according to the notation of Figure 2.2. The $X \times Y$ region was controlled via percentage p of the field dimension, $X = p \cdot x - R$ and $Y = 2p \cdot y$. Parameter p was the input parameter to the whole method.

We also used the concept of *default vectors* for both registration directions, i.e. whether we are registering fields adjacent to one another in horizontal (as in Figure 2.2) or vertical direction. The very first registration in given direction determined the default vector for that direction. Every consecutive registration result was compared to the default vector. If the difference between default and found vectors was not significant, then the default vector was updated via reaveraging with the new vector. Otherwise, the default vector was used instead as the result of given registration. The difference was judged upon the respective differences in both coordinates. In particular, if the difference of at least one coordinate was more than 10 pixels, the found vector was compensated. This mechanism was used to detect misalignments and enabled us to embed fields when required.

Registration as well as final global coordinate determination was performed according to the sequence given in the first step of the method. Global coordinate $(0, 0)$ was set to be the top left corner of the first field in the sequence which was determined in the step (2) of algorithm in Figure 2.1. For every field in the sequence, except for the first one, the registration was computed with all adjacent fields that did not have the final global coordinate already established. Global coordinate of every field in the sequence, except for the first one, was determined as the weighted average from coordinates suggested from adjacent fields that did have the final global coordinate already established. The suggestion was based on the respective global coordinate and shift vector. The weight associated to the corresponding edge in the underlying graph representation of the mosaic was used.

2.3.1 Alignment evaluation methods

The stochastic sign change. The stochastic sign change (abbreviated to SSC) is, perhaps, the most simple and basic test for similarity. We determined the sequence of differences from corresponding pixel intensities in the common overlap given by evaluated alignment. We counted occurrences of the following situations:

- difference value change from strictly positive (above zero) to zero or less,
- difference value change from strictly negative (bellow zero) to zero or above,
- difference value remained zero.

The count is divided by the length of the sequence and that is returned as the evaluation of given alignment. SSC evaluation is simply one-pass evaluation following the idea that the more sign changes there are the more similar these two images in their common overlap are.

The last situation is for the sake of exactly the same data from both overlap area. In such circumstances the pixel intensity differences are zero obviously. That would lead to no sign changes at all while the alignment is perfect, in fact.

The normalization allows for comparison of evaluation values of alignments representing different size of overlaps. The larger the evaluated overlap is the longer the sequence is which allows for possibly more changes. On the other hand, we must restrict the size of overlap from being too small since the fraction might get higher because of the division by small number. Note the area $R \times Y$ of restricted alignments in Figure 2.2. We used $R = 5$ pixels. The range of possible evaluations is $\langle 0, 1 \rangle$.

This method presumes that the data from both adjacent images is identical except for non-correlated additive noise with the zero mean value and symmetric probability density function. Moreover, the higher the brightness difference between registered fields is, the less sign changes will occur. Sign changes will occur sparsely in such case and the SSC's expressibility will become poorer from the similarity point of view. At some level of brightness difference the SSC will completely lose the ability to point out optimal alignment. The similar behavior would happen if the noise won't be centered at zero value.

The sum of absolute valued differences. This evaluation (abbreviation is SAVD) seems to have its root in the least-square criterion which is very popular measure of similarity in computer science. We computed the overall sum of absolute values of differences of intensities of corresponding pixels. The sum is then divided by the overlap size for the same reasons as in the sign change criterion. The small overlap restriction is preserved. SAVD is defined as the MAX constant minus the computed sum — in order to be consistent since we defined that evaluation rises as the given alignment is closer to the optimal one. The best alignment is then achieved when the sum is equal to zero value, hence the evaluation is equal to MAX value. Any sub-optimal alignment will differ at some pixel pair resulting in strictly above zero value of the sum. The sum, in fact, expresses the average difference of corresponding pixel intensities and since the maximum difference of 8-bit color is 255, we set the MAX constant to the value of 255.

SAVD suggests the least-square criterion. The square function, which purpose is in fact to turn the negative values into positive ones, is substituted with the absolute value function in this case. The advantage is its less sensitivity to outliers — corresponding pixel pairs which differ notably in comparison to others. This improves the similarity evaluation when the noise is present provided the noise won't over-buzz the image itself. In other words, the noise won't affect the total sum as much as it would affect in least-square criterion leaving this way the noise-free pixels to control the total value of the sum. SAVD's range of possible evaluations is $\langle 0, \mathrm{MAX} \rangle$.

It is slightly faster than SSC and more reliable too. It is again one-pass evaluation. The images to be registered should be identical, small variance in brightness and noise is acceptable.

The normalized cross-correlation coefficient. Also known as Pearson r-coefficient sometimes also referred as linear or product-moment correlation [16], abbreviated to NCC. The basic implementation is typically two-pass evaluation [2] which can be straighten into a one-pass evaluation [18]. We implemented the NCC with the range of possible evaluations to be $\langle -100, 100 \rangle$. The optimal alignment was reached for the 100.

This evaluation measures the extent to which the intensity values of corresponding pixels are "proportional" to one another. The term proportional means linearly related. The higher the NCC is the better can be every pixel pair from overlap described with single linear equation. For instance, this implies that NCC should work when additive

noise is present with expectation equal to additive coefficient and symmetric probability density function. Multiplicative noise with expectation equal to linear coefficient should be handled too. Thus, NCC evaluation should still find the optimal alignment of images with different brightness.

The correlation ratio. We implemented the correlation ratio (CR) evaluation according to [13, 12] where it is compared to the Woods criterion and to the mutual information measure. While the NCC evaluation handles just linear dependency, the CR should handle any functional dependency [13].

The idea behind is, at first, to estimate the "dependency" of $p_i(x, y)$ data on $p_j(x', y')$ data, $j \in Sur[i]$, from the given overlap and, at second, to quantify this dependency. We would like to find the function ψ^* which satisfies the formula:

$$\psi^* = \min_{\psi} Var[p_i(x' + t_x, y' + t_y) - \psi(p_j(x', y'))], \tag{2.3}$$

where $\psi : \langle 0, 255\rangle \rightarrow \langle 0, 255\rangle$ represents the functional dependency, $(x' + t_x, y' + t_y) \in part_{i\rightarrow j}$, $(x', y') \in part_{j\rightarrow i}$ and (t_x, t_y) is examined shift vector (given by the examined alignment). Var is symbol for statistic variance (the second order central moment).

The correlation ratio can be computed in one-and-one pass. That means to scan once the entire overlap to get some statistics and then to process the statistics. The method is still rather fast. Unfortunately, the description of theoretical background as well as implementation details to this evaluation method are beyond the scope of this thesis. Instead, readers are encouraged to review [13, 18].

We implemented CR with the range of possible evaluations to be $\langle 0, 100\rangle$. The higher the evaluation is the more optimal alignment is currently examined. In case of the optimal alignment the functional dependency is fully explained by some ψ and, thus, the variance in equation (2.3) will be zero, i.e. minimum and therefore $\psi \equiv \psi^*$. The restrictions regarding the noise volume hold as for NCC. Even in spite of the fact that the CR evaluation is considered to be an extension of NCC in terms of more complex functional dependencies it can handle.

The mutual information. We also implemented the mutual information (abbreviation MI) as an alignment evaluation method. This method also deals with probability estimations. Unlike CR, which is based on functional dependency, the MI is based on statistical independency. The basic literature regarding MI and registration of large images can be considered [22] followed by papers [20, 21], technical reports [23] and notably [8]. We also implemented a variant to the method [25] originally suggested by Viola [22]. The difference was due to the probability functions estimated from joint histogram.

The idea of MI is basically like this. We can think of the image data within the overlap as of a sequence of trials. All possible trial results are numbered. Trial numbers are all possible pixel values and are stored in the set T. We can imagine U to be the discrete random variable with the probability density function P_U. Similarly, V will be discrete random variable with P_V. Both variables are represent by data within the overlap of two adjacent fields. If we can estimate the joint probability density function P_W of U and V so that P_U and P_V are the marginal probabilities, W is the discrete random vector $W = (U, V)$, then we can base the evaluation on the following two ideas. Two random variables are statistical independent if the equation 2.4 holds. Two random variables are

statistically maximum dependent if the equation 2.5 holds,

$$\forall u, v \in J : \; P_U[U = u] \cdot P_V[V = v] = P_W[U = u \wedge V = v], \tag{2.4}$$

$$\forall u, v \in J : \; P_U[U = u] = P_V[V = v] = P_W[U = u \wedge V = v]. \tag{2.5}$$

The MI evaluation method is defined for two discrete random variables according to the information theory:

$$MI(U, V) = I(U|V) = I(V|U) = H(U) + H(V) - H(U, V). \tag{2.6}$$

The $I(U|V)$ denotes the information measure from information theory where the amount of information that V has about U ($I(U|V)$ is expressed as the Shannon entropy of U ($H(U)$) subtracted by the conditional Shannon entropy of U provided V ($H(U, V)$).

We found another definition of mutual information in [5] on pages 24, 25:

$$MI(U, V) = \sum_{u,v \in J} P_W[U = u \wedge V = v] \log \frac{P_W[U = u \wedge V = v]}{P_U[U = u] \cdot P_V[V = v]}. \tag{2.7}$$

This measures the mutual information between U and V via the dependency degree using the Kullback-Leibler distance between the numerator and denominator of the fraction in the logarithm. This is in correspondence with the idea explained by equations 2.4 and 2.5. Both definitions of MI are equivalent, i.e. equation 2.6 can be converted into 2.7 and vice versa.

The general pitfall of the MI method is the estimation of joint probability density that rises the computational complexity of the evaluation to one-and-one-squared pass method. That can be understood as scanning the entire overlap in order to compute some statistics that is separately evaluated later with quadratic complexity to the length of statistics, i.e. to the color depth per pixel in the input images.

2.3.2 Registration optimization

The drawback of voxel-based registration methods is the need to explore the parameter space of expected transformation. Time consumption for deciding one registration depends on the searching of the space. That, in turn, depends on the dimensionality of the space, the domains of parameters and last, but not least, the computational demand of evaluation method (including size of processed data).

The influence of mentioned aspects can vary. However, it is always safe, while it is still reasonable, to reduce as much as possible everywhere. We narrowed the parameter space dimension by searching just for translation, i.e. search over just two parameters. We narrowed their domains into a reasonable large area where optimal alignments can be expected. We also lowered the number of examined parameters until the optimal ones are found by optimizing the search in the parameter space.

The problem of searching the parameter space is, in fact, the problem of searching the global maximum. We used two optimization techniques that we called n-step optimization technique and the gradient ascend search. We will allow ourselves in the following text to use the terminology that we are searching vectors.

The n-step optimization technique resembles the multiscale approach, or sometimes also called a pyramidal approach. The space is searched in $\lfloor \log_2(n) \rfloor + 1$ iterations. In the

first iteration, the technique examined every n-th vector and selected the best evaluated one (x_{best}, y_{best}). In every consecutive iteration, n was lowered to its half ($n := \lfloor n/2 \rfloor$), the search space was narrowed into $\langle x_{best} - n, x_{best} + n\rangle \times \langle y_{best} - n, y_{best} + n\rangle$ and it was again scanned every n-th vector for a new (x_{best}, y_{best}) vector. The iteration stoped whenever n equals to zero, i.e. after the finest search in a small region around, possibly global, maximum.

The gradient ascend search is a classical iterative optimization technique. We started the iteration with the default vector for given direction and denote it (x_{best}, y_{best}). In each iteration, better evaluation was searched around in the neighborhood of all 8 adjacent vectors (adjacent vectors were differing in some of its coordinate by at most one pixel from the given vector (x_{best}, y_{best})). If no better vector were around, the iteration would stop. It is guaranteed that the last (x_{best}, y_{best}) vector described the best alignment around, possibly the optimal one.

We optimized the very first registration in both directions with the n-step optimization technique because it searched all space of allowed shift vectors. The first registration in given direction defined the default shift vector for that direction. Every consecutive registration in given direction was then optimized by gradient ascend technique.

2.4 Mosaic composition and stitching

The final step of this method was a stream stitching — the third step. At this stage, our developed software already knew final global coordinate of every field in the mosaic. All grayscale (8-bit per pixel) parts were released from memory.

The stream stitching worked in two directions. It could assemble a part of final large image from the fields present in the given grid row. This is called the horizontal direction stitching, see Figure 2.3. Whereas, the vertical direction stitching worked with two already assembled grid rows and connected (stitched) them together with the least memory requirements as possible. That means to store as much pixel lines of the first assembled grid row as possible into a hard drive while the currently stored line is still not interfering with the second assembled grid row (adjacent grid rows are overlapping). The whole process was iterative starting with horizontal direction stitching of the top most grid row. In each iteration, the following grid row was assembled and connected to the upper part by vertical direction stitching. The process finished after connecting the last grid row and storing remaining pixel lines of the final large image into a hard drive.

The stitching itself was implemented as smooth passing over from one data to adjacent one. The intention is to display original data, then suddenly quickly and seamlessly come over to the adjacent original data. We tried to change the source of data continuously within the overlap area.

The passing was controlled by weights assigned to every pixel from the overlap. The output was computed for every pixel of the overlap using simply the weighted average of corresponding pixel values from both image data. Consider the horizontal direction stitching for instance and see the left picture in the Figure 2.4. Owing to the orthogonality of mosaic grid and the fact that fields cannot be arbitrary rotated, the overlapping region given by every reasonable alignment is a rectangle. In particular, every pixel line from the given rectangle (overlap, in fact) contains the same number of pixels. Let us denote the

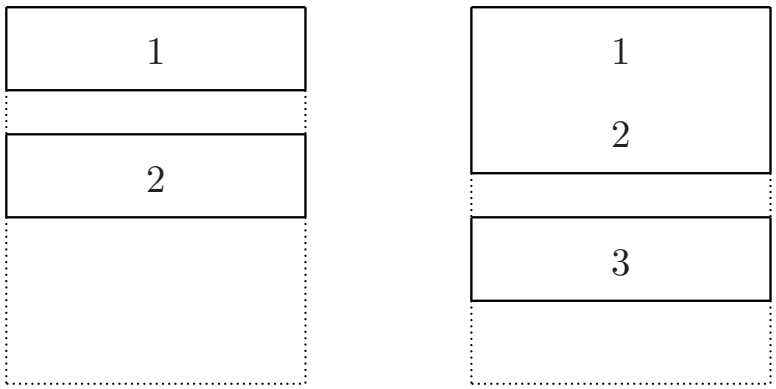


Figure 2.3: The illustration of stream stitching process. Lines 1, 2 and 3 were each separately the product of horizontal direction stitching. The stream stitching firstly connected lines 1 and 2 by the vertical direction stitching. In the next iteration it appended line 3 and so on.

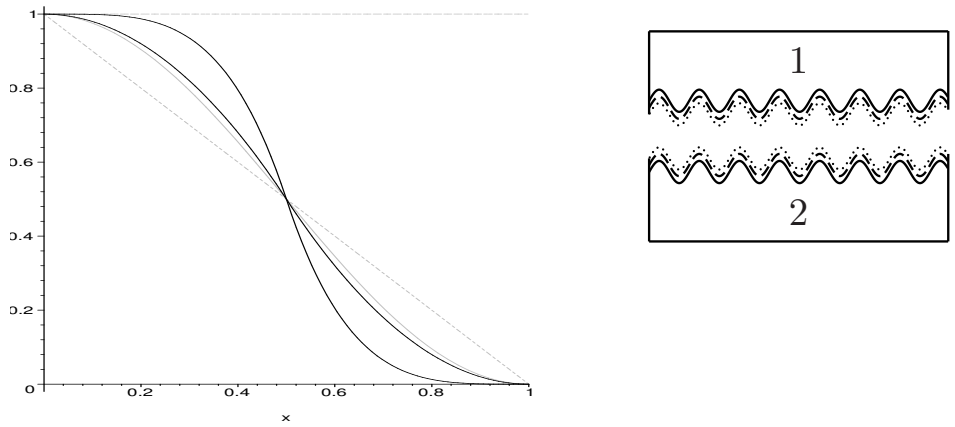


Figure 2.4: The left picture displays several weight distributions. The horizontal dashed line is the constant weight, the gray dotted is a linear weight, the gray solid is weight based on cosine function. The thin black solid curve is based on the 2nd order polynomial while the thick black solid curve is based on 4th order polynomial. For detailed expressions refer to [18]. The right picture visualizes the "zig-zag" method. It also tries to illustrate the weights distribution within the overlap. The highest weights are attached to pixels far from the edge whereas the weights get lower as the distance is shortening. The lowest weights are assigned to pixels along the dotted curve.

length of pixel line with l. The domain of weight function $w : \langle 0, 1\rangle \to \langle 0, 1\rangle$ is "stretched" to interval $\langle 0, l\rangle$. Obviously, the pixel values with coordinate $x \in \langle 0, l\rangle$ from left field are weighted by $w(x/l)$ while the corresponding pixel values from right field are weighted by $1 - w(x/l)$.

We proposed several weight functions, refer to Figure 2.4 for list of them. Notably, the weight function $w(x) = 1, x \in \langle 0, 1\rangle$ corresponds to overlapping the data originating from the left field over the data from right field. We also experimented with a "zig-zag" weight. The idea is to stitch the data in a narrow stripe which is irregularly meandering from top to bottom of overlap. We used 20 pixel wide stripe. More details can be found in [18].

Chapter 3

Results and discussion

3.1 Registration order

The shape of specimen permitted a lot of background area in the mosaic. In particular, it often happened that the background region was spread over two adjacent fields. In such cases, overlaps contained merely or solely background. The background provides very limited amount of information regarding the alignment of such fields. Thus, registering such fields seemed to be more or less a lottery that, in addition, often complicated the final stitching. For instance, misalignment of two fields shifted the global coordinates of fields in particular grid row resulting in wrong mosaic. We provided a mechanism that enabled us to overcome problems with background in overlaps.

The purpose of a value, that was attached to every edge of underlying graph representation, was to evaluate the quality of overlap with respect to registration methods implemented. Since global coordinate of each field must be determined, each field must be processed at least once. Hence, we would like to visit each vertex in the graph. Moreover, we can only visit some of those vertices that have some of their neighbors already visited. We allowed ourselves to move in the graph only along edges. Furthermore, we preferred to move first along higher valued edges in order to visit fields that are possibly easier to register, in other words, that are possibly less probable for registration failures (misalignments) to occur. The proposed modification to Prim algorithm satisfies all that requirements.

The measure $Mes[i \leftrightarrow j]$ represented major influence to the order establishment. It managed to emphasize fields with tissue in overlap to the overlaps with background. The justification of the design of equation (2.2) can be as following. Imagine two sheets of paper that have to be registered. Since the sheets are empty, or at least one of them, we are unable to determine the correct alignment. There is simply no information to do that. When we draw a dot on each sheet we gain something. We get translation parameters if we align those dots. Depending on the transformation expected we might still need more information, for instance we still cannot decide the rotation. If we draw another pair of points we might establish the rotation (or get confused...). Let's assume the new pair of dots did not denied the translation parameters. Hence, we got rotation parameter in addition. If we draw line on both sheets we again get more information even when we are expecting just the transformation and rotation, i.e. the amount of information already

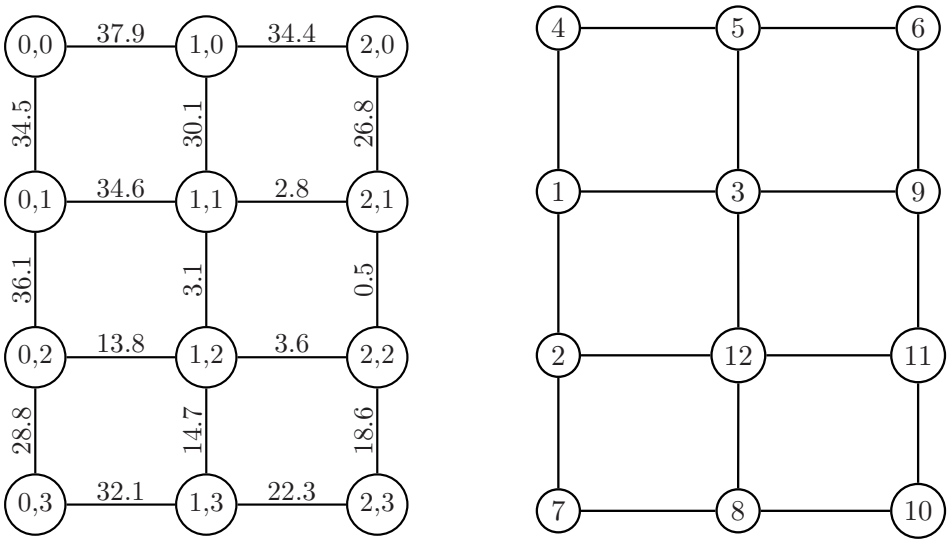


Figure 3.1: The left image displays the graph representation of illustration data from Figure 1.1. Notice the difference of weights. Low values corresponds to overlaps with background regions while, in contrast, high values corresponds to overlaps with containing solely a tissue. The associated weight is normalized value of equation (2.1). The right image displays the order that was established from given data.

extracted seems to be enough. The lines can ensure us that the parameters are correct. They gave us more confidence that observed dots were not noise etc. To sum it up, it seems that the more objects is present on the sheets, the better a human being can register these two sheets. The values given by suggested measure behaves according to this claim.

The Figure 3.1 gives an example. The illustration from introduction was used and the corresponding graph representation was created. The Figure 3.1 also displays the order established for such data. An example of somewhat larger mosaic is in the Figure 3.7 (on page 26) where the fields from the first 1/3 of the registration order are emphasized.

Even that low measure values corresponded to the overlaps where registration failed, the correct threshold for determination of the possibly worse-registrable overlaps is generally hard to guess. The texture of background may be arbitrary in general. For that reason, we belive that it is much safer to put confidence on the mounting stage property rather than on the measure itself. The mounting stage property will be described in the next section. The mounting stage property is more probable to hold and independent on background's texture.

3.2 Comparison of alignment evaluation methods

3.2.1 Robustness comparison

We made use of 3D graph in order to visualize the behavior of alignment evaluation because the evaluation behavior is central to the registration. Evaluation value is displayed for

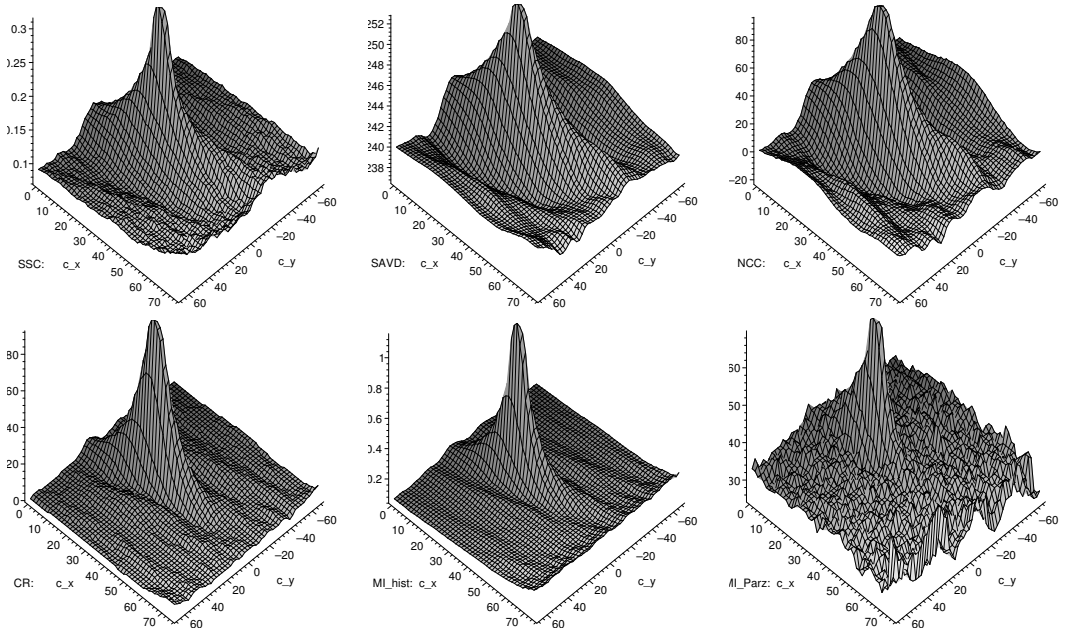


Figure 3.2: Examples of alignment evaluation method characterization. The graphs are computed for registration of horizontally adjacent fields — Figure A.1 and Figure A.2. By using notation from Figure 2.2: $(c_x, c_y) \in X \times Y, X = \langle 0, 76 \rangle, R = \langle 77, 86 \rangle$ and $Y = \langle -68, 68 \rangle$. The term MI_hist stands for MI with probabilities estimated from histogram and the term MI_Parz stands for MI with probabilities estimated using Parzen window estimator (see [22, 18]).

every reasonable shift vector (c_x, c_y). The Appendix A contains images that were used in all tests through out this section.

The selected evaluation methods are robust. Moreover, the Figure 3.2 shows that the evaluation methods are smooth, contain single peak at the shift vector representing the optimal alignment and the peak base is rather wide. In addition, the normalizations of evaluation methods seemed to improve the evaluation values correctly so that they can be directly compared regardless the size of underlying overlap or whatsoever. Note the behavior in Figure 3.2 that evaluation slightly rises when c_x approaches the R interval, notably for SSC, SAVD and MI. This is accounted for a side effect of normalizations. We observed similar graphs on all tested data borrowed from the collaborating Department of pathological anatomy.

We tested the behavior of alignment evaluation methods to several image quality degradations. Figures 3.3, 3.4 and 3.5 document the behaviors when the right field image was made more brighter, unsharped and covered with noise, respectively. The robustness may be due to the size of the overlap given by every reasonable alignment. The stripe of prohibited alignments always established reasonable amount of data present in every overlap. The evaluation method was, thus, provided with enough information to make a decision.

We also observed a feature of evaluation methods that it actually doesn't really matter for what color channel the registration is performed. We tried to register images in every color channel separately and in grayscale. We observed that there are always at least two

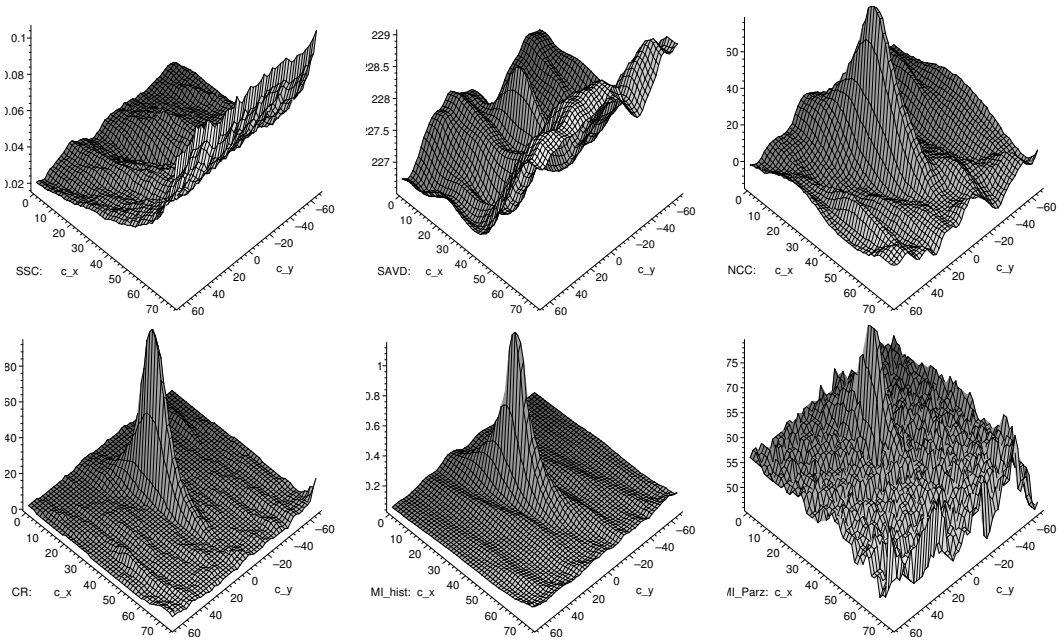


Figure 3.3: Examples of alignment evaluation method characterization. The right field image was made more brighter than the original.

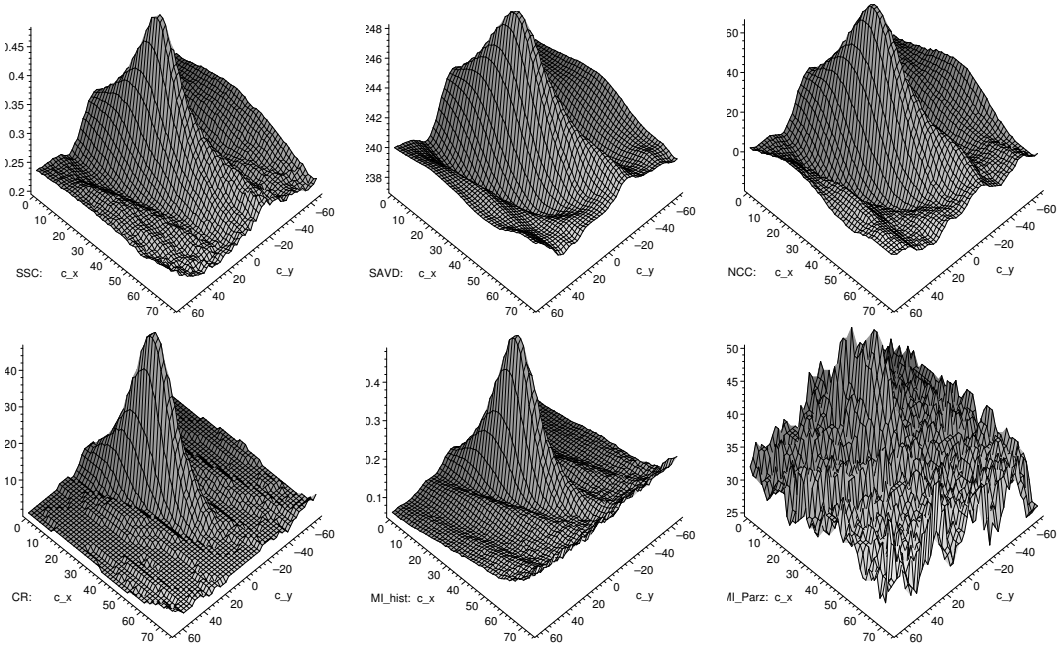


Figure 3.4: Examples of alignment evaluation method characterization. The right field image was jittered, bigger objects were retained, edges lost their sharpness.

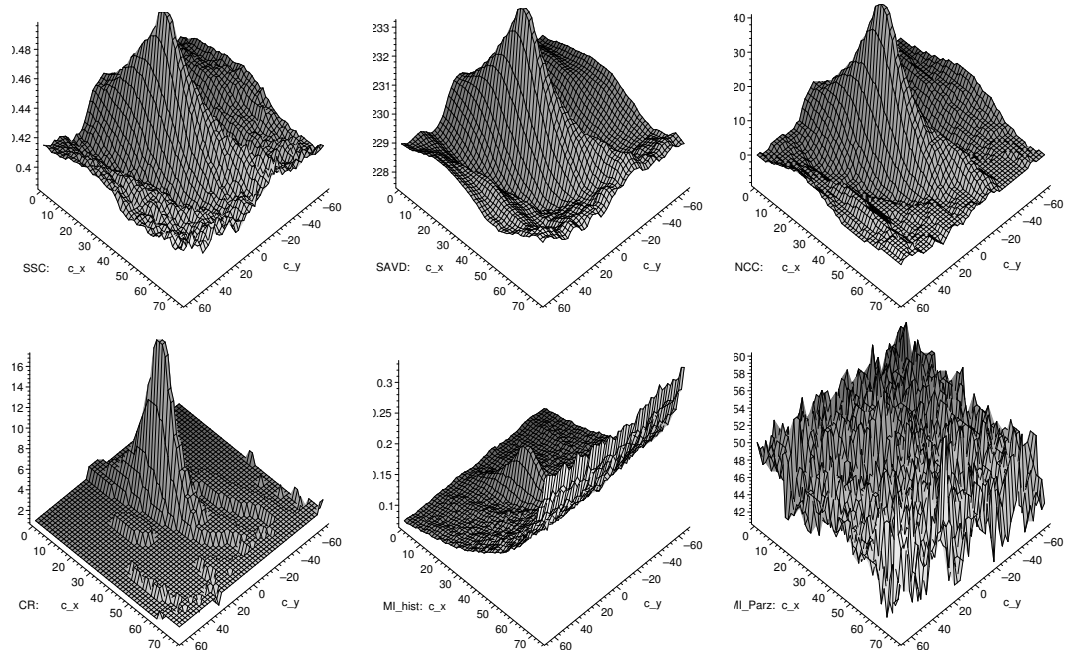


Figure 3.5: Examples of alignment evaluation method characterization. The right field image was covered with random, equally spread noise. The original structures in the image were still recognizable.

color channels with their shift vectors close to the grayscale one. The specimen was tinted into somewhat pink-to-red color before acquisition. Hence, we presume that it suppressed information present in either green or blue color channel. Nevertheless, the information aggregated into the grayscale seemed to be enough for optimal alignment determination. It allowed us to store parts of fields in just 8-bit color depth which introduced considerable memory savings.

The registration results discovered a mounting stage property. Since the controlling software drives the mounting stage to move always the same amount of units in a given direction, the shift vectors for given direction seemed to obey normal distribution characterization. The concept of default vectors then estimated the mean values for both directions. We made use of this feature and the robustness of evaluation methods for detecting the registration failures as described earlier. Drawback of this approach is in the first estimation of mean shift vector for given direction. We were successful by establishing an order, as described in the previous chapter, and providing the first estimate of mean shift vector from the first registration result.

The shape of alignment evaluation methods enabled us to use optimization techniques. The evaluations exhibited smooth changes while continuously moving through the parameter space of shift vectors. Also the shape of graphs is favourable to optimize searching parameter space. Namely, the n-step optimization technique could not be fooled by local maximum since there is one huge peak with base much broader than value of n. Typically, small values for $n \in \{4, 6, 8, 12\}$ were satisfactory while reliable, since the technique got better overview of the parameter space owing to the finer step. The continuous-like behavior of evaluations was good for gradient ascend optimization technique from the obvious

	Entire search	4-step tech.	8-step tech.	16-step tech.	Gradient ascend
SSC	55.323s	3.372s	1.027s	0.465s	0.031s
SAVD	44.746s	2.701s	0.821s	0.369s	0.036s
NCC	63.568s	3.995s	1.215s	0.546s	0.050s
CR	74.594s	4.746s	1.441s	0.648s	0.060s
MI_h	121.567s	7.835s	2.296s	0.990s	0.085s
MI_P	873.925s	56.460s	15.468s	5.384s	0.352s

Table 3.1: Speed comparison of alignment evaluation methods. Every column displays the amount of time required for one registration of typical pair of adjacent images. The column "Entire search" represents non-optimized registration. The gradient ascend optimization started its search from default shift vector established over the whole mosaic from which given pair was extracted. The default shift vector was not equal to the highest evaluated shift vector. Experimental times were measured on 1.5GHz processor (3014.65 bogomips).

reason.

3.2.2 Time consumption comparison

Alignment evaluation methods can be also compared according to the time consumption. Table 3.1 summarizes our results. Notice two facts. At first, the n-step optimization techniques introduces a great speedup even for small values of n. At second, the speedup, in general, is in the order of magnitude regarless the evaluation method. Moreover, our implementation of mosaicking program spends considerable more time when accessing the data stored on the peripheral than computing the final global coordinates (with optimized registration computation).

3.3 Establishment of final global coordinates

Final global coordinate of every field was determined from some of its adjacent fields according to the algorithm described in the previous chapter. The contribution of each adjacent field was weighted relevant to the measure $Mes[i \leftrightarrow j]$. That naturally controlled the influence. Whenever the global coordinate was supplied from adjacent field to which the measure is poor, the influence to the final global coordinate of given field will be lower since there is higher risk of misalignments resulting in wrongly estimated final global coordinate.

This framework proved to establish reliable final global coordinates. We can support this claim from inspection of several large mosaics. The suggested global coordinates to selected field from its adjacent fields did not usually differ in more than 3 pixels in x or y coordinate in the most cases.

3.4 Stitching and stitch camouflaging

Owing to the memory limitation constraint some kind of stream data processing had to be implemented. We were also looking for simple and fast algorithm. The input data typically suffered from almost negligible translation in some area of the overlap region. We disliked the idea of input data resampling in order to match another data to which it is being stitched. We felt that interpolation errors would result in the same blurring effect as in the adopted approach. Hence, we implemented stitching by weighted average without testing any other possible solution.

The perception of stitch by human eye (brain) must be well understood in order to tune the weight function. The weight function is the most important parameter of selected approach. We tested several weight functions and several sizes of overlap. From the knowledge we learnt, we can conclude following statements. Some are demonstrated in the Figure 3.6. When images are not identical, changes will occur and will be noticeable even in spite of the fact that stitched overlap contains pixel intensities derived from intensities originating from both overlaps. The weight smoothness aims against the rapid changes. We presume that the weight should slowly change its values so the eye can get used to the changes. Once the eye (human brain, respectively) gets used to changes in displayed image data, changes then became less noticeable for the human being. Nevertheless, since the real data lacks the optimal alignment, the best suboptimal alignment must be used. Consequently, in the larger region, that displays wide stitch due to slowly changing weight function, will accumulate more artifacts. For instance, the blurring of data which is caused by roughly the same weights used in averaging, Figure 3.6. The change should be, therefore, reasonably quicker. Weights should change more rapidly, resembling continuous smooth step-like function.

The slope of the weight can be regulated by the size of overlap (that is an input parameter to Lucia DI software). We report hardly noticeable stitches when using slowly changing weight function, based on the 2nd order polynomial for instance, stretched to length of around 40 pixels. Alignment evaluation methods proved to be relatively independent on the overlap size which allows for controlling the quality of stitches by regulating the size of overlap. It seems, from our observations, that reasonable narrow overlap provides us with nice stitches. On the other hand, larger overlap puts more confidence on the registration and gives more space to maneuver when stitching.

We gained even better results after employing the "zig-zag" stitching method. It is a logical conclusion drawn from the preceding paragraphs. Successful setting was a 20 pixel wide stripe together with the weight function based on the 2nd order polynomial. The stitching was further enhanced by letting the narrow stripe continuously, smoothly and *randomly* meander along direction perpendicular to the stitching direction. The idea behind meandering is to move the stitching area somewhere within the overlap without giving any piece of information where. For instance, a human being wouldn't normally notice the stitch in some certain row if there weren't similar row with more noticeable stitch.

If the data from both overlaps are perfectly identical, i.e. each corresponding pixel pair holds exactly the same intensity value, then the selection of weight is unimportant as long as the sum of corresponding weights equals to 1. The weights preserve the overall brightness in the region. The suggested stitching method allows for arbitrary change of

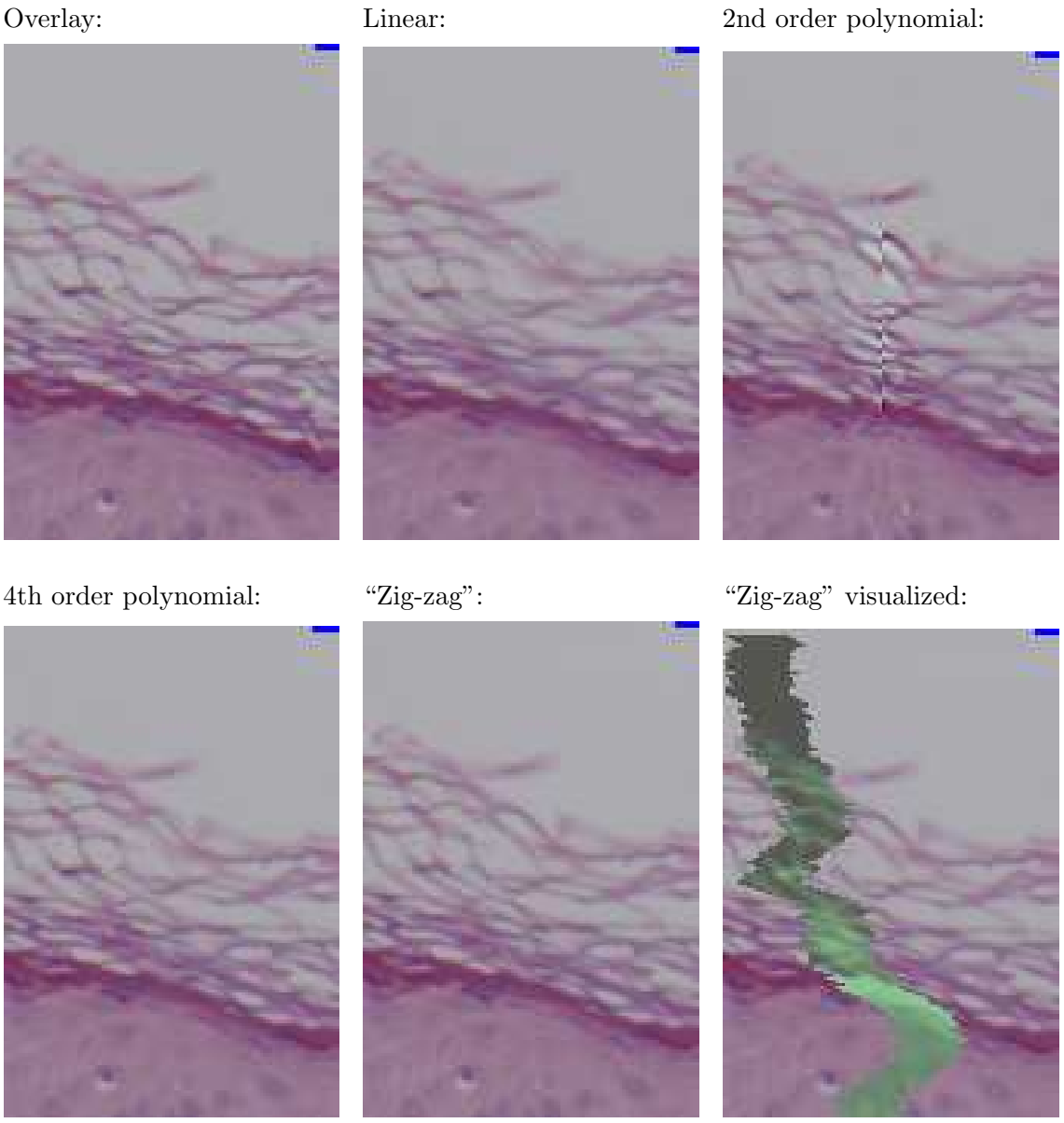


Figure 3.6: The figure demonstrates application of six different weights on the same overlap data and the same given alignment. The class of weight function used is described in the text above each picture. The "Zig-zag" visualized is the ordinary "zig-zag" method with the difference that each pixel value from the area of narrow meandering stripe was inverted, i.e. $p \to 255 - p$. A sample pass of "zig-zag" method is visualized in this way. Each picture is a rectangle constantly drawn from given horizontal direction stitch. Notice the different artifacts emerging.

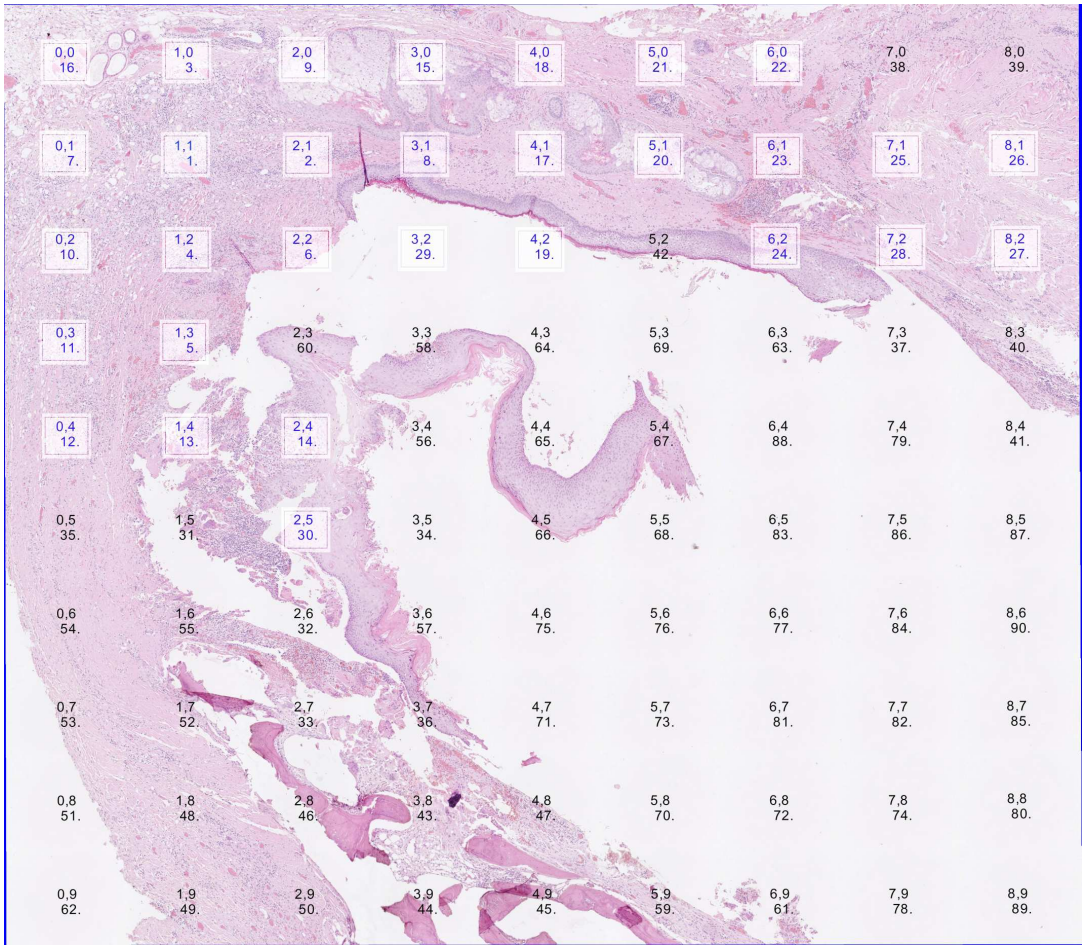


Figure 3.7: This is an example of mosaic created from 90 fields. Each field designated with its grid coordinate and the its number in the registration order. The first 30 fields in the order are emphasized by enlightened frame to demonstrate the performance of registration order.

brightness in a local region by adjusting the sum of weights at appropriate positions. This might counterbalance illumination errors of particular optical instrument. It should handle, for example, mosaicking of images acquired with brighter left-hand side than the right-hand side of images.

Final note is about the memory usage of this stitching algorithms. The mosaic from Figure 3.7 is a grid of 10 rows times 9 columns with overlap of adjacent fields to be 7%. The overall image size was thus 315MB. Before the registration order establishment all grayscaled parts were loaded into memory. This represented the first memory usage peak. The memory usage was 2.8MB at that moment. The second usage peak was achieved when vertical direction stitching was conducted. The memory usage was 20.6MB at that moment. To sum it up, the maximum memory required was up to 6.6% the size of the final large image.

Chapter 4

Conclusion

This thesis was about the image mosaicking problem. We considered given input images, slightly overlapping arranged into a grid, displaying biomedical specimen — typically tissues of skin or organs. The goal was to seamlessly stitch input images into one large image. We described the method and discussed some of its aspects. We have also presented results obtained using a special in-house developed software that implemented presented method.

We presented an algorithm that can establish registration order. The novelty is in the fact that it enabled us to handle registration of images with background in the overlap. The description is accompanied by examples and some discussion.

A decent theoretical background to the topic of alignment evaluation methods was supplied. Text is properly accompanied with citations of original sources so that anyone can develop system by himself/herself with the help of this document. We have tested major algorithms for evaluating the optimality of given alignment. These algorithms are representatives of the voxel-based family, i.e. they compute various statistics directly on the raw image data instead of extracting features from images in order to discover the optimal alignment from these features. The image preprocessing can be of course performed but our results prove that it is not necessary in this case — some alignment evaluation methods handled even very noisy images. We have observed some difficulties with the mutual information matching method. Mainly with the version that uses the Parzen-window technique for estimating probabilities.

The time consumption of the entire image mosaicking problem was mainly due to the registration step. The reason is in the amount of data that must be processed during the evaluation of every alignment. We have published a table in this thesis which summarizes our results regarding the optimization of registration step. We compared among alignment evaluation methods using either full parameter space search or optimized parameter space search. The table shows encouraging information. Up to three orders of magnitude speedup can be achieved by employing the special two level optimization scheme. The scheme makes use of discovered property of mounting stage.

Finally we have found a satisfactory solution for stitching of images when these are almost identical. Our solution is fast and accurate in that it preserves the information contained in the data. That means, for example, that it preserves shapes and colors of objects. The uniqueness of the stitching method can be considered innovative in this field since it focuses on *where* and how the stitching is performed. The output of stream stitch-

ing algorithm is deemed satisfactory even from the pathological analysis point of view. However, there is always a place for improvement. The stitch can be meandered through the regions displaying low mutual difference, along the minimum error path [15]. In this thesis, it was demonstrated that the stream stitching can be efficient with memory when composing final large image. In particular, the large image doesn't have to necessarily fit into computer memory at the moment.

We used data from the Department of pathological anatomy at Brno Faculty Hospital — Bohunice. The developed software is still in use there.

The work was also presented at the international conference. The paper, accepted into the proceedings of this conference, is included in the Appendix.

Chapter 5

Acknowledgments and notes

Presented work has been partly supported by the Ministry of Education of Czech Republic (Grant No. MSM-0021622419).

The software program used free file formats processing libraries [9, 17]. For that reason we are obliged to include the following lines:

TIFF library:

Copyright (c) 1988-1997 Sam Leffler

Copyright (c) 1991-1997 Silicon Graphics, Inc.

JFIF JPEG library:

This software is copyright (C) 1991-1998, Thomas G. Lane.

All Rights Reserved except as specified in the `README` file provided with the Independent JPEG Group's software.

In order to be complete with citation of used sources, we must cite [10] for providing reference to the programming language used. Also cite [1] for reference to LaTeX program that was used for typesetting this document.

Bibliography

[1] R. Bednář. Latex. Web pages. reference manual, URL: `http://www.cstug.cz/latex/lm/frames.html`.

[2] P. Bourke. Cross correlation. Web pages, September 1996. URL: `http://astronomy.swin.edu.au/~pbourke/analysis/correlate`.

[3] L. G. Brown. A survey of image registration techniques. Technical report, Columbia University, January 1992.

[4] M. Budíková, Š. Mikoláš, and P. Osecký. *Teorie pravděpodobnosti a matematická statistika*. Masaryk University in Brno, 1998. ISBN 80-210-1832-1.

[5] M. Čapek. *Registrace snímků z konfokálního mikroskopu*. Disertační práce, Czech Technical University in Prague, 1999.

[6] Q. S. Chen. *Image Registration and its Applications in Medical Imaging*. PhD thesis, Faculty of Applied Sciences, Vrije Universiteit Brussel, 1993.

[7] J. Feit, V. Ulman, W. Kempf, and H. Jedličková. Pořizování obrazů o velmi vysokém rozlišení metodou skládání. *Česko-Slovenská patologie a soudní lékařství*, 40/49(1), 2004.

[8] S. Gilles. Description and experimentation of image matching using mutual information. Technical report, Oxford University, 1996.

[9] E. Hamilton. *JPEG File Interchange Format*, September 1992. URL: `http://www.ijg.org/files`.

[10] J. Kučera. PB161 Programování v jazyku C++. Web pages. URL: `http://www.fi.muni.cz/usr/jkucera/pb161/`.

[11] W. K. Pratt. *Digital Image Processing*. Wiley, 1991. ISBN: 0-471-37407-5.

[12] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proceedings MICCAI'98*, volume 1496 of LNCS. Springer Verlag, 1998.

[13] A. Roche, G. Malandain, X. Pennec, and N. Ayache. Multimodal image registration by maximization of the correlation ratio. Technical report, Institut National De Recherche En Informatique Et En Automatique, 1998.

[14] J. Serra. *Image Analysis and Mathematical Morphology*. Book, 1982.

[15] P. Somol and M. Haindl. Novel path search algorithm for image stitching and advanced texture tiling. In *The 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2005*, pages 155–218. University of Western Bohemia, 2005.

[16] Electronic statistics textbook. Web pages. URL: `http://www.statsoft.com/textbook/stathome.html`.

[17] *TIFF 6.0 Specification*, June 1992. URL: `http://www.libtiff.org`.

[18] V. Ulman. Image composition. Master's thesis, Faculty of Informatics, Masaryk Univerzity, 2003.

[19] V. Ulman. Mosaicking of high-resolution biomedical images acquired from wide-field optical microscope. In *EMBEC'05: Proceedings of the 3rd European Medical & Biological Engineering Conference*, 2005.

[20] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, pages 137–154, 1997.

[21] P. Viola, W. M. Wells III, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. In *Medical Image Analysis*. Oxford University Press, 1995.

[22] P. A. Viola. *Alignment by Maximization Of Mutual Information*. PhD thesis, Massachuttes Institute of Technology, 1995.

[23] P. A. Viola and W. M. Wells III. Alignment by maximization of mutual information. In *International Conference on Computer Vision 1995*, 1995.

[24] M. Šonka, V. Hlaváč, and R. Boyle. *Image Processing: Analysis and Machine Vision*. O'Reilly, 1999.

[25] M. Zaffalon and M. Hutter. Robust feature selection by mutual information distributions. Technical report, IDSIA Switzerland, June 2002.

[26] B. Zitová and J. Flusser. Image registration methods: a survey. *IVC*, 21(11):977–1000, October 2003.

Appendix A

Sample images

In this chapter we will present five images that were used in the chapter 3.2.1. See the captions for more details regarding respective images.

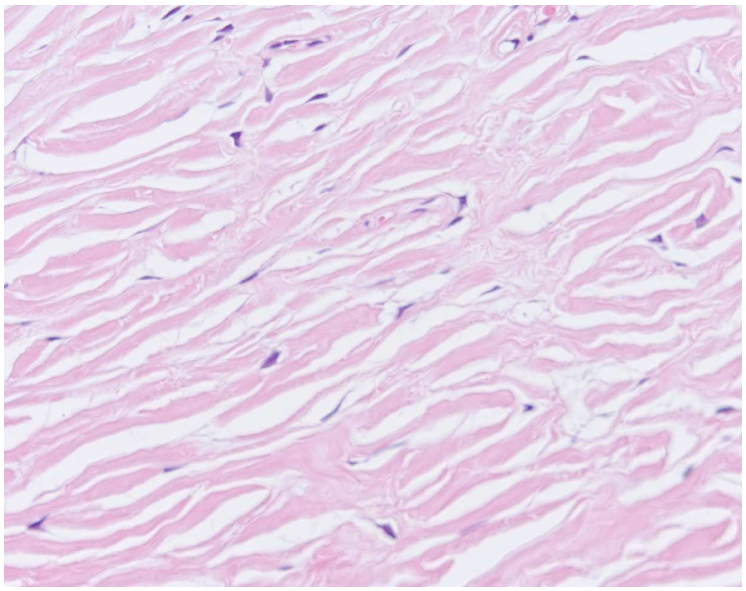


Figure A.1: This is the reference (the left) image. All tests, from the subsection 3.2.1 starting at page 19, were performed on this image data. The examined adjacent image was to the right of this one so the overlap area is on the right-hand side of this image.

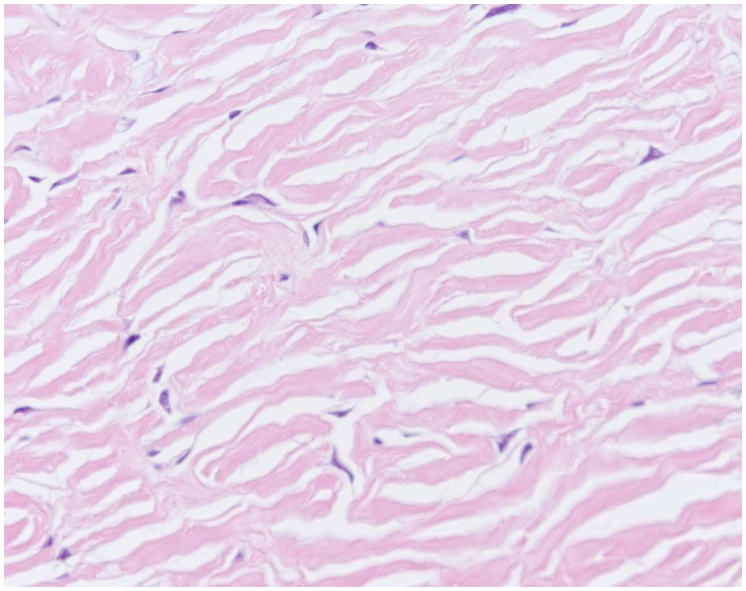


Figure A.2: This is the registered (the right) image. The overlap is, therefore, located on the left-hand side of this image. Notice, both reference and registered images are very similar not only in their overlap. Also the nature of texture as well as contrast and brightness are almost the same. This is good for voxel-based registration techniques since there exists nearly perfect match. On the other hand, this is not good for voxel-based registration techniques since there might be more reasonable alignments.

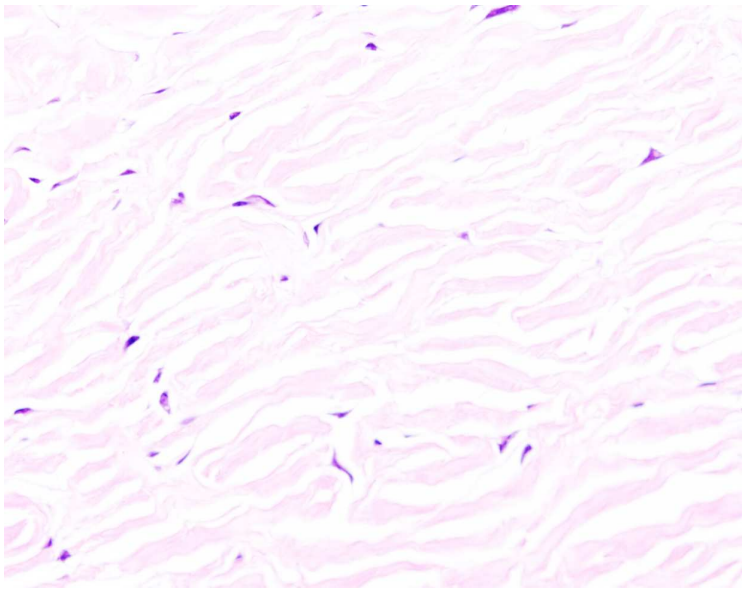


Figure A.3: This is the registered image that was adjusted to become more brighter in comparison to the original one. No artificial noise was added.

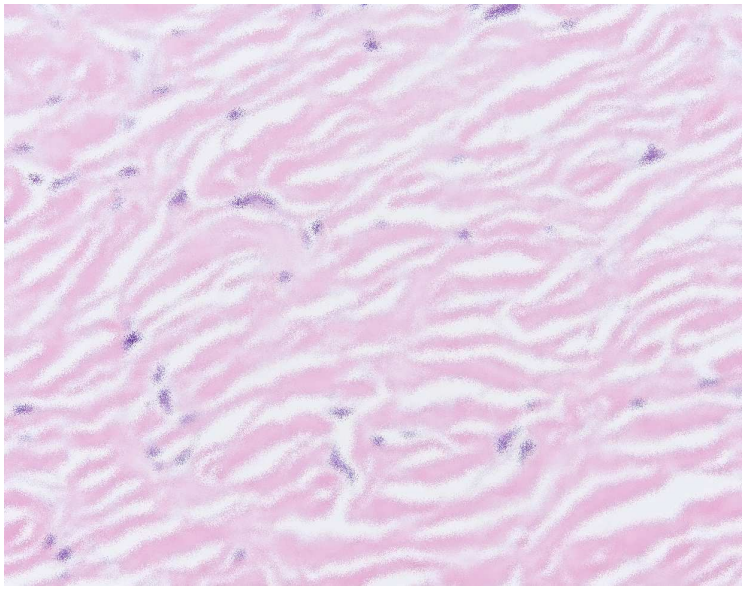


Figure A.4: This is the registered image that was adjusted to become less sharp. No artificial noise added.

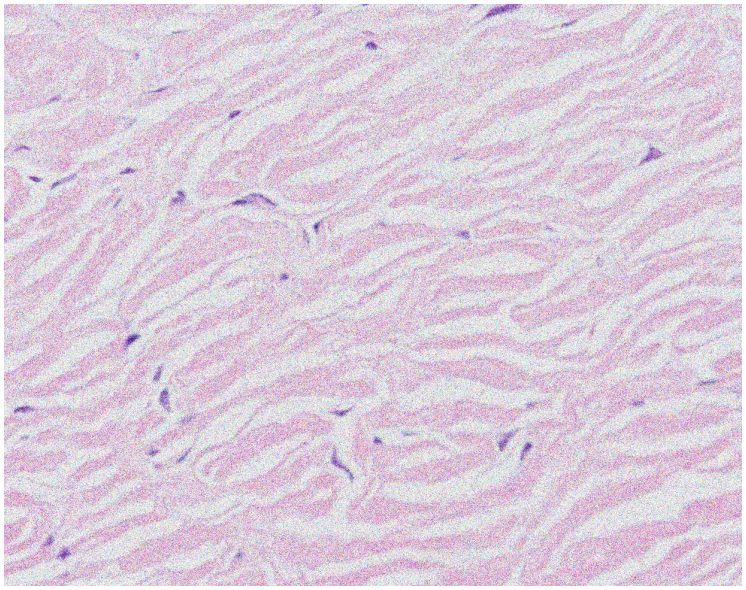


Figure A.5: This is again the registered image that is the original image to which a random noise was added. No brightness or sharpness adjustments.

Appendix B

Original papers

This chapter contains two published original papers that are reprinted at the end of this document starting from the following page.

The first paper is entitled "Pořizování obrazů o velmi vysokém rozlišení metodou skládání" and was published in the journal "Česko-Slovenská patologie a soudní lékařství" in 2004 [7]. The paper provides broader information regarding the image mosaicking problem. Not only the methodology itself is described. Also the context and reasons for doing this are mentioned as well as ideas that lead us to the decision for doing it exactly this way. It is written in Czech language.

The second paper is entitled "Mosaicking of High-Resolution Biological Images Acquired from Wide-Field Optical Microscope" and was accepted to the conference proceedings of EMBEC'05 (3rd European Medical & Biological Engineering Conference) in 2005 [19]. This paper focuses on the subject closely. It describes and discusses merely the registration and the order in which it is conducted. It also contains discussion regarding the mosaicking of images in a slightly general case.

Pořizování obrazů o velmi vysokém rozlišení metodou skládání

Feit, J.[1] Ulman, V.[2] Kempf, W.[3]
Jedličková, H.[4]

[1]Patologicko-anatomický ústav FN Brno
[2]Fakulta informatiky, Masarykova Univerzita, Brno
[3]Institut Dermatologie, Univerzitní nemocnice, Curych
[4]Dermatovenerologická klinika, FN U Sv. Anny Brno

Adresa autora:
MUDr. Josef Feit, CSc
Patologicko-anatomický ústav LF MU Brno
Jihlavská 20
662 63 Brno
e-mail: jfeit@ics.muni.cz

Souhrn
Při získávání mikroskopických obrazů o velmi vysokém rozlišení metodou skladu výsledného obrazu z jednotlivých dílů jsme narazili na některé problémy. Mezi nimi byla nutnost refokusace mezi jednotlivými dílky. S tím souvisely problémy se spojením obrazů, které si vzájemně zcela neodpovídaly a oblast spojení byla zřetelná. Byl vyvinut program překonávající některé problémy při spojování obrazových dílků, který pracuje se všemi díly naráz a hledá optimální pořadí spojení dílů. Jednotlivé dílky spojuje strmým gradientem, který probíhá po náhodně generované křivce. Program dává dobré výsledky i při spojení obrazů s pozadím či otvory ve snímané tkáni. Metoda postupného snímání a následné montáže obrazu byla využita i pro snímání sbírky kožních lymfomů ve spolupráci s Institutem pro dermatologii Univerzitní nemocnice v Curychu. Takto vzniklá digitální sbírka je veřejně k dispozici jako kapitola šesté verze Hypertextového atlasu dermatopatologie na `www.muni.cz/atlases`.

Klíčová slova: digitální mikrofotografie, počítačová analýza obrazu, interaktivní atlas, dermatopatologie, hypertext

Abstract
In order to acquire microscopic images of very high resolution by composing large images from individual parts several problems had to be solved. One of them was the necessity to adjust the focusing level when moving from one part to another. Re-focusing lead to problems with joininig the image parts, which did not correspond exactly and the area of image fusion was noticeable. A computer program was developed to overcome these problems. Our program worked with all the image parts together to find their optimal order for image fusion. Individual image parts were joined using a steep gradient running along a randomly generated curve. This method gave good results even in images with background or holes in the tissue. The method of composing large images from individual parts was used for digitizing the skin lymphoma collection of the Institute of Dermatology, University Hospital, Zurich. This collection of digital images is a part of the 6th version of Hypertext atlas of Dermatopathology at `www.muni.cz/atlases`.

Keywords: digital microphotography, image analysis, interactive atlas, dermatopathology, hypertext

V současné době se v patologii uplatňují metody digitální mikrofotografie. Jsou to metody, které znamenají kvalitativní změnu v mikrofografické dokumentaci a uplatňují se ve výuce, dokumentaci, tvorbě referenčních publikací, v telepatologii i v samotné diagnostice.

Digitální mikrofotografie umožňuje nejen okamžité získání obrazů bez závislosti na vyvolávací laboratoři, ale též získání obrazů o velkém rozlišení, snímání velkých ploch preparátu i uplatnění speciálních metod při digitálním zpracování výsledných snímků (například sesazení obrazů získaných při různých vlnových délkách při vizualizaci výsledků FISH). K tomu přistupuje i možnost sdílení a výměny obrazů po počítačové síti. Lze očekávat, že tyto metody povedou během několika let k zásadním změnám v publikování i výuce a že ovlivní i samotnou diagnostickou práci (například možností vzdálené konzultace).

Mikrofotografie (ať již na klasický film nebo digitální) má limity dané rozlišovací schopností. Tyto limity jsou dvojí: médium pro záznam obrazu (film nebo světločivý prvek kamery) a dále rozlišovací schopnost optického systému (především objektivu). Kromě toho se fotograf musí vyrovnat s problémem tloušťky snímané tkáně v kombinaci s hloubkou ostrosti objektivu. Sumace objektů ležících nad sebou nepříznivě ovlivňuje kvalitu snímku, naopak u snímků o velkém rozlišení (malé numerické apertuře a tedy malé hloubce ostrosti) chybění struktur snímané tkáně může vadit.

Již dlouhou dobu se používají techniky, které umožňují některá z těchto omezení překlenout, zejména slepování výsledných fotografií z jednotlivých dílků. Dílky byly buď snímány vedle sebe (za účelem pokrytí velkých ploch preparátu) nebo v různých fokusovacích rovinách (výsledné montáže měly velkou hloubku ostrosti a umožňovaly například sledování vláknitých struktur v silných řezech).

V současné době se začínají používat metody montáže obrazů, kde systém (mikroskop, digitální kamera, skanovací stůl a software) skládá velké obrazy z postupně snímaných obrazů běžné velikosti. Tato metoda umožňuje získat obrazy o velkém rozlišení a překonat omezení daná optikou. Tyto obrazy se dají využít ve výuce i v diagnostické praxi (2, 3, 7).

V dnešní době jsou k dispozici další metody (např. konfokální mikroskopie) a dále je možnost prohlížení výsledných snímků pomocí počítače. Zobrazovací software je schopen nabídnout funkce, které u klasických tištěných snímků nejsou k dispozici: přístup k detailu (zoom), vertikální proostřování podobně jako u mikroskopu (s výběrem vhodné fokusovací roviny nebo zpracování proostření na způsob videosekvence nebo zpracování a sumace více rovin do jednoho snímku). Obrazy o velkém rozlišení jsou k dispozici na některých www stránkách (6, 10), na (5) je možné i měnit zaostřovací roviny.

Software dále umožňuje obrazy různým způsobem anotovat, třídit a rychle vyhledávat. Lze říci, že dnes již výhody počítačového zobrazení histologických snímků svou kvalitou i komfortem předčí obrazy tištěné.

Prostředí Internetu je pro šíření obrazové informace ideální (8). V dnešní době lze říci, že Internet je dostupný všude. Rychlost přenosu a přístupnost dnes nabývají forem, které

měli autoři atlasu na mysli v době, kdy vznikal Internetový dermatopatologický atlas (1996) (4). Největšími překážkami v současné době je pro naši republiku specifická situace v cenách a kapacitách pro datový přenos, daná pozicí státem podporovaného dominantního telekomunikačního operátora.

Pro patologii má obrazová informace zásadní význam. Lze předpokládat, že v krátké době se digitální mikrofotografie zcela prosadí. V článku budou popsány naše zkušenosti se získáváním obrazů o vysokém rozlišení pro digitální dermatopatologický atlas a diskutovány některé problémy.

Materiál a metody

Starší snímky z atlasu byly získány digitální skanovací kamerou Leica S1 v rozlišení 5000×5000 px (obrazových bodů). Následně byly upraveny do výsledné velikosti maximálně 2500 px. Obrazy o vyšší velikosti byly získávány manuálním skladem obrazů v obrazovém editoru. V současné době používáme automatický systém pro snímání obrazů o vysokém rozlišení metodou skladu obrazu z jednotlivých dílků.

Snímání obrazů probíhá na systému Lucia DI (Laboratory Imaging, Praha), který řídí skanovací stůl (Märzhäuser 2D), digitální kameru (Nikon DXM 1200) a mikroskop (Leica DMLB, objektivy HC PlApo 10/0.4, HC PlApo 20/0.7, HCX PlApo 40/0.85 CORR a HCX PlApo 100/1.35 Oil Imm, pro přehled i objektivy PL Fluotar 2.5/0.07 a HC PL Fluotar 5/0.15). Preparáty jsou připraveny obvyklou cestou, nicméně pozornost je věnována dobrému napnutí řezu a standardní tloušťce krycího skla a montovacího média.

Obrazy o velkém rozlišení jsou získávány metodou snímání a následného spojení jednotlivých dílků. Systém Lucia DI řídí fokusování před každým záběrem, 3D rekonstrukci snímků a ukládání snímků na disk. Jednotlivé dílky jsou do výsledného obrazu spojeny zvlášť vyvinutým programem. Nakonec jsou obrazy digitálně zpracovány (barevná korekce, kontrast, velikost) programem Adobe Photoshop 6.0, anotovány a vloženy do atlasu. Získané obrazy obsahují jak přehlednou informaci tak i detail (viz obr. 1).

Problémy při snímání obrazů o velkém rozlišení

Snímání a sklad obrazů umožňuje použít objektivů o velkém zvětšení i pro snímání velkých ploch preparátu. Tyto objektivy mají velkou numerickou aperturu a tudíž malou hloubku ostrosti. Není možné nastavit rovinu skanovacího stolu dostatečně přesně tak, aby snímání proběhlo v jedné fokusovací rovině. Ideálně rovné není ani podložní sklo ani vlastní tkáň. Snímaná plocha při použití objektivu 40× a skladu 10×10 dílků je přibližně 4×3 mm (dílek 1200×1000 px). Přitom vertikální posun objektivu o 0.5 μm má již zásadní vliv na ostrost obrazu.

Je proto nutné, aby před každým snímaným obrazem systém zaostřil. Systém Lucia DI má možnost fokusování pomocí software (systém snímá více rovin a vyhodnocuje tu nejvhodnější na základě analýzy obrazu). Tento postup se ukázal jako dostačující.

Zásadní vliv pro kvalitu výsledného spojení má korekce obrazových dílků s ohledem na homogenitu pozadí. Systém Lucia DI umožňuje sejmutím několika zorných polí bez obrazu sestavit referenční obraz pozadí a vliv nehomogenního osvětlení odstranit (shading correction). Tato funkce se ovšem uplatní i při snímání klasických digitálních snímků.

Testovali jsme také vliv 3D rekonstrukce na kvalitu výsledného snímku. Zkoušeli jsme každý dílek obrazu sejmout v několika rovinách (typicky 3 nebo 5) a provést pro každý dílek obrazu rekonstrukci (v podstatě tak vzniká montáž ostrých částí obrazu z různých snímaných rovin). Funkce pro 3D rekonstrukci je součástí systému Lucia DI (obr. 2). Tento postup se osvědčil pro obrazy snímané menším zvětšením (objektiv 20×), pro objektiv 40× nebyl většinou nutný a při snímcích imerzním objektivem někdy selhával a někdy se objevily ve výsledném obraze artefakty (světlejší kontury kolem některých částí jader). Nevýhodou je značná časová náročnost procesu (snímání jednoho obrazového dílu ve více rovinách trvá cca 70 s, obraz 20×20 se tedy snímá řadu hodin).

Problémy při skládání obrazů

Jednotlivé části obrazů jsou snímány s mírným (5 %) překryvem. Obrazová informace z oblasti překryvu je použita pro přesné stanovení vzájemné polohy dílků obrazu před konečným spojením vzájemným prolnutím. Dokonalé spojení je teoreticky možné jen tehdy, pokud si okrajové oblasti dílků přesně odpovídají. Vzhledem k tomu, že mezi snímáním jednotlivých dílků je nutno přeostřit, oblasti se liší a dokonalé spojení není možné. Další skupina problémů souvisí s postupným narůstáním obrazu, kdy je nově nasnímaný dílek napojen na předchozí. Pokud oblast překryvu neobsahuje dostatek informace pro přesné spojení (například je snímáno pozadí), potom je možné dílek napojit jen na základě informace z kalibrace skanovacího stolu, což není zcela přesné. Pro spojení nekontrastních oblastí obrazu nebo pozadí tato nepřesnost nevadí. Pokud se však meandr snímání vrátí zpět do tkáně, nepřesná poloha dílků může bránit přesnému spojení nového dílku s dílkem předešlým a zároveň s dílky předchozí řady.

Z těchto důvodů byl sestaven program, který pracuje s již nasnímanými a na disk uloženými dílky obrazu (9). Tento program vyhodnocuje obrazovou informaci v překryvných oblastech, hledá polohu jednotlivých dílků a zároveň i pořadí, v jakém se mají dílky spojovat (obrazové hrany s dobrým kontrastem napřed, hrany s pozadím nakonec).

Spojování předem pořízených obrázků probíhá ve třech krocích. V prvním kroku jsou nejprve načteny okraje (také nazývány jako hrany) všech spojovaných obrázků, překonvertovány do šedotónní stupnice (8 bitů na pixel), uloženy do paměti a následně ohodnoceny.

Ohodnocení má vyjádřit vhodnost hrany k hledání správného přeložení obou sousedních obrázků. Jako míra vhodnosti se osvědčila suma absolutních hodnot rozdílů jasů sousedních pixelů.

Druhým krokem je nalezení tzv. vektorů posunutí (1), které definují polohu dílků ve výsledném obraze. Obrázky společně s ohodnocením jejich hran lze chápat jako graf, kde vrcholy jsou právě obrázky. Pořadí hledání vektorů posunutí je pořadí, v jakém zpracovává vrcholy Primův algoritmus hledání maximální kostry s tím, že začátek je v obrázku, který má maximální součet ohodnocení všech jeho hran (čili poskytuje nejvíce informace k připojení jeho sousedů).

Pokud nelze spolehlivě stanovit polohu daného dílku, použije se implicitní vektor získaný z dobře definovaných obrazových hran. Využívá se toho, že použitý skanovací stůl je velmi přesný (nalezené vektory posunutí u jednoho složení se většinou liší maximálně o 3 pixely v každé souřadnici). Typické použití implicitních vektorů posunutí je připojení obrázků, které obsahují v hraně spojení výhradně pozadí.

Vyhledávání vhodného vektoru posunutí je prováděno jako úplné prohledávání definičního oboru obou souřadnic; každý vektor je opět ohodnocen. Hledá se vektor, který realizuje globální extrém (většinou maximum). Ohodnocování vektorů popisuje, jak dobré by bylo napojení obou obrázků právě při tomto vektoru. Byla testována různá kriteria (11): Stochastic Sign Change, Sum of Absolute Valued Differences, Normalized Correlation Coefficient a Mutual Information.

Byla implementována tzv. n-kroková optimalizace (hledá se extrém přes každý n-tý vektor, v okolí extrému přes každý $n/2$-tý vektor, atp.). Velké urychlení je dosaženo už pro $n = 4$ (na procesorech s frekvencí okolo 1,5 GHz pak trvá jedna hrana přibližně jednu vteřinu), úspěšně odzkoušeny byly i hodnoty $n = 8$ nebo $n = 12$. Čím větší hodnota n, tím větší pravděpodobnost nenalezení správného vektoru posunutí.

Posledním krokem procesu skládání je spojení všech obrázků v jeden. Jako dobré skrytí přechodu mezi obrázky se osvědčilo prolnutí jednoho v druhý na úzkém pásu (okolo 20 pixelů), který probíhá a náhodně se přitom klikatí společnou oblastí překryvu obou obrázků (obr. 3). Obrázky se k sobě přikládají způsobem zleva doprava a shora dolů. Vždy se znovu z disku načte celý řádek matice obrázků, složí se v jeden celek a ten je celý připojen k dalšímu řádku. Proces se dobře vyrovnává s pozadím i s děrami uvnitř preparátu.

Program je psaný v jazyce C++ (především pro OS Linux) a pracuje neinteraktivně, má široké možnosti konfigurovatelnosti. Nároky na operační paměť jsou velmi malé.

Tímto způsobem je možné snímat obrazy značných velikostí (teoreticky bez omezení). Při velikosti dílku 1200×1000 px a počtu dílků až 25×20 dosahují výsledné obrazy velikosti přes 1 GB (`tiff`). Tyto obrazy jsou dále zpracovány (barevná korekce, úprava kontrastu, zaostření), zmenšeny a archivovány. Po převedení do formátu `jpeg` (JFIF) jsou popsány, anotovány a vkládány do atlasu.

Popsaný postup byl použit pro digitalizaci části sbírky kožních lymfomů Dermatologického institutu Univerzitní nemocnice v Curychu. Lymfomy jsou snímány ve zvětšení 20×, 40× i 100×. Vzhledem k tomu, že literatury o kožních lymfomech není mnoho, předpokládáme, že veřejně dostupná sbírka asi 150 obrazů doprovázená stručnou textovou informací přispěje k lepší orientaci v této problematice.

Závěr

Velkoplošné snímání histologických preparátů vyžaduje adaptivní sledování roviny fokusu. Byly popsány vlastnosti programu, který spojuje jednotlivé díly obrazu přes problémy vzniklé změnou roviny ostrosti mezi jednotlivými dílky. Metodika je uplatněna při snímání nové kapitoly o kožních lymfomech i nových obrazů v Internetovém atlasu (`www.muni.cz/atlases`).

Tato práce byla podpořena výzkumným záměrem
MZ č. 00065269705.

Literatura

[1] **Čapek, M.**: Registrace snímků z konfokálního mikroskopu. Disertační práce, České vysoké učení technické v Praze, 1999.

[2] **Costello, S., Johnston, D., Dervan, P., O'Shea, D.**: Development and evaluation of the virtual pathology slide: a new tool in telepathology. J. Med. Internet Res., 2003, 5, s. e11.

[3] **Dee, F., Lehman, J., Consoer, D., Leaven, T., Cohen, M.**: Implementation of virtual microscope slides in the annual pathobiology of cancer workshop laboratory. Hum. Pathol., 2003, 34, s. 430–436.

[4] **Feit, J., Jedličková, H., Matyska, L., Dvořák, K.**: Multimediální atlas kožní patologie. Čs. Patol., 2001, 37, s. 23 – 27.

[5] **Laboratory Imaging (LIM Praha)**: Virtual microscope, demonstration. `http://www.lim.cz/webvsl/index.php`, 2003.

[6] **Munavalli, G., Dick, F. R., Cowper, S. E.**: Virtual microscopy: teaching dermatopathology with flashpix images delivered via a streaming web-based image-server. `http://www.pathmax.com/apiii/exp.html`.

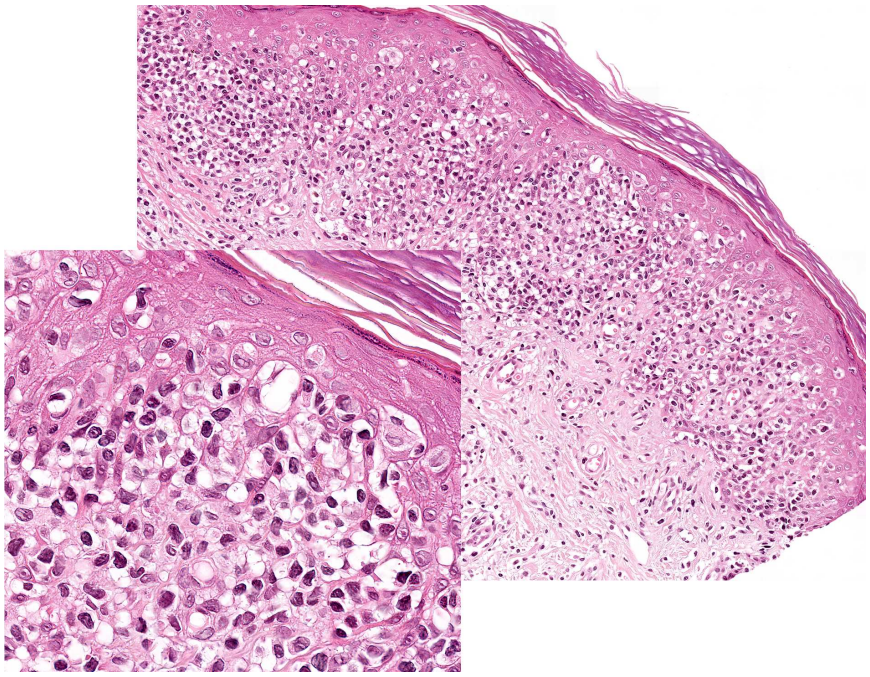
[7] **Romer, D., Yearsley, K., Ayers, L.**: Using a modified standard microscope to generate virtual slides. `http://virtualmicroscope.osu.edu`. Anat. Rec., 2003, 272B, s. 91–97.

[8] **Ryška, A., Dědič, K.**: Patologie v síti Internet. Čs. patol., 1998, 34, s. 43 – 46.

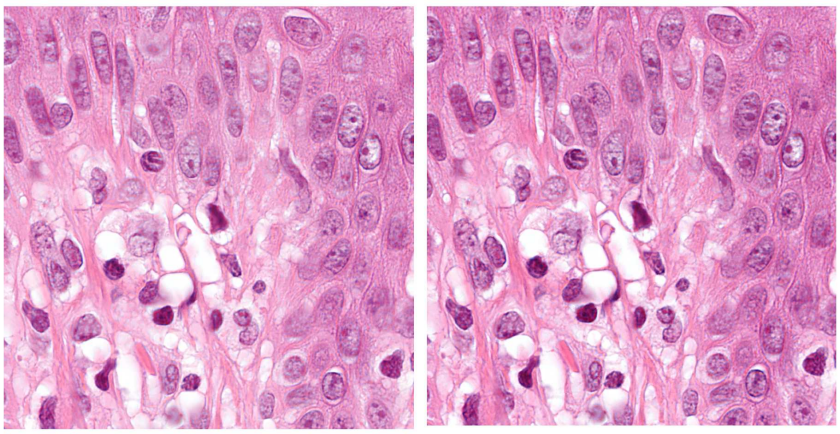
[9] **Ulman, V.**: Skládání velkých obrazů. Diplomová práce, Fakulta informatiky, Masarykova Univerzita, Brno, 2003.

[10] **University of Pittsburgh, Center for Pathology Informatics**: Virtual slide. `http://virtualslide.upmc.edu`.

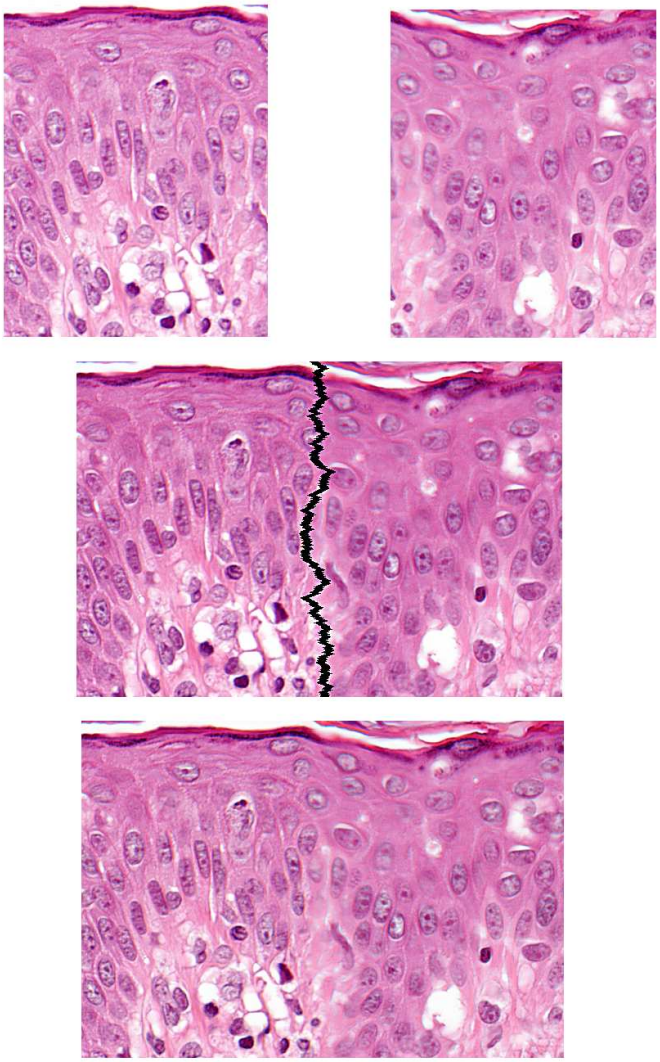
[11] **Viola, P., A.**: Alignment by maximization of mutual information. Disertační práce, Massachuttes Institute of Technology, 1995. Dokument je dostupný na URL `http://www.ai.mit.edu/people/viola/research/publications/PHD-thesis.pdf`.



Obrázek 1:



Obrázek 2:



Obrázek 3:

Legenda k obrázku 1

Celkový přehled i detail v jednom obraze (pagetoidní retikulóza, 100×).

Legenda k obrázku 2

Srovnání detailu obrazů, pořízených prostým sejmutím (100× imm.) (vlevo) a rekonstrukcí z několika rovin zaostření (vpravo). U rekonstruovaného obrazu je místy patrný detail, který u prostého obrazu byl mimo fokus.

Legenda k obrázku 3

Dva dílky obrazu před složením (40×). Dílky byly získány prostým sejmutím. Protože došlo k přeostření, překryvná oblast pravého dílku je mimo ideální rovinu fokusu. Snímek je spojen prolnutím po nepravidelně probíhající křivce.

MOSAICKING OF HIGH-RESOLUTION BIOMEDICAL IMAGES ACQUIRED FROM WIDE-FIELD OPTICAL MICROSCOPE

V. Ulman*

* Laboratory of Optical Microscopy, Masaryk University, Brno, Czech Republic

xulman@fi.muni.cz

Abstract: Large 2D high-resolution color images were acquired from wide-field optical microscope. The specimen was from the field of pathology of tissues. Each large image was obtained by stitching from a grid of smaller images. Separate acquisitions required registration and stitching of adjacent images. The novel use of special order for registration allows for easy processing of images with solely background. The order is determined from graph representation based on the grid. In this way, a reliable registration confidence test could be provided. The final large image was composed by stream stitching process because of imposed memory limitations. The stitches between adjacent images themselves were hidden by meandering technique. The methodology is described and several aspects are discussed in this paper. Our experience, gained from practical application of our system in the department of tissue pathology, supports the claim that the system is robust, fast and accurate.

Introduction

Digital microphotography becomes more and more popular in pathology of tissues. One of few enhancements we gain from transition into digital world is the possibility to acquire *large* 2D images at high resolution. Such big images can be annotated and stored into a database as reference images. High-detailed reference images can be used, for instance, when examining another data, for teaching purposes or even in tele-pathology.

We were acquiring high-resolution 2D color images from optical microscope by composing smaller images (fields) of specimen. This solution enables us to acquire every field at the limits of given optical setup, namely at high magnification and resolution possible. The fields were arranged into an orthogonal grid spreading over the entire region of interest of a given specimen.

Nevertheless, high lateral resolution of attached CCD camera is better than the resolution of the movement of mounting stage. This and the mechanical matter of stepper motors implies that the lateral movement of specimen is not described sufficiently at the resolution of fields (images). Adjacent fields were therefore acquired with small overlap providing information for correct alignment of fields.

Another issues stem from the thickness of specimen and from the almost-perpendicularity of specimen plane. The system had to refocus on every field. Due to this and the imperfection of optics, the overlaps of adjacent fields were not exactly identical and some smoothing had to be performed while stitching fields.

The registration was even more complicated because of the structure of specimen. There were fields displaying solely background due to a hole in the specimen or non-convex shape of it. Determining the correct alignment of such fields is hardly possible even for operating personal.

Last, but not least, constraint required a really large mosaicked image (e.g. more than 1GB) to be created using a computer with much less of physical memory (e.g. 0.5GB).

Materials and Methods

The specimen samples were mounted on a 2D moving stage (Märzhäuser, Germany) and acquired with CCD Nikon DXM 1200 camera (Nikon, USA), microscope Leica DMLB (Leica, Germany) with 10x to 100x objectives (lens HC PIApo 10/0.4, HC PIApo 20/0.7, HCX PIApo 40/0.85 CORR and HCX PIApo 100/1.35 Oil Imm). The system is driven by Lucia DI software (Laboratory Imaging, Czech Republic). Fields, comprising a grid, are acquired row by row, each odd row from left to right while each even row from right to left — a meandering scan.

Lucia DI software refocused at each particular field. Sometimes to get a sharp picture, a stack of images was acquired in which each image was focused at different distance. Montage from focused parts of images from 3D stack was conducted resulting in 2D sharp field.

Thus, grid coordinate of each field was known. All fields had the same dimension, typically 1232×972 pixels, and all were in 24-bit colors. Adjacent fields were acquired with overlap, typically 5–10 percent of the field dimensions.

After the acquisition, all fields were stored into separate files and ready for further three-step processing. The goal was to determine a good order for processing of fields as well as to register adjacent fields resulting in global pixel coordinates attached to every field. Coordinates were then used in the final third step where sort of stream stitching process was creating final large image.

Let us denote a set Im to be the set of all fields and $Sur[i]$, $i \in Im$, to be the set of all adjacent fields, i.e. fields that are, if they exist, to the left, top, right and bottom relative to the given field i. A set of pixel coordinates, in the coordinate system of field i, of a part corresponding to overlap between fields i and j is designated as $part_{i \to j}$.

Coordinates of the most right pixel line and the most bottom pixel line are excluded from the set for the sake of equation (2). Pixel value of field i at coordinate (x,y) will be $p_i(x,y)$.

Boundary *parts* of every field were converted into 8-bit grayscale and stored into computer memory. A *part* is actually an edge of a boundary frame of an image. It is slightly wider than the overlap that was used during acquisition of individual fields. It was expected that every part contained the real overlap. Some parts are outlined in Figure 1 by dashed lines.

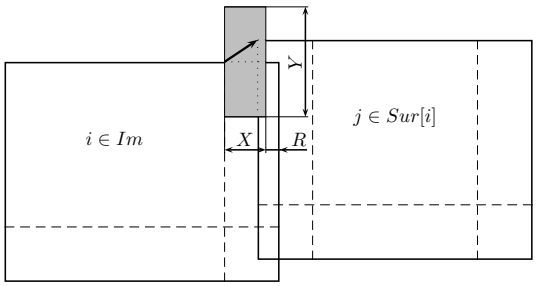


Figure 1: Parts and possible alignments

The registration order is established from weighted graph representation (Im, Mes) of the grid. Vertices of the grid are fields themselves, the set Im. Edges are just between two adjacent vertices (fields). The weight of every edge is described by $Mes[i \leftrightarrow j], \forall i \in Im, \forall j \in Sur[i]$. Modified Prim algorithm for finding maximum spanning tree is used, Figure 2. It is the order, in which fields are

(1) for $\forall i \in Im$ do $Pr[i] := 0$ and $Val[i] := -\infty$
(2) $i' := \max_{i \in Im} \left(\sum_{j \in Sur[i]} Mes[i \leftrightarrow j] \right)$
(3) while $\exists i \in Im : Pr[i] = 0$ do
(4) for $\forall j \in Sur[i'] : Pr[j] = 0 \wedge Val[j] < Mes[i \leftrightarrow j]$ do $Val[j] := Mes[i \leftrightarrow j]$
(5) $Pr[i'] := 1$, PRINT(i')
(6) $i' := \max_{i \in Im : Pr[i] = 0} (Val[i])$
(7) end while

Figure 2: Modified Prim algorithm

printed in the given algorithm (step (5)), that is used for fields registration and computation of global pixel coordinates. The weight is given by equations (1) and (2):

$$Mes[i \leftrightarrow j] = Mes[i \rightarrow j] + Mes[j \rightarrow i], \qquad (1)$$

$$Mes[i \rightarrow j] = \sum_{(x,y) \in part_{i \rightarrow j}} |p_i(x,y) - p_i(x+1,y)| + \sum_{(x,y) \in part_{i \rightarrow j}} |p_i(x,y) - p_i(x,y+1)| . \qquad (2)$$

The correct alignment of two adjacent fields was established by voxel-based registration methods [1, 2]. These methods test all reasonable alignments, the $X \times Y$ area in Figure 1, and evaluate each of them. The best alignment should have the highest evaluation. In our implementation, the search for translation of corresponding parts was just enough. The result of registration was a two-elements vector, thick arrow in Figure 1, estimating the best alignment of two adjacent fields.

Two *default* vectors were maintained, one for horizontal and one for vertical alignments. The very first registration in given direction determined the default vector for that direction. Every consecutive successful registration in the same direction improved the respective default vector. Improvement was done via re-averaging so far computed successful registration vectors in given direction. Registration was considered successful whenever the found registration vector did not differ more than 10 pixels from default vector in some of its elements. Otherwise, the found vector was ignored, default vector was supplied and no improvement was calculated.

Searching the alignment space was reduced by two optimization techniques. The very first search in given direction was improved using n-step optimization technique. First, this technique tests every n-th alignment among all from those in $X \times Y$. Then, it searches every $(n/2)$-th alignment in a $n \times n$ surroundings of the, so far, highest evaluated alignment. The last step repeats with $n := n/2$ and ends when $n = 1$. After the default vector for given direction was established, every consecutive search was optimized using gradient ascend technique. In this case, the search starts with the default vector. Neighboring vectors are examined and the highest one is selected for the next iteration. The iteration stops whenever no better alignment is around.

Global coordinates of the top left corner of every field $i \in Im$ were determined immediately when alignment vectors between i and all $j \in Sur[i]$ were computed. The global coordinate $(0,0)$ was in the top left corner of the field i' determined in the step (2). The modified Prim algorithm (Figure 2) ensures the property that whenever global coordinate of i is being computed, there exists at least one field from $Sur[i]$ that has its global coordinate already established allowing to set the global coordinate of i in this way.

Global coordinates of every top left corner were used when creating final large image. The original 24-bit color images (fields) were loaded into memory from a given grid line and stitched together according to associated global coordinates. The first two grid lines were assembled separately and stitched together. Then, as much as possible pixel lines were stored into the output image file (and removed from memory). The third grid line was assembled and stitched with the rest of the first two grid lines. The process was repeated by storing as much pixel lines as possible and proceeding with next grid line until all grid lines are processed.

A smooth transition within an overlap was utilized. Horizontal transition between fields occurred when assembling a grid line and vertical transition occurred when stitching two grid lines. Transition was implemented as

weighted sum of both intensities of corresponding pixels. The weights were controlled by two continuous functions: one was smoothly lowering influence of image data while another was raising influence of the counterpart image data.

Results

We have tested five registration methods, namely the stochastic sign change (SSC), the sum of absolute valued differences (SAVD) [1], the normalized cross-correlation coefficient (NCC) [1, 3, 4], the correlation ratio (CR) [5, 6] and mutual information [7, 8]. Two versions for estimation of underlying pixel intensity probability densities in mutual information method were tested: estimation from joint histogram (MI_h) [9] and estimation using Parzen estimator (MI_P) [10]. The time consumption of tested registration techniques is presented in Table 1 where each value (time) is for the same particular registration. The last column was measured on different registration in the same grid since the used default vector was the result from previous columns.

Table 1: Speed comparison of registration techniques

	Entire search	4-step tech.	8-step tech.	16-step tech.	Gradient ascend
SSC	55.323s	3.372s	1.027s	0.465s	0.031s
SAVD	44.746s	2.701s	0.821s	0.369s	0.036s
NCC	63.568s	3.995s	1.215s	0.546s	0.050s
CR	74.594s	4.746s	1.441s	0.648s	0.060s
MI_h	121.567s	7.835s	2.296s	0.990s	0.085s
MI_P	873.925s	56.460s	15.468s	5.384s	0.352s

The most important observations were, perhaps, the robustness of voxel-based registration methods and the movement behavior of the mounting stage. Tested registration methods proved that it is enough to search on just grayscaled data for correct alignment. This introduced big memory savings since parts could have been stored in just grayscale. Furthermore, all methods except MI_P exhibited smooth evaluation of alignments and performed equally well under normal circumstances. The smoothness is illustrated in Figure 3 where x and y axes constitute a region in a plane of evaluated translational vectors. Each vector represents unique alignment. The vertical axis describes the evaluation. Domain of tested alignments is demonstrated in Figure 1, where the registration of adjacent fields in a grid row is outlined, in the gray area. Using the notation from both figures it holds $(x,y) \in X \times Y$. The $R \times Y$ area was excluded from evaluation. In this particular example of Figure 3 the overlap was set to 7%. Thus, it was: $X = \langle 0, 76 \rangle$, $R = \langle 77, 86 \rangle$ and $Y = \langle -68, 68 \rangle$. The smoothness enabled us to make use of optimization techniques which introduced acceleration that can be seen in Table 1.

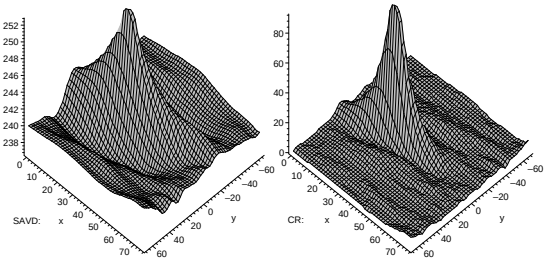


Figure 3: Alignment evaluation

The observed distribution of translational vectors in given direction resembled normal distribution. This was in agreement with expected behavior of moving stage because the controlling software always made it move by exactly the same number of distance units. Small fluctuations about mean value are due to better lateral resolution of the optical setup. The aim of default vector, for the direction under consideration, was to estimate the mean value of the stage. This reasoning enabled us to use the default vector as a starting alignment for gradient ascend optimization technique which, in fact, only refined the registration for given situation.

The robustness of registration technique and the default vectors provided the solution for detection of registration failures. It occurred, from time to time, that registration of two adjacent fields failed. Typically, the background formed more than two-thirds of overlap or there were at least two equally probable alignments — graphs, as those in Figure 3, contained more than two peaks. The deviation from expected behavior was detected using default vector as described in the previous section.

The registration order was very important because of default vectors estimation. We were successful with modified Prim algorithm, Figure 2, which builds a maximum spanning tree on the most robust edges — robust from the registration point of view. The robustness was indicated by the measure given by equations (1) and (2). The proposed measure emphasized overlaps with non-constant texture that displayed some edges (i.e. structure) which, in turn, was expected to guide the registration process. Especially, the measure was low for overlaps with background only. The ordering for highly-scattered non-convex specimen acquired using a grid of 10×9 fields is demonstrated in Figure 4. Every field is designated by its field coordinate and its number according to registration order. Fields, with its coordinate depicted in a frame, were among thirty fields that had its global coordinates established first.

A transition smoothing was performed when stitching adjacent fields. Since the overlap data were not strictly identical, a simple overlay of, say, left field over right field was not satisfactory. Notice the right-hand side of Figure 5A where the overlay is noticeable. The direction of transition is horizontal in Figure 5. Also note that the overlay can be implemented as weighted sum with one weight function constant at value 1 and another weight

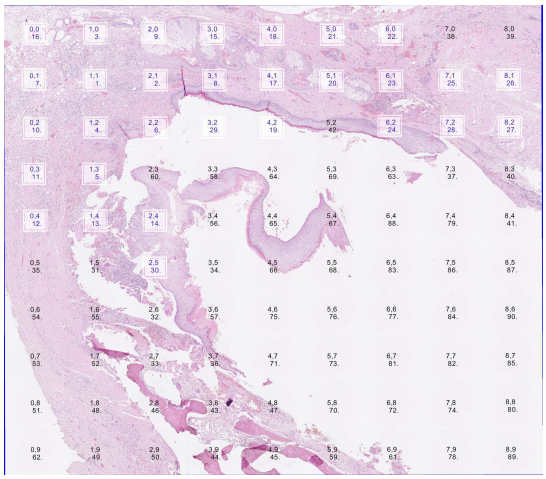


Figure 4: Mosaicked image and the registration order

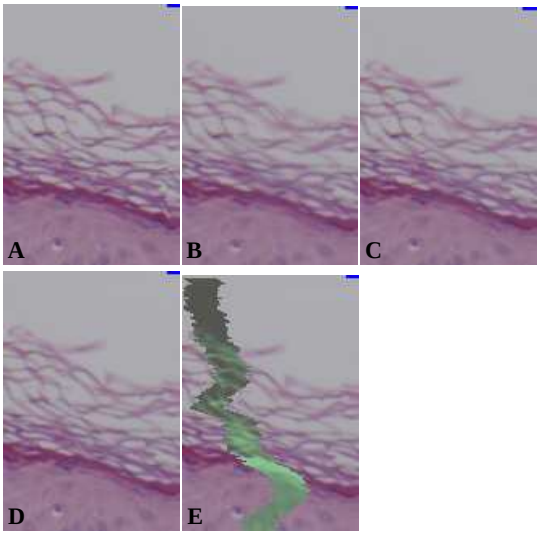


Figure 5: Demonstration of few transition techniques

function constant at value 0. A slightly better adjustment of weights represented the linear weight function ranging from 1 to 0 and from 0 to 1, respectively, over the entire overlap. Pitfall of such smooth transition is visible in the middle vertical stripe of overlap in Figure 5B. A kind of blurring is visible there due to similar values of both weight functions and due to data shift induced by non-identical overlap data. We got better outcome by using quadratic or even biquadratic weight functions (ranging again from 1 to 0 and vice versa), Figure 5C. Such weights performed more rapid and yet smooth transition making the blurred stripe narrower thus less noticeable. A "zig-zag" technique further improved the outcome of polynomial weights by narrowing the transition stripe, in our implementation to 20 pixels, and by letting it to meander along the axis perpendicular to transition direction. Result of this technique is shown in Figure 5D. Visualization of meandering transition stripe is in Figure 5E.

The stream stitching of fields was selected because of memory constraint. It also holds a pleasant property, from the implementation point of view, enabling to always stitch along the whole edge of adjacent fields or adjacent assembled grid lines. Furthermore, this property also determines the minimum memory requirement. The memory subsystem must be able to store at least two grid lines of entire input 24-bit color images (fields). Depending on image dimensions and the size of overlap, this requirement was more demanding than the requirement to store, at the moment, all 8-bit grayscale parts from all fields in a memory subsystem.

Discussion

The nature of pathological specimen, which typically displays some tissue, predetermines the voxel-based registration methods. Tissues hardly ever contain some specific features that can be extracted. Furthermore, these features should be present within overlaps of all adjacent fields. Consequently, the feature-based registration techniques are out of question.

Voxel-based techniques proved to handle images of tissues well. Unfortunately, the search through the parameter space is really time demanding. Even in spite of the fact that we search only for translational vectors (i.e. two dimensional space). In our particular experiments the evaluation of alignments behaved well which enabled us to use optimization techniques. In this way, we were able to outweight the time consumption and gain a really fast implementation.

However, in general case, we cannot be confident of the shape of the alignment evaluation of an arbitrary data without any prior analysis. An example of a shape of an evaluation is in Figure 3. Moreover, the Gaussian property of moving stage cannot be expected in advance either. In such situations we always have to search the entire registration parameter space. If the shape of evaluation is smooth then we may use the n-step optimization technique to speed up the registration process since it is quite a general optimization technique (it resembles classical pyramidal approach to search for global extrema). A huge speed up can be gained even for small n, i.e. $n = 4$, while the base diameter of peak is usually larger than basic step. Still, we may remain helpless without the Gaussian property of moving stage — for instance, when the shape of evaluation is not unimodal.

The Gaussian property of moving stage allowed us to handle fields where registration was not clear. We tested the reliability of registration by computing deviation from default vector which estimated the mean movement of moving stage for that particular direction. This solution worked well. We could have used weights from described graph representation to detect potential registration-failure fields instead. But we would still have to decide what to do with such fields, how to determine

the global coordinates. In our approach the default vectors become handy in such situations. We also believe, though we have not tested, that detection of successful registration by deviation from default vector is more powerful since it is not based directly on specific image data. In fact, the detection of failure is based on the behavior of registration.

The order establishment is crucial for setting the default translational vectors. In our implementation the responsibility is on the measure given by equations (1) and (2). The measure should represent the applicability of given overlap to image registration procedure. The higher the measure is the more information the part poses which is expected to be better for registration methods. The adjacent fields with background in their overlap are less weighted than fields with tissue in overlap since the background is expected to be more uniform. The texture of background is more solid with a few noticeable pixel intensities changes in comparison to texture of arbitrary tissue.

The aim of transition was to preserve the original information as much as possible. The stitching process left original data untouched and moved into adjacent image data as fast as possible. A small area of entire overlap was computed from both image data which typically resulted in a decent blur. We've adopted this solution because of streaming nature of final large image composition which, again, enabled us to work with whole field's edge or assembled grid line.

Conclusions

We have described a software solution for obtaining large 2D color images in high-resolution microscopy. For this purpose we developed a special program which can run very fast while still accurate as much as possible. The entire system can efficiently make use of digital microscope and an ordinary personal computer for acquiring large-scale high-resolution color images of pathological specimen.

However, the presented methodology doesn't have to work on general image mosaicking problem satisfactorily. For example, we are expecting the orthogonal grid of fields, which may pose a strong requirement in general, although it is quite natural in microscopy. The selection and processing order of techniques was focused on microscopy of tissues. The parameters of techniques were tuned for particular optical setup. We tried to discuss some aspects of our solution and suggest what to do when some of expected constraints are not met.

The system is still in use in The Faculty Hospital Brno in combination with Lucia DI software. The mosaicking process itself works fully automatic on several different kinds of tissue. The registration in combination with default vectors computes global coordinates well in respect to the stitching. The smoothed transition was not deemed harming by pathologists. The stitching process, as described, didn't produce any artifacts that could violate the analysis.

Acknowledgements

Presented work has been supported by the Ministry of Education of the Czech Republic (Grant No. MSM-0021622419).

References

[1] M. Čapek. *Registrace snímků z konfokálního mikroskopu*. Disertační práce, Czech Technical University in Prague, 1999.

[2] L. G. Brown. A survey of image registration techniques. Technical report, Columbia University, January 1992.

[3] M. Budíková, Š. Mikoláš, and P. Osecký. *Teorie pravděpodobnosti a matematická statistika*. Masaryk University in Brno, 1998. ISBN 80-210-1832-1.

[4] P. Bourke. Cross correlation. Web pages, September 1996. URL: `http://astronomy.swin.edu.au/~pbourke/analysis/correlate`.

[5] A. Roche, G. Malandain, X. Pennec, and N. Ayache. Multimodal image registration by maximization of the correlation ratio. Technical report, INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE, 1998.

[6] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proceedings MICCAI'98*, volume 1496 of LNCS. Springer Verlag, 1998.

[7] P. A. Viola. *Alignment by Maximization Of Mutual Information*. PhD thesis, Massachuttes Institute of Technology, 1995.

[8] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, pages 137–154, 1997.

[9] M. Zaffalon and M. Hutter. Robust feature selection by mutual information distributions. Technical report, IDSIA Switzerland, June 2002.

[10] S. Gilles. Description and experimentation of image matching using mutual information. Technical report, Oxford University, 1996.