Software Analysis of Mining Images for Objects Detection

Lacezar Licev [1], *Marek Babiuch* [2], *Radim Farana* [3], *Jan Tomecek* [4]

The contribution is dealing with the development of the new module of robust $FOTOM^{NG}$ system for editing images from a video or mining image from measurements for subsequent improvement of detection of required objects in the 2D image. The generated module allows create a final high-quality picture by combination of multiple images with the search objects. We can combine input data according to the parameters or based on reference frames. Correction of detected 2D objects is also part of this module. The solution is implemented into $FOTOM^{NG}$ system and finished work has been tested in appropriate frames, which were validated core functionality and usability. Tests confirmed the function of each part of the module, its accuracy and implications of integration.

Keywords: *Digital image processing, image registration, Phase correlation, Threshold, Fourier transform, $FOTOM^{NG}$, Netbeans platform, Java.*

Introduction

The aim of the work is developing of a new module for combining images from video or frames from mining measurement for later processing and the correction of 2D objects in the image.

The paper is divided into several parts. The first section describes digital image processing and its applications. There is described a method of combining images, which is used to create a high-quality picture or picture with higher resolution.

The next section describes the design of the module. It was used NetBeans platform with Netbeans IDE, a library for fast creation of applications in image processing such as Java Advanced Imaging and video library Xuggler, which you can use to encode and decode video.

Furthermore, the paper describes the implementation of the module and the results of the module testing on the mining images. The generated module was subsequently added to the latest version of $FOTOM^{NG}$ system [1].

Digital Image Processing

Definition of Image

When talking about the real picture, we're talking about an image as a multidimensional signal. The real image may be defined as a two-dimensional function $f = f(x,y)$, where x and y are coordinates in the plane and the amplitude of f is called intensity for each pair of coordinates (x,y). The image is represented discretely in the computer; in order to digitally process the image, it is necessary to digitize, which means convert an image from a continuous space into discrete, by the methods of sampling and quantization. A digital image is composed of a finite number of elements that have a specific value and location. We're talking about image points known as the pixels. [2]

Convolution

Convolution is the most basic mathematical operations in the processing and analysis of digital image. It plays an important role in edge detection and application of image filters. Convolution as a simple math operator produces from two functions the new one. Convolution of two functions $f(x,y)$ and $h(x,y)$ is marked with $*$. In the two-dimensional continuous space is the convolution defined by the relationship

$$f(x,y) * h(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(a,b) * h(x-a, y-b) \, da \, db \qquad (1)$$

[1] *Assoc. Prof. Ing. Lacezar Licev, CSc.*, VSB - Technical University of Ostrava, Czech Republic, Department of Computer Science, lacezar.licev@vsb.cz

[2] *Ing. Marek Babiuch, PhD.*, VSB - Technical University of Ostrava, Czech Republic, Department of Control Systems and Instrumentation, marek.babiuch@vsb.cz

[3] *prof. Ing. Radim Farana, CSc.*, VSB - Technical University of Ostrava, Czech Republic, Department of Control Systems and Instrumentation, radim.farana@vsb.cz

[4] *Ing. Jan Tomecek*, VSB - Technical University of Ostrava, Czech Republic, Department of Computer Science, jan.tomecek@vsb.cz

Discrete two-dimensional convolution - The points originate from finite space, if we define this space as $\Omega = \{(x,y)|x = 0,1,..,M-1; y = 0,1,..,N-1\}$ and le the signals $f(x,y)$ and $h(x,y)$ are from the considered area, then

$$f(x,y)*h(x,y) = \sum_{r=0}^{M-1}\sum_{s=0}^{N-1} f(r,s)h(x-r,y-s) \tag{2}$$

Let us consider the function h as a matrix, which is called the convolution matrix, convolution core or convolution filter. The calculation is performed so that the matrix is attached to the appropriate place in the image, and each of the overlapped pixels in the image is multiplied by the value of the corresponding location matrix, thereby we perform the sum of these values. The sum of the values gives the final value of the pixel.

For the implementation of discrete convolution is necessary to solve the boundary parts of the image matrix, where some elements of the convolution masks may be located outside of the image matrix with overlap. The solution is to expand the image matrix by auxiliary points on half the size of the convolution masks or by adding the values in positions mirrored by the image border.

Convolution masks - Design and creation of convolution masks is based on complicated mathematics and took too much time, therefore it is used primarily previous been created, in which is a well known result. Defined masks have different dimensions, that can be used to set the required filter characteristics and as 2x2, 3x3, 5x5, etc. The most famous are Median filters and Gaussian filter. There are filters used for other purposes, such as thresholding, edge detection, blur and focus.

Thresholding

The simplest method of image segmentation is the process called thresholding when the image detected by the whole area. The main idea of thresholding is to use the fact that the different objects in the image have different levels of intensity (brightness). One of the possible implementation of thresholding is to determine the threshold value of brightness so-called threshold. If the pixel brightness level is greater or equal to the threshold value, then the pixel is indicated as the object. The pixel with brightness levels lower than the threshold is marked as a background.

The hardest and most important fact for thresholding is the correct determination of the threshold value. The optimal solution is to let calculate the threshold value in the image by automatic method. These methods can be based on an analysis of the histogram image when the histogram image contains one or more dominant peaks. For images with bimodal histogram is advisable to choose a threshold minimum value between the two peaks of the histogram. For images with more modal histogram is suitable to use adaptive thresholding, so that each area contains only a single dominant peak.

Fourier Transform

The Fourier transform is used to convert from the time domain to the frequency and back. It is also used in the analysis of the frequency spectrum, when the Fourier transform is used to detect edges, adjust image quality, image segmentation, image reconstruction and image compression [3]. Picture transform is mainly used to transfer the image to a shape that is for the processing much more advantageous. Let us denote the input image f and image after Fourier transformation as F, then the

Discrete Fourier transform

$$F(k,l) = \frac{1}{\sqrt{MN}}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)e^{-j2\pi(\frac{mk}{M}+\frac{nl}{N})}, k = 0,1,\ldots,M-1, l = 0,1,\ldots,N-1. \tag{3}$$

Backward discrete Fourier transform

$$f(m,n) = \frac{1}{\sqrt{MN}}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1} F(k,l)e^{j2\pi(\frac{mk}{M}+\frac{nl}{N})}, m = 0,1,\ldots,M-1, n = 0,1,\ldots,N-1. \tag{4}$$

Convolution theorem - The Fourier transform of convolution functions in the spatial domain is the product of the Fourier images of functions in the frequency domain and vice versa product in the spatial domain corresponds to convolution in the frequency domain. Convolution is in this case considered cyclic.

$$\mathscr{F}\{f(x,y)*h(x,y)\} = \frac{1}{\sqrt{MN}}F(k,l)G(k,l) \tag{5}$$

Shift theorem - According to shift theorem can be argued that the shift of an object in the spatial domain leads to a phase shift in the frequency domain as follows

$$f(m - m_0, n - n_0) \Leftrightarrow F(k,l)e^{-j2\pi(\frac{m_0 k}{M} + \frac{n_0 l}{N})} \tag{6}$$

Fast Fourier Transform - Fourier transform itself does not achieve very good time complexity $(M^2 N^2)$ in two-dimensional space, so it has devised a new faster Fourier transform, called Fast Fourier Transform. Here N x N matrix is divided by the odd-and even-numbered positions, and this process continues until only one element remains after halving. The requirement for the use of fast Fourier transform method requires a square size of the image and dimensions must be equal to multiply of two.

Phase Correlation

Phase correlation is based on the method of fast Fourier transform, which aims to determine the size of the shift between two partially overlapping images. We can say that the phase correlation is a normalized cross-correlation[4].

If we have in R^2 defined function $f1(x,y)$ and $f2(x,y)$ and:

$$\mathscr{F}\{f_1(x,y)\} = F_1(\omega_x, \omega_y) \tag{7}$$

$$\mathscr{F}\{f_2(x,y)\} = F_2(\omega_x, \omega_y)$$

where $\mathscr{F}$ denotes Fourier transform. Then let's assume mutual displacement of functions:

$$f_2(x,y) = f_1(x + \Delta x, y + \Delta y) \tag{8}$$

Using Fourier's theorem we get the displacement relationship in the frequency domain:

$$F_2(\omega_x, \omega_y) = F_1(\omega_x, \omega_y)e^{j(\omega_x \Delta x + \omega_y \Delta y)} \tag{9}$$

Then we convert function F_1 to komplex-valued and acquired function $F_1^*(u,v)$, we multiply by equation:

$$F_2(\omega_x, \omega_y)F_1^*(\omega_x, \omega_y) = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \tag{10}$$

The symbol $*$ denotes the complex conjugate. The resulting image is called the cross spectrum between $F1(u,v)$ and $F2(u,v)$. We must obtain the resulting normalized cross spectrum for computing the relative displacement.

$$\frac{F_2(\omega_x, \omega_y)F_1^*(\omega_x, \omega_y)}{|F_2(\omega_x, \omega_y)F_1^*(\omega_x, \omega_y)|} = e^{j(\omega_x \Delta x + \omega_y \Delta y)} \tag{11}$$

Now we easily derive the relative displacement Δx and Δy. After execution of reverse Fourier transformation we obtain Dirac delta fiction with center at (Δx and Δy):

$$\mathscr{F}^{-1}\left(\frac{F_2(\omega_x, \omega_y)F_1^*(\omega_x, \omega_y)}{|F_2(\omega_x, \omega_y)F_1^*(\omega_x, \omega_y)|}\right) = \mathscr{F}^{-1}\left(e^{j(\omega_x \Delta x + \omega_y \Delta y)}\right) = \delta(\Delta x, \Delta y) \tag{12}$$

The suggested procedure is assumed real functions with infinite domain of values. In image processing, the procedure must be applied to the discrete image function of finite size. The solution of this problem is using a discrete Fourier transform version assuming periodic enlargement of image functions when Dirac delta function is replaced by a unit impulse. It is proved that under these assumptions, the results are still valid [5]. The resulting displacement of images is determined by searching the cross power spectrum in the spatial region, and finding the maximum, which corresponds to the term $\delta\delta(\Delta x, \Delta y)$ of the equation (12).

Image Stitching

Image stitching is the process of combining multiple image frames of the same scene with overlapping parts. The process is in most cases solved by software, but there are already devices when the entire process is handled directly in hardware device that creates the final image during the acquisition of images. Pictures can be taken at different times, from different view or sensors. This method is widely used in image processing, when the method is used to create images with high resolution, such as digital maps, satellite images, and photos of the sky or a high-quality picture for using as a reference image for further processing. The combination of the image can be divided into step registration and fusion of images [6].

Images Registration

Image registration allows determining the geometric transformation between images. Registration determines the degree of transformation of the image, with the help of which we can reconstruct the image, which correspond to the reference image. One of the most popular methods for image registration is the phase correlation. It is based on the Fourier transform and the aim of this method is to obtain a displacement between the two images, which are partially overlapped. Phase correlation has a sharp peak at the shift between images in the cross-power spectrum. Methods for image registration are classified according to various viewpoints and can be classified according to the possible criteria:

Dimension - registration methods can be divided according to time and space dimension of reference image with the source image (2D/2D, 2D/3D, 3D/3D).

Areas of local and global transformations - The global transformation taking transforming the image as a unique complex, unlike local transformation, where different type of transformation with different parameters applied to each part of the image.

Automatic vs. Interactional - Registration methods can be divided into automatic, semi-automatic and fully manual. The manual method of registration is performed fully by user. Semi-automatic method performs several registration steps automatically, but depends on the user to verify the accuracy of the result. A fully automatic method is performed automatically for all steps of registration.

Domains - The registration process can take place in the frequency or spatial domain. Spatial methods work in the domain of image by searching intensity or symptoms in the image. Registration methods which operate in the frequency domain find parameters of transformation for using in the spatial domain.

Modality - It is a division of the multimodal and monomodal. monomodal include all the pictures that were taken by the same method. The multimodal methods are compared to images taken from different modalities (different instruments, different physical principles).

Images Fusion

The process of joining images comes into play when the source image pixels are mapped to the reference image and we need to decide how these images are further processed to achieve the best possible result. In the case that the registration process was carried out correctly, the only fusion of image is executed. However, in real images may be blurred, or we can see visual transition between images or view so called ghosts. We need to determine which pixels can be used and what weight give them in the process of joining of images to create a best clear image [6]. The simplest method to create the final image is to take the middle value of each pixel of the images. It calculates cut-out (border) of the first image, where the second image will be transformed. Then take the points of cut-out of the first image and by reverse transformation is determined location, where it would be projected in the second image. It specifies the color of the place and it will set the point of the first image. Brightness is determined in two ways, either to find the nearest real point (nearest neighbor), or interpolates the color of the four nearest neighbors [7]. If we determine the brightness function of first image as I and similarly in the second image J, it will be applied to:

Method of the nearest neighbor

$$I(u,v) = J(round(H^{-1}(u,v))) \tag{13}$$

Bilinear interpolation

$$I(u+\alpha, v+\beta) = (1-\beta)((1-\alpha)J(u,v) + \alpha J(u+1,v)) \\ + \beta((1-\alpha)J(u,v+1) + \alpha J(u+1,v+1)), \tag{14}$$

$$\text{where}(x+\alpha, y+\beta) = H^{-1}(u,v), \alpha, \beta \in (0,1); x, y \in N$$

Design, Implementation and Testing

$FOTOM^{NG}$ system was developed at the Department of Computer Science FEI VSB-TU Ostrava, and it is a system for image processing [8]. The main requirement for the system was modularity, so that the system was developed by modern technology, and for any modification or addition of module as simple as possible. The aim was to create a new module that can be easily integrated into the existing system $FOTOM^{NG}$ to take advantage of all the elements of the system $FOTOM^{NG}$. System was developed in Java Netbeans Platform version 7 in NetBeans IDE 7.0.1 and have

been utilized already implemented JAI libraries and Xuggler which was achieved much faster to write source code that are easy to understand and quick enough [9]. The development of a module has been divided into two logical parts. The first part deals with the analysis and correction of images with the goal to create high-quality images for later use. The second part perform analyzes and correction of geometric objects in the images in system $FOTOM^{NG}$.

Image Correction

- **Specification of requirements** - Requirements for the first part of the module are:

    - Load images and videos of various formats.

    - The loaded video decode into the images and allow to create video from pictures after their correction.

    - View all loaded images on the screen along with a detailed view of the selected image from the series.

    - Fill in the missing parts of objects in an image by user-defined color or by average values of all images or by the values of the reference image.

    - Allow the user selection from a series of images and choosing reference image for correction.

- **Design GUI** - GUI appearance was defined from specification of requirements. The next step was the choice of controls and adding them to an existing graphic design of $FOTOM^{NG}$ system. The appearance of the GUI is shown in figure 1.

Fig. 1. GUI of new module for image correction

- **Correction** - The most important part of the module was to find a shift between two partially overlapping images so that the images then went align and complement their missing parts. Delivered test images have shift between themselves at the range from three to five pixels. Deformation of images was the similar; the magnitude of deformation was from two to four pixels. Phase correlation will be used to detect movement. Pre-processing step take place before correction, thus removed from the image imperfections that could lead to an inaccurate determination of displacement and bad filling of area.

- **Removal of noise** - The noise is removed at the first step by applying a median filter, set a matrix of size 5x5. We need more highlight objects of interest and remove uninteresting objects after removing the noise. One of the methods of achieving this is to use thresholding methods, when properly chosen threshold in our case it was the automatic detection threshold of the histogram.

- **Registration** - As already mentioned, the most important is to determine the displacement between two partially overlaid images, and it can be obtained by registering the image. For registration of images were created class *ImageCorrectionProcess* and *ImageRegistration*. Another process is divided to two parts. The first part is the calculation of the delta and setting its calculation. The second part combines the average frame and adds the missing parts. Next, the shift for the frame is selected, through which the image is aligned and missing parts of the image are complemented by value of the average image or user-defined color. This corrected image is then saved.

- **Modified images saving** - Image is stored after the correction by replacing the defective image or as a new image. We have possibility to choose graphic format of saving image BMP, PNG or GIF. It is possible to set the quality of the saved image, if you select JPEG as the storage format. You can choose 100%, 75%, 50% or 25%.
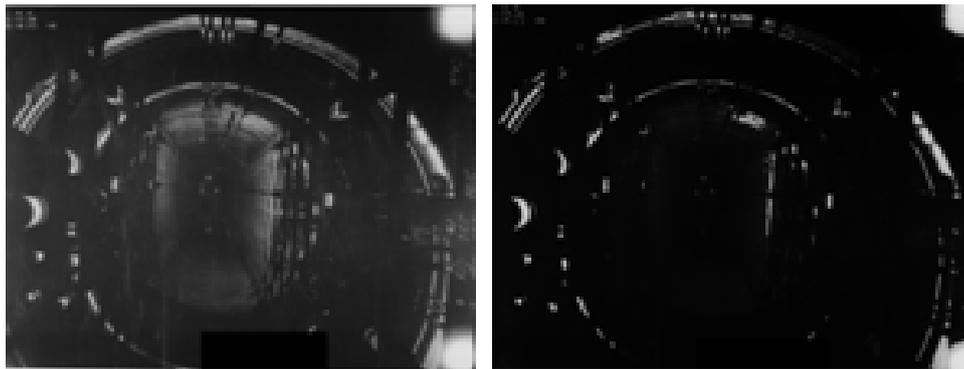
Fig. 2. Left)Using the median filter on mining image. Right)Using thresholding on mining image

Correction of Geometric Object on Mining Image

- **Analysis of existing 2D objects in** $FOTOM^{NG}$

    - Point, see Figure 3a).

    - Intersection, see Figure 3b).

    - Grid, see Figure 3c).

    - Matrix, see Figure 3d).

    - Polygon, see Figure 4a).
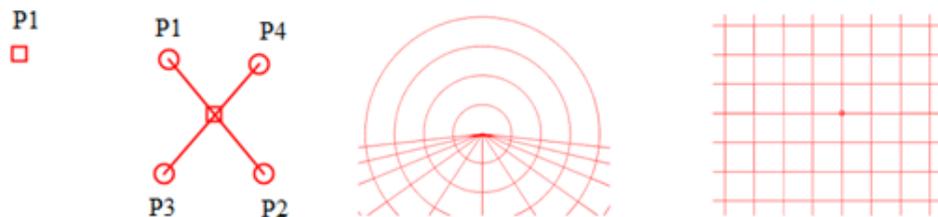
    - Circle, see Figure 4b).

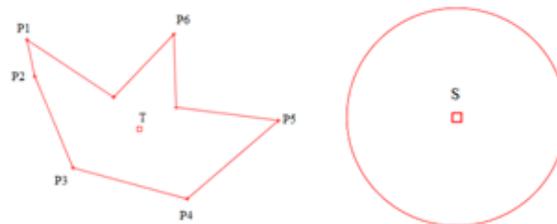*Fig. 3. a) Point b) Intersection c) Object grid d) Object matrix*

Fig. 4. a) Object Polygon b) Object circle

The common parameter of all the objects is the name and the position of their center or center of gravity in the examined image. Polygon is the most widely used tool to help define areas which are not regular. The area can be defined manually, by gradually defining individual items or you can use the tool to automatically define an object for large or highly detailed objects, that automatically finds object boundaries, and it defines the polygon. Object circle shall be used primarily for building a circular shape.

- **Specification of requirements** - Correction has to be done through supplementation or modification of an existing object in an image. For correction of edited objects are used objects from other images that have the same names. Added or modified objects will be possible to distinguish by user-defined colors. Objects of reference frames are compared after correction with objects of existing picture. If the object name is not the same

as an existing object images, the object is added to the current frame, if so, the object is modified. Treatment involves setting the average values of the 2D picture of the objects, the circle with a radius at the intersection and the intersection points of lines, all observed properties of the object. If the user chose the possible addition of added and modified objects by color, objects are marked with color selected by the user.

Testing

Testing was conducted on test images and videos. The average processing time per test video lasting 26 seconds and containing 671 images lasted one and a half hours for processing in a single thread, a processing time for one of two shots lasted around eight to nine seconds. Determining the size shift between images was processing time on average seven to nine seconds. Phase correlation spend on average 70% of the total processing time.

In Figure 5, we can see how they work image correction. Image (a) in fig. 5 is mine picture, which is incorrect in specific part in which the light trace is missing. For correction of this track was chosen reference frame (c). Image (b) in fig. 5 1 is the resulting corrected image based on a reference image (c).

The quality and success of the correction depends primarily on appropriately selected reference frames. The difference in intensity or in histogram between the reference image and a defective image should be as small as possible in order to properly determine location of missing part of the image, which have to be filled by average value or a user-defined color.

Another important factor is the degree of deformation of the image when the deformation of the object of interest exceeds its own boundary in the correct object state, which can lead to addition of parts that are correct and don't require correction. This way leads to poor adjustment and correction can be described as a failure.
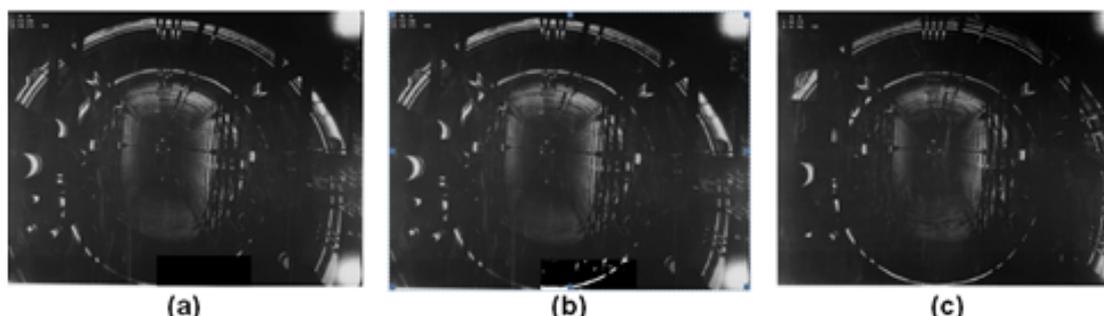

Fig. 5. a) Image with missing part, b)Image with correction, c) Reference image.

The second part of the testing carried out on images containing objects of FOTOM who had poorly defined its properties or frame does not contain the desired objects. These objects have been added or modified by objects from reference frames. Figure 6 shows the course of a test portion of this module.
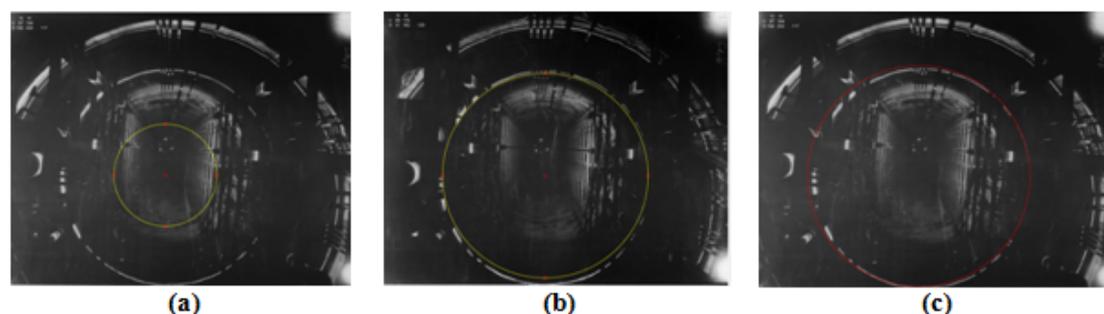

Fig. 6. a) Image with objects for correction, b) Referential image, c) Corrected object from the image (a) using by referential image (b).

Conclussion

The aim of this work was to develop a module for correcting the objects in the image and combination of images for creating of one high-quality image. Images with the new module can be processed from images and from the video. The module was implemented into the $FOTOM^{NG}$ system as photo-image-correction module. The solution is suitable illustrated and finished work has been tested on appropriate images. The major functionality and usability was validated. Tests confirmed the function of each part of the module, its accuracy and implications of integration. Based

on these tests, we can say that the new module is ready for deployment. $FOTOM^{NG}$ is a unique and comprehensive tool for working with mining images. At the present time are developed two additional modules for the $FOTOM^{NG}$ system. The first module is developed for finding the subject in the frame and a second module for the segmentation of objects in an image using active contour method and level-sets.

References

[1] L. Licev, M. Babiuch, T.Fabian, and R.Farana, *Develompent of methods for the processing of mining images using genetic algorithms.* Acta Montanistica Slovaca, vol.17, pp. 218-223. FBERG Kosice, 2012. ISSN 1335-1788.

[2] J.Firt and R. Holota, *Digitalizace a zpracovani obrazu.* [online], 2008. Available from: <http://home.zcu.cz/~holota5/publ/DigZprO.pdf> [cit. 2013-03-26].

[3] T.Zikmund, *Mereni rychlosti objektu pomoci Fourierovy transformace.* VUT Brno, Diplomova prace, 2008.

[4] H.Shekarforoush, M.Berthod, and J.Zerubia, *Subpixel Image Registration by Estimating the Polyphase Decomposition of the Cross Power Spectrum.* In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1996. Available from: <http://cil.cs.ucf.edu/pdf/foroosh_cvpr96_polyphase.pdf> [cit. 2013-04-06].

[5] E.Sojka, *Digitalni zpracovani a analyza obrazu.* VSB TU Ostrava, 2000. ISBN: 80-7078-746-5.

[6] R.Szeliski, *Image Alignment and Stitching.* A Tutorial [online], 2013. Available from: <http://research.microsoft.com/pubs/70092/tr-2004-92.pdf> [cit. 2013-03-10].

[7] M.Krcmar, *Spojovani obrazu.* [online], 2013. Available from: <http://cmp.felk.cvut.cz/cmp/courses/pvi2000/Labs/Ulohy/krcmar.1/1.htm> [cit. 2013-04-22].

[8] L.Licev, *System FOTOM 2008 a vizualizace procesu mereni.* In: Proceeding of GIS Ostrava 2009, 2009. ISBN: 978-80-87294-00-0.

[9] H.Bock, *Platmorma NetBeans.* Computer Press, Brno, 320p., 2010. ISBN: 978-80-251-3116-9.