Approaches for Automatic Comparison of Laser Burned Samples

Jana Hájková

Abstract — **This paper deals with the problem of simulation of laser burning process and comparison of a burned samples during process of automatic verification of this simulation. At the beginning it gives a short overview of the project of laser simulation which is processed at the University of West Bohemia. The basic facts are described to understand the whole approach well.**

The main part of the paper concentrates on the possible methods for comparison of the samples burned by the laser into specific kind of material. Possible alternatives of differences between samples are discussed. Further on, it considers usage of these methods for the verification of the laser simulation system. Problems of verification and their possible solutions are discussed.

At the end of the paper future plans for the project are introduced. The stress is laid on the various alternatives of system verification and its automation.

I. INTRODUCTION

Our work has been created as a part of the large project which is solved by several departments of our university. The aim of the whole project is to develop the real laser system (a sophisticated HW device plus supporting SW) which consists of several parts. This laser system will serve for really advanced physical experiments which are executed as a part of scientific research at the Department of Physics. The HW laser itself services the Department of Physics. The second group is from the Department of Cybernetics, who is responsible for direct controlling of the laser device. The task for our group from the Department of Computer Science and Engineering is to develop a software which will be able to simulate a burning process. The last partner of the project is the commercial firm Lintech who will use the system for common technical tasks. This is the second usage of the system. The whole project and its parts are in detail described in [1].

Somebody could think that such devices already exist for commercial use. That is partly true, because they do not enable to take all possible advantages of up-to-date lasers (e.g. affecting all burning parameters, ...). Moreover, they do not enable to use simulation. The usage of the simulation can serve for various kind of optimization and saves time and recourses that would be needed for repeating unsuccessful experiments.

J. Hájková is with the Department of Computer Science and Engineering, University of West Bohemia, Univerzitní 22, 306 14 Pilsen, Czech Republic (phone: +420 377 632 434, e-mail: hajkovaj@kiv.zcu.cz)

Simulation of the whole burning process is a large software system. Its one part is a preparation of data describing a laser capability. The main part of this preparation is processing of a huge amount of real samples of laser burned hole. These holes were burned into a specific material in exactly defined procedure. Samples are scanned
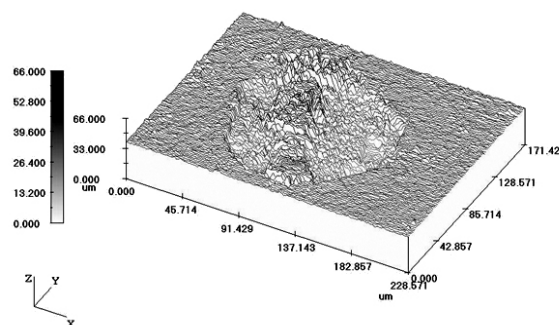

Fig. 1. Height map of the measured sample. The sample of 100 pulses burned into one point into cermet.

by confocal microscope. The task of processing of these samples is to prepare an accurate description of them, which will serve as a main source for simulation of a burning.

A. Data Acquisition

The way of data acquisition can be described in several steps. First, the material and laser parameters are defined and the set of real samples is burned. A set of input data consists of several samples with different counts of laser pulses burned into the material into one point. Each sample (an example can be seen in Fig. 1) is measured separately by the confocal microscope and saved in the form of height map and represented by an array of real values. As can be seen in Fig. 1, a part of the material surface is modified by the laser beam during the burning process, while its surroundings remain unchanged. You can see that the sample is not a hole only, but some tang is always an integral part of the sample. Due to we will use a term "a pulse", which means "a shape of a material formed by one laser pulse". During the data preprocessing the pulses are detected and extracted from all input samples and they are saved as modified height maps. The basic level of the material is represented by the zero values, the positive values represent material upon the basic material level and the negative ones represent material which has vaporized during the burning. This format enables simple usage of pulses for the simulation. The methods for pulse extraction and also the simplified simulation process are described in [2].

After the input data is preprocessed, it is used for the simulation. For better imagination we can use a very simple

example. The task is to burn a pattern of a given shape and dimension into selected material. Without the simulator we would have to repeat burning with various laser settings, until the result agrees to our requirement optimally. While the simulator can test all these settings in the virtual way and find the optimal parameters and laser beam intensity for the real burning. The only input it needs is the data set that represents the characteristic of the laser and the material. For the right function, the simulation has to be verified.

The ways of simulation verification are described in Chapter II. Chapters III and IV specify possibilities of samples comparison and overlapping that are used during the verification. Our results in the comparison of samples and the simulation verification are described in Chapter V. Chapter VI concludes the paper.

II. Simulation Verification

For the correct function of the simulation the system has to be verified. During the verification the burned samples are compared with the real ones. The real samples for verification are gained by the same method as it was described in Chapter I. The system verification can be done manually by comparing and evaluating results visually or it can evaluate results automatically.

The manual evaluation can be used for single samples. Selected sample is burned by the simulation and compared with another one that is really burned into the material and measured by the confocal microscope. In dependence on the result of comparison, the simulation is adjusted. The comparison is done for several combinations of material and several number of burned pulses or other type of burning sample (e.g. burning along the trajectory).

To test and evaluate the system in a more global way a broad range of samples has to be simulated and compared. An automatic verification is used for speeding it up. In this phase, the whole experiment designed for verification can be evaluated at once. For each single compared case several parameters are computed. Based on these parameters the results can be marked as accurate or problematic. Problematic sample has to be reoperated or the simulation has to cover such problems so that the simulation becomes more precise. After the system is verified it is able to burn the samples as realistic as possible.

A question arises – which parameters to use to represent the difference well enough. Possible methods and approaches are described in following chapters.

III. Samples Comparison

Each time during the verification two samples are compared, the real one and the sample created by the simulation. At the beginning we should discuss possible variants of samples and location of the pulse in the sample. Pulses can be shifted towards each other, they can be scaled or they can have even different shape. Several basic possibilities are shown in Fig. 2 from top view and in Fig. 3

from the side view as the cross-section curves. First sample (a) is compared with the other ones. The other sample can be shifted (b), scaled (c) or can differ from the first one in the shape (d). Of course, all these differences can be combined together. A question arises, which one of the simulated pulses is more similar to the real one and what pulse characteristics influence the difference most.

Which of described modifications is most problematic? The most uncomplicated situation is in the case of vertical
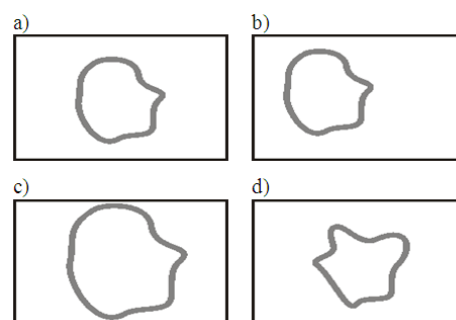

Fig. 2. Possibilities how can the sample (a) differ from another one which is shifted (b), scaled (c) or has different shape of pulse (d). All variants are examined from the top view.

shift (Fig. 3b). Both samples can be aligned to the same height by recounting the surface of the basic material. This case arises quite often, because it corresponds to different heights of basic material levels of both compared samples. Solving of the horizontal shift (Fig. 2b) is a little bit more
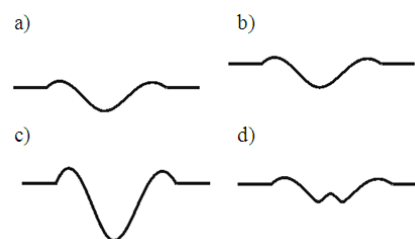

Fig. 3.: Possibilities how can the sample (a) differ from another one which is shifted (b), scaled (c) or has different shape of pulse (d). All variants are in the form of cross-section curve examined from the side view.

complicated, methods for pulse detection described in [2] are used. The problem of scaled sample (Fig. 2c and Fig. 3c) has to be solved by the simulation model modification. The worst situation arises in the case, when pulses differ in the shape as it is visible in Fig. 2d and Fig. 3d. This could be solved partly by changing the input experiment used for the simulation model setting and its verification and also probably by modification of simulation itself.

Of course, to get correct comparison results, samples must be saved with the same precision and resolution or they have to be recomputed before the comparison itself.

Also the roughness of surrounding material can increase rate of inaccuracy of compared samples. The surface of the sample is not optimally smooth, but the roughness of the

original material and local defects increase faults in the computation. That is why we would not get a zero difference after comparison of two real samples with the unchanged material (samples without any pulses) measured in different places. It shows the fact that even samples of the same basic material will differ. This happens also in the case of samples in which the pulse fills the majority of the sample surface. To increase the precision we have to evaluate only the area of burned pulse and closest surrounding. For this purpose, the methods and techniques used for automatic pulse detection described in [2] can be used successfully.

Let's suppose now that we detected burned areas in both samples and we want to compare them. Even though we expect minimal space around the burned area it is necessary to overlap both samples over each other so that we really compare only burned areas. Usable methods, possibilities and problems of samples overlapping are described in Chapter IV.

Samples are represented as height maps, so that the problem is about to compare two surfaces described in the form of uniform rectangular grid. This representation can be transferred into the format of grayscale image. So thus the problem of two burned samples comparison can be solved by methods and algorithms for image comparing [5]. Methods based on image comparing are not the only possible way how to solve the problem of sample comparison. As a promising approach seems to be the usage of the Principal Component Analysis [4].

We have to decide which information we need to get as the result of samples comparison. Some methods give us one value which expresses total difference. Such methods are suitable for automatic verification. The other group of methods shows difference in the form of real array which can be visualized as a grayscale image. This approach enables better localization of problematic parts. However they are unsuitable for automatic processing.

In following three subchapters, methods used for images comparing which can be used for samples comparison, are described and their usage is showed on concrete samples.

A. MSE

This method computes the mean square error (MSE) of both samples. It is defined as the difference of pixel heights of samples squared for the error highlighting. Differences for all pixels are summed and divided by the sample surface dimension. The resultant value expresses average error of each pixel. The calculation is given by (1), where W and H represent width and height of the samples, $real(i,j)$ and $sim(i,j)$ indicate single points at the given position in the sample height map grid of the real and simulated sample.

$$MSE = \frac{1}{W \times H} \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \left(real(i,j) - sim(i,j) \right)^2 \qquad (1)$$

It is easy to see that for more identical samples the lower result of MSE method is computed. Let us compare two similar samples with the same pulse counts from the real experiment (5 pulses burned into one point in the steel; shown in Fig. 4a-b). Samples are very similar, but nevertheless they differ a little. The MSE for this case reached the value of 0.385. The other case is comparison of two absolutely different samples (samples with 5 and 50 pulses burned into one point in the steel; shown in Fig. 4b-c); MSE = 4.186.

The results of MSE computation for sample height maps are much smaller than in the case of grayscale images. The reason is simple, gray scale images consist of values in the interval <0, 255> while values representing the sample surface reach the maximal values of about a few tens. Moreover, before the computation itself, basic levels of both samples are shifted to the same height and so the differences among values of both samples come more near. After the sum of individual differences is divided by the total dimension of the sample, the result is typical between 0 and 10.

During the MSE computing for the whole sample a computational error arises. It results from among of the points on the basic material level where the zero difference is computed. These points do not add any increment into the total sum, but they are included into the sample dimension and so they distort the result. That is why it is important to
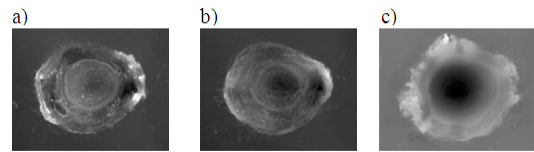

Fig. 4. a-b) Two similar samples with 5 pulses burned into one point into the steel; c) sample with 50 pulses burned into one point into the steel.

compute the MSE only in the area of pulse to get the most accurate results.

B. Difference Image

Difference image is in principle a visualization of the MSE. Color of the pixel is defined as the difference of heights of both samples at the corresponding position recounted to grayscale interval. The dependence can be simply described by (2), where min and max represent minimal and maximal values of difference. The method differs a little from the typical difference image counting, where the absolute value of difference is used. By this modification we preserve information about the order of sample heights (we can find out from the image which sample has the higher value representing the surface in the given pixel).

$$dif[i,j] = 255 * \frac{(real(i,j) - sim(i,j)) - \min}{\max - \min} \qquad (2)$$

The color of the basic material level represents the level of zero difference; all points which are darker are in the

positions, where simulated sample surface was higher than the surface of the real burned sample. In the opposite cases where the simulated sample does not reach heights of the real sample, the pixel is lighter. The points with highest difference are white (real sample is higher) and black (simulated sample is higher).

Results of comparison of the samples described in Chapter A are shown in following images. Two different samples with 5 pulses burned into one point were compared and its result can be seen in the first column of images. The comparison of sample with 5 pulses and the sample with 50 pulses burned into one point are placed in the second column. All samples are burned in the steel. Difference images are depicted in Fig. 5a-b. Together with difference images also horizontal cross-sections going throw middle of the samples are shown. In Fig. 5c-d cross-sections of difference images are visualized and in Fig. 5e-f cross-sections of both simulated (black curve) and real (gray curve) sample are presented.
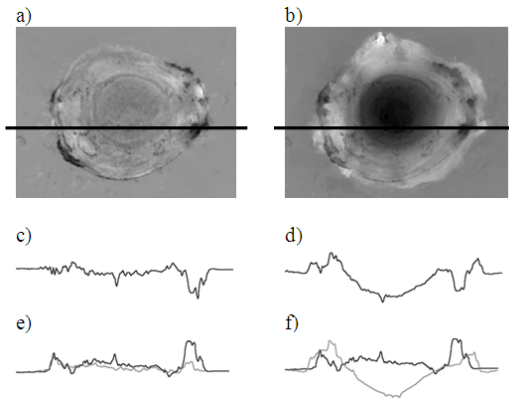
Fig. 5. Difference images of comparison of two similar samples (the left column) and two absolutely different samples (the right column). a-b) Difference images, c-d) horizontal cross-sections of difference images, e-f) horizontal cross-section with curves for both simulated (black curve) and real (gray curve) sample.

C. *Logarithm Operator Adaptation of Difference Image*

For difference images with broad range of values, small differences can not be distinguished well, because they are in the image represented as very similar or even the same gray color. If we are interested in these small difference values, we have to highlight them. It is possible by using the logarithm operator [3]. We can use logarithmic function as a mapping function for difference image modification. During this operation each pixel value is replaced with its logarithm.

For logarithmic modification of images it is typically used function in (3). It expresses logarithm operator mapped on the standard difference image with the absolute value of difference. The basis of logarithm does not influence the result; there can be used e.g. natural logarithm or the base of 2 or 10 logarithm. The multiplicative constant of the expression ensures scaling to values in the interval <0, 255> representing gray color. Because the logarithmic function is not defined for 0, the value of 1 is added to the parameters of logarithmic function used in the expression.

$$difLog[i,j] = \frac{255}{\log(1+255)} * \log(1 + Abs(real(i,j) - sim(i,j))) \quad (3)$$

For the comparing of samples the logarithm operator has to be used in a modified way. We have to distinguish points, where the height of simulated sample is higher than the surface of the real one (such points are in the difference image placed under the zero level) from the opposite case. That is why we have to apply logarithmic operator separately on the points above and under the basic material level. The material level can be declared as zero, because both samples have been aligned before the sample comparison.

The modified expression (4) goes from the value dif[i,j] computed by (2). The value M is counted for each of two cases separately as a number of gray colors between the material base level and maximal difference in the relevant direction. For example, if the basic material level is represented by the gray color with the intensity 160, all pixels above the material are spread into 95 levels - they have the value in the interval <0, 159>. Value of the fraction at the beginning of the expression differs for converting for the points above and under the level of basic material.

$$difLog[i,j] = \frac{M}{\log(1+M)} * \log(1 + dif[i,j]) \quad (4)$$

The results can be seen in the following images; in Fig. 6a the original difference image is shown, in Fig. 6b the difference image modified with the logarithmic operator can be seen. The difference of both images is visible well. The original difference image defines the color equally in dependence of the displayed color. From this representation we can get good imagination of heights distribution. On the other hand, from the logarithmical modified image we can better find out small differences in the neighbourhood of the material level.

IV. SAMPLES OVERLAPPING

To get optimal result for simulation verification it is necessary to overlay burned pulses as exactly as possible. Otherwise the error of comparison is enlarging in dependence on relative shift of burned area in the sample. The manual overlapping is possible, but it can not be used for automatic verification and results evaluating. That is why the system has to be able to find the best matching position independently.

One possible way how to match pulses automatically is to compute any parameter representing the difference of both samples (e.g. MSE) in a given position for all possible positions of two samples. It could be the optimal solution, if it were not for the problem of speed. E.g., for two samples of the size 1024×768 values would be the parameter computed 768432-times. If one calculation takes 1ms, the best position of samples would be decided after approximately 128

minutes. That is why we have to reduce number of repetitions as much as possible, but at the same time we have to preserve sufficient precision of the result.

First, we can eliminate all the positions, in which only borders are overlapped (as shown in Fig. 7). We can suppose that no optimum will be probably among these combinations.
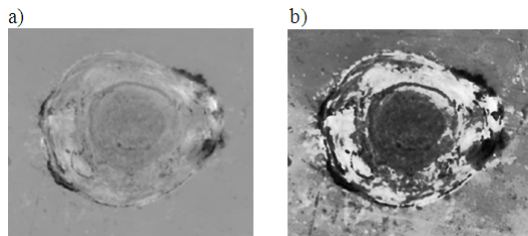

Fig. 6. Difference images of two compared samples a) the original difference image, b) difference image with logarithm operator modification.

We have to find the best starting position (a guess of optimal position of overlapped samples) and from there we have to start searching the real optimal position of overlaid samples in its surrounding. We can use, e.g., the centre of mass of the sample as the starting position (the method of centre of mass modified calculation is described in [2]). The centre of mass can be computed for both compared samples and samples can be shifted to overlay their centers of mass. Another approach offers the usage of the method of automatic pulse detection, to detect pulses in both samples and to overlay the selected areas. In the second method we can directly take advantage of the pulse detection for the faster and more accurately computation of difference only in the selected area.

After finding the starting position we just have to find the optimum by small shifting of samples over each other. In further positions the step for shifting is rougher, in the close surroundings the shift step is chosen smaller. For each of these combinations the difference of the samples is computed, so the algorithm of difference parameter computation has to be quick enough to reach the computation acceleration.
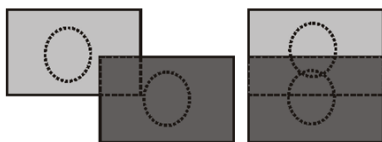

Fig. 7. Positions of overlapped samples which can be most likely excluded from searching for optimum.

V. RESULTS

The first thing to be solved during the project was the data preparation, processing and its visualization. Because no tool for our format of data was available so far, we had to start development of a new software tool that would help us with data processing. Many functions for pulse detection and extracting, samples viewing, modification or comparing were required. Plenty of functions which are divided into several categories in dependence on the phase of data processing were designed and implemented. The most often used functions were matched with key short cuts for the operations acceleration.

The tool works in three modes of data processing. The first one serves for sample viewing and exploring. The second mode serves for the simulation of burning, in which the extracted pulses are used. The third mode is used for samples comparing and simulation verification. It enables to compare any two samples and to determine the size of difference as the MSE coefficient. Also the visualizations of both samples and their difference image and logarithm operator adaptation of difference image computation belong to basic functions of this tool. It enables the manually and automatic comparison of samples and it will be used for the simulation verification.

Presently, we use the comparison functions for exploring of measured real data. We dispose with data set burned into steel, which contains samples of 1, 2, 5, 10, 20, 30, ..., 100 pulses burned into one point of the material. Each number of pulses was burned five times and saved as a separate sample. For the next process of our work we have to explore all similar samples (it means samples with the same number of pulses burned into the same material) and to detect differences and similarities.

For the result demonstration we decided to use three samples with 1 pulse (Fig. 8a-c) and three samples with 100 pulses burned into one point (Fig. 8d-f) burned into the steel.

All results are summarized in following tables. Each table contains for the selected combination of samples the cross-section curves for both compared samples, difference image and cross-section of difference image. Finally the MSE is computed. It should be highlighted that the MSE is computed only in the area of detected pulse (as it was explained in the introduction chapter). In Table I results of comparison of two absolutely different samples b) and f) are described, Table III and Table II summarize results for similar samples a-c) and d-f).
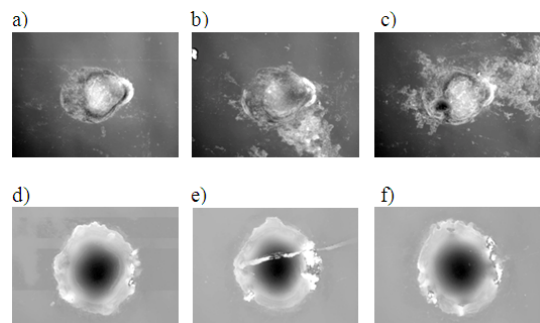

Fig. 8. Similar samples with a-c) 1 pulse; d-f) 100 pulses burned into one point into the steel.

VI. CONCLUSION

The height difference between the maximum and minimum is in the case of samples a-c) around 3.5μm. In the majority of samples any local defect affects the surface of the

sample evidently. The reason is simple – the height of local defect is approximately 2μm, which is in comparison with the sample surface a radical distortion of the surface. On the other hand, the height difference of samples d-f) is around 14μm and so the local defect does not affect the result so much and it is relatively insignificant.

The verification and comparison should help to create the reliable simulation system with the ability of automatic selection of optimal result. Parameters of the optimal simulation can be used for setting up the real device. This approach should avoid executing and burning of redundant experiments, which is money and time consuming.

Our actual results show that the task of samples comparison is not trivial and requires comprehension of a

TABLE I
RESULTS OF COMPARISON OF TWO ABSOLUTELLY DIFFERENT SAMPLES
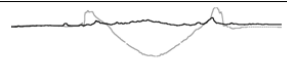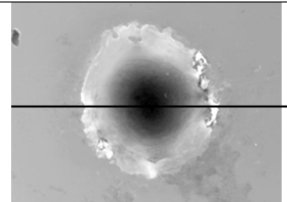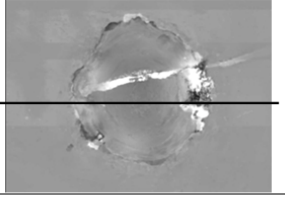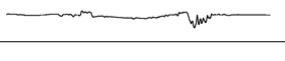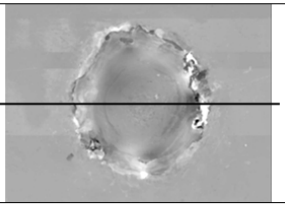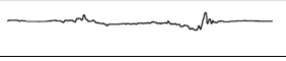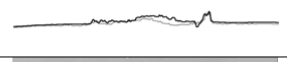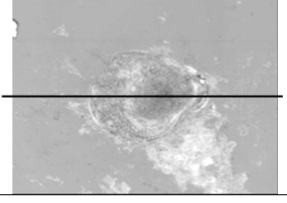
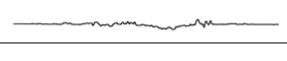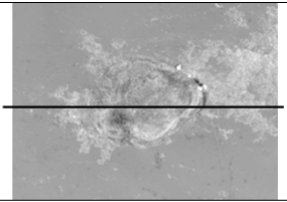|	*(MSE = 10.924)*
both cross-section curves of compared samples	
difference image	
difference image cross-section	

TABLE II
OVERVIEW OF COMPARISON RESULTS A-C)

Comparison of samples d) and e)	*(MSE = 2.429)*
both cross-section curves of compared samples	
difference image	
difference image cross-section	

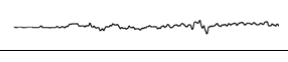| *Comparison of samples d) and f)*	*(MSE = 1.175)*
both cross-section curves of compared samples	
difference image	
difference image cross-section	

TABLE III
OVERVIEW OF COMPARISON RESULTS D-F)

Comparison of samples a) and b)	*(MSE = 0.028)*
both cross-section curves of compared samples	
difference image	
difference image cross-section	

Comparison of samples a) and c)	*(MSE = 0.094)*
both cross-section curves of compared samples	
difference image	
difference image cross-section	

Of course, we have many future plans for our further research. They are divided into several groups in dependence to which activity it is related to. All of them concern the part we are working on – data processing, software simulation and result visualization. The other (especially hardware) parts of the project have of course their own individual plans for the future.

REFERENCES

[1] Hájková, J., Herout, P., "Laser Simulation," *Proceedings of the 7th International Conference APLIMAT 2008*, ISBN: 978-80-89313-03-7, pp 823-828, Bratislava, 2008.
[2] Hájková, J., "Methods of Pulse Detection in Laser Simulation," *Proceedings of the 3rd International Conference on Software and Data Technologies ICSOFT 2008*, ISBN: 978-989-8111-57-9, pp 186-191, Porto, 2008.
[3] Logarithm Operator, March 2008,
URL: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/pixlog.htm>
[4] Smith, I. L., *A Tutorial on Principal Component Analysis*, 2002,
URL: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.
[5] Zapletal, J., *Application of radial Basic Functions*. Diploma Thesis, University of West Bohemia, Pilsen, 2007.

number of smallnesses, which influence the final result. One of critical tasks is the automatic pulse detection.