Mobile system for estimation of the internal parameters of distributed cameras

Przemysław Mazurek

Abstract: Estimation of the internal parameters is very important task for photo- and videogrammetric applications. Such operation for multiple cameras is time consuming task and the method based on the mobile backlight illuminated calibration board is proposed. Proposed system uses wireless link for transmitting calibration results with the pattern features detection results. OpenCV library is applied for calibration process and libdc library for real-time acquisition of IIDC compliant camera measurements.

Keywords: Camera calibration, Distributed acquisition systems, Embedded systems

1. Introduction

Multiple camera systems are used in numerous applications [1]. Two or more cameras could be used for 3D scene reconstruction. Motion capture systems used for estimation of the human motion are good examples [1,9] of 3D scene reconstruction applications. They are used in the computer animation instead manual key-framing, ergonomics, biomechanics, etc. Multiple camera surveillance systems could be also used for the position and pose estimation. Such systems are dedicated for tracking of multiple humans on the street or air terminals for further detection e.g. abnormal behaviors (thief, terrorist, drunker, lost child, ill person, etc.).

Relative positions of cameras should be estimated with the high accuracy for best results and this is a subject of estimation of the external parameters. There are three positions and three angle-orientation parameters so there are six external parameters of particular camera.

The number of the internal parameters of camera is dependent on the camera model [2,3]. The most well-known parameters are: the focus, focus length (zoom), iris diameter, and parameters of the geometrical distortions. All of them could be estimated during calibration for photogrammetric/videogrammetric applications.

Combined estimation of external and internal parameters during the one-time calibration operation is possible but is not convenient because there are much more unknown parameters. Local minima during the optimization process used for the estimation of parameter could give wrong results. Optimization process for the reduced number of parameters is much faster and reliable.

Calibration is human supported operation, and the automation of this process is very hard to do. Estimation of the internal parameters could be automatic using special test bed and overall process could be fast and precise. Such test beds are used by lens or camera manufactures.

Real camera has a lot of distortions dependent on camera settings. Fixed lenses should be mounted and focused on the final location. Optical axis is not in the center of sensor due to limited tolerance of mechanical and optical elements [11]. Complete calibration for all possible settings is not possible. Typical calibration process is related to the specific location and settings of camera.

An example figures 1 and 2 shows importance of calibration. Distortions for the narrow angle lenses are not visible and sometimes are omitted but for wide angle lenses they are very well visible and important for human and machine vision applications [15,16].

Fig. 1. Example subject of the internal parameters estimation (large scale geometrical distortions for fish eye lens)

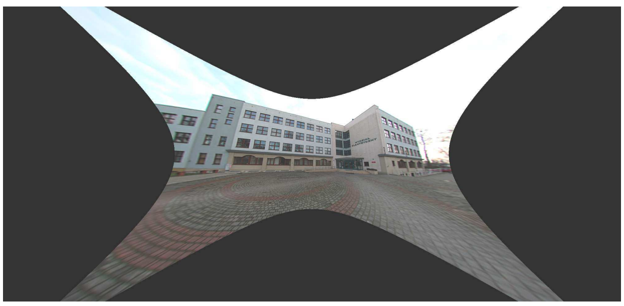

Fig. 2. Undistorted image (after crop) for 180deg diagonal lenses

Camera calibration process is based on calibration objects typically and the calibration pattern is used during estimation of the internal parameters. Dots or chessboard patterns are used typically and the two examples of them are shown in Fig.3. There are numerous variants of them, dependent on the calibration software tools. Spatial relations between feature points of the calibration pattern (dots or corners) are a priori known. Distortions could be measured by the comparison detected features from the acquired image and the pattern model.

Chessboard pattern are typically used with multiple presentations (different positions and orientations). Pattern selection is dependent on the image processing. Corner detection algorithms are used with subpixel accuracy for the estimation of the features (corners of the black squares).

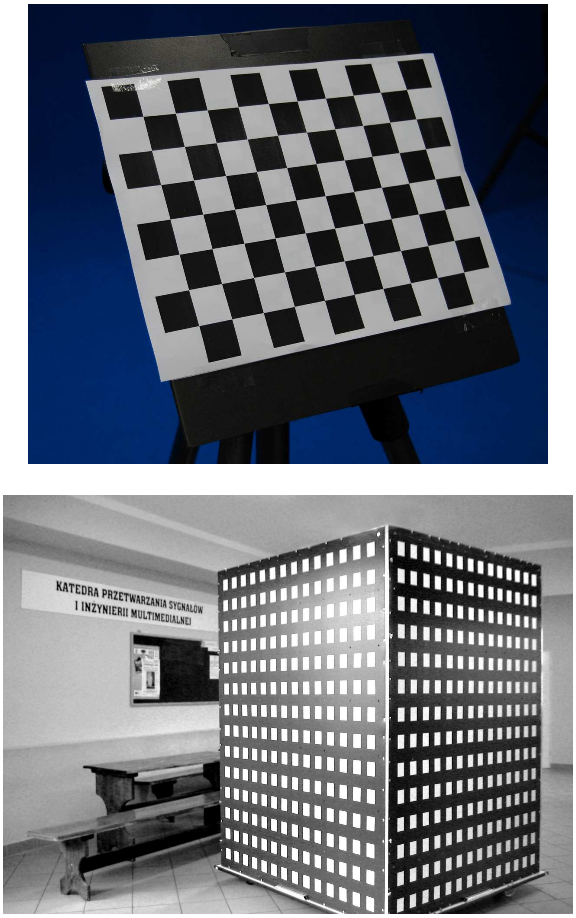

Fig. 3. Example calibration patterns (simple flat A4 size pattern and very large calibration object with four faces)

Multiple presentation of the calibration pattern at different position, distance, and angles are necessary for estimation of the internal parameters what is time consuming operation for human operator. Cameras could be placed at different height and angles, and tens or hundreds of different presentations (images) are necessary for proper calibration.

Calibration speed-up in necessary especially for the motion capture systems – there are tens of cameras. This process should be repeated every time if the camera parameters are modified what occurs when focus and iris rings are used.

The calibration pattern should be presented in near distance to the camera. Most calibration tools needs complete pattern presentations.

The calibration pattern is not physically ideal and presentation from larger distances is preferred for a spatial noise suppression (errors related to the positions of features is reduced). Multiple presentations with different positions (distances and orientations) reduce pattern errors and assign irregularities to the local distortions of the lenses more reliable.

Pattern presentation before camera mounting is not possible every time. Blocking screws in lenses are available for the quality optics only. Mechanical problems related to the camera are especially for mobile robots.

2. DISTORTION MODELS

There are a few image distortions models used for image correction. There are two kinds of models – the parametric and non-parametric. The parametric models [2-4] use simple distortions models based on the polynomials so the number of parameters is limited to few only. Correction of distortion is quite simple from algorithmic point-of-view [12] if the parameters are knows but not all kinds of distortions could be corrected.

OpenCV library [2,13] uses parametric model with the two kinds of distortions: radial (1) and tangential (2):

$$\begin{aligned} x_{corrected} &= x\left(1 + k_1 r^2 + k_2 r^4\right) \\ y_{corrected} &= y\left(1 + k_1 r^2 + k_2 r^4\right) \end{aligned} \tag{1}$$

$$\begin{aligned} x_{corrected} &= x + \left[2 p_1 y + p_2\left(r^2 + 2x^2\right)\right] \\ y_{corrected} &= y + \left[p_1\left(r^2 + 2y^2\right) + 2 p_2 x\right] \end{aligned} \tag{2}$$

where (x,y) is the original location on the image of the distorted point and ($x_{corrected}$,$y_{corrected}$) is the new location. Radial distance from the optical axis (image center) to the particular position is marked as ‘r’. Estimation of correction parameters is the subject of optimization process.

Non-parametric models are based on the vector field for detailed description of global and local distortions. Distortion parameter is the vector set. Number of vectors is dependent on spatial sampling density.

High quality lenses with minimized local distortions are very expensive so calibration process gives ability of application of much cheaper lenses with comparable quality for machine vision applications.

3. CALIBRATION METHOD FOR DISTRIBUTED CAMERAS

Correction of distortions is possible in camera software/hardware but the most typical solution is correction in the central computer or cluster (for multiple cameras). Such systems are simpler to maintain and camera will be cheaper, dedicated only to the image acquisition, not for processing.

In most papers related to the estimation of the internal parameters a laboratory conditions are assumed. For close camera and computer (with calibration software) case a presentation of the calibration pattern is straightforward because computer screen could be used for visualization of camera image and calibration results. Distant camera needs cooperation of two peoples (the first is responsible for presentation of the pattern and the second is responsible for the process monitoring and directs movements/orientation of the pattern by voice commands). A more convenient system is necessary for convenience and quality of calibration.

Such system should be wireless, battery powered, and a necessary data obtained from feature detection and calibration process should be forwarded to the pattern operator.

The most significant disadvantage of typical systems is the light condition sensitivity. Calibration patterns are passive typically so light conditions in room influent on results (especially on the features detection algorithm).

LCD screen could be used as good source of calibration pattern. This is a not a new idea [6,8,17]. Light emitting panel gives better contrast in comparison to the printed pattern boards and are less dependent on surrounding light conditions. Application of LCD screen for human operator is not so good due to weight (panel, power supply – usually 230V AC, pattern generator), so another lightweight device should be prepared.

The proposed solution is LCD casing (plastic only) with the LCD panel backlight and transparent pattern attached to the front of the backlight. Complete LCD panel consist from four elements: the screen (necessary to remove), a set of backlight diffusing layers, metal frame, and light source with a power supply (usually 12V DC powered). The cold cathode backlights should be not used because there is possible a current flow over operator's body and LED based backlight (e.g. border LED) should be used for improving safety. Typical LCD monitor is grounded and the display glass is also an insulator. After removal of display such condition are not fulfilled especially for battery powered devices. Complete solution for the calibration pattern is shown in Fig.4.

Fig. 4. LCD backlight with calibration pattern

Pattern visibility is orientation dependent. For outdoor locations and bright rooms there is a surrounding light that support of the pattern visibility. For indoor location situation vary (Fig.5) and the backlight pattern is more reliable especially for negative declination angles. Both variants are good calibration pattern sources for the positive declination angles due to top-mounted lamps in the room.

The calibration pattern should be operated manually and feedback from the camera and computer is necessary. Applications of the additional LCD screen attached to the calibration pattern is not possible so additional LCD monitor should be used. Such monitor attached to the harness of mechanical stabilizer ("steadicam") is shown in Fig.6. Mechanical construction fixes position of the monitor (about 7" diameter) and is the mounting platform for Li-ion battery necessary (the calibration pattern and monitor power supplies) and for wireless link.

Calibration software uses Linux ported OpenCV library [13] in continuous way so all accepted images by feature detection algorithm are used by the calibration algorithm. Only feature points coordinates are stored instead all contents of images for reduction of memory consumption.

Correctly detected features are marked on the current image and this image is transmitted to the portable monitor using 1-bit image coding and Bluetooth interface using asynchronous serial mode. The libdc library is used (Firewire 1394 IIDC camera are supported) for acquisition of images from particular camera.

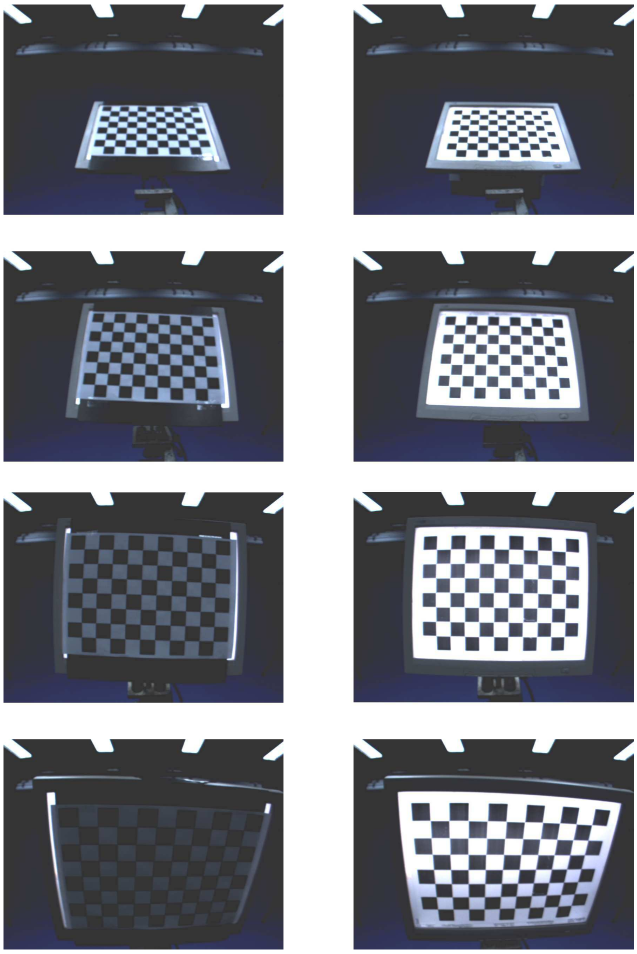

Fig. 5. Pattern visibility for different declination angles. Passive (left) and backlight (right) variants of the calibration pattern. Declination angles (from up to down): +60 deg, +30 deg, 0 deg, -30 deg

Fig. 6. Calibration using proposed system (from left: harness from mechanical stabilizer ("steadicam") and attached monitor, backlight-based calibration pattern, and camera)

Transmitted image data from the computer are received by the microcontroller and converted in to analog TV signal for the monitor (Fig.7). Application of low-cost microcontrollers for image displaying is possible but software generated video image requires a lot of computation power. Resolution of such software generated images depends on computation power and availability of hardware peripherals so bit-bang mode has higher requirements in comparison to the synchronous serial interfaces.

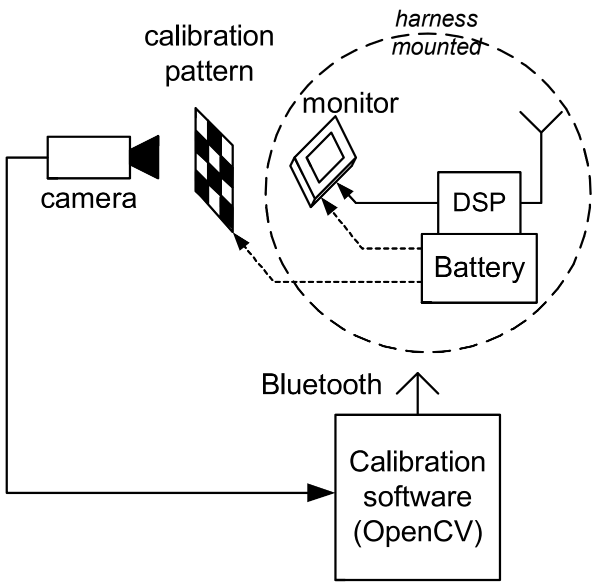

Fig. 7. Schematic of the proposed system

Additional microcontroller resources are necessary for support of wireless transmissions. A lot of modern microcontrollers have enough computation power (e.g. ARM7) but they have very slow peripheral devices (especially GPIO ports) so alternative choose was made. The DSP56303 signal processor [5] (Motorola/Freescale DSP56303EVM module) is used as a controller for video generation and communication.

This processor has a very fast GPIO ports and synchronous interfaces. ESSI (Enhanced Synchronous Serial Interface) [5,14] is used for improving performance instead generation using software only method.

A continuous block of 24-bits could be transmitted with minimal intervention of the main core (using DMA) so overall module has enough computation power for data processing (reception of image data using Bluetooth module). In the further modification a compression algorithm will be added. ESSI port has up to three serial outputs so video image quality could be also improved by applications of them as inputs of 3-bit D/A converter for grayscale or color output. Currently 1-bit image is used.

4. Network Mode of ESSI

Generation of analog video signal is especially convenient due to Network Mode of ESSI (Enhanced Synchronous Serial Interface) availability in DSP563xx family [5,14]. Similar technique could be used for other DSP families not necessary from Freescale (previously Motorola Semiconductor). Network mode is based on the TDM (Time Domain Multiplexing) transmission scheme with up to 32 time slots. Such transmission mode is available especially in DSP processors used for the telecommunication purposes. ESSI could be configured with multiple serial bus masters network but in this application there is only one master (DSP) and only one slave (Digital-to-Analog Converter). Application of other transmission mode is possible with more instruction to process.

Video signal generation using bit-bang mode and GPIO lines is possible but much more costly so image compression and similar task are hard to do especially if simple 8-bit microcontroller is used. ESSI has very important advantage in comparison to the typical serial synchronous channel peripherals – there are 3 synchronous channels with independent data registers so up to 3 lines could be used for bit streams transmission. An additional frame sync and clock output lines are not used in this application. In particular DSP there are two ESSI and synchronization of them is possible so there are 6 lines available.

Another advantage of particular DSP is the internal word size – 24-bits per word so up to 24-bits could be transmitted with single interrupt.

There are the two interrupts used for transmission. The first one is related to the word transmission and the second one is related to the last time slots. The output data shift registers are filled by the first interrupt service routine. The second one interrupt is used for selection of bit pattern for next line because single block of time slot correspond to the single line. Such formulation uses internal (ESSI) register for loop creation without DSP core intervention. Additional modulo addressing modes are used for bit patterns repetition.

There are three memory blocks related to the video signal generation. The first one consist bit patterns for synchronization and corresponding waveforms are shown in Fig.8. Synchronization pattern is outputted by dedicated ESSI synchronous data line. Zero-value corresponds to the synchronization level of video signal (0V) and one-value to the black level (0.3V). Values between 0.3V and 1.0 V are related to the visible video levels, and the 1.0V is white value (Fig.9). In the current implementation a 2-bit D/A Converter is used with black and white output video signal. In the first one block there are 6 synchronization patterns according to the specification of PAL video format.

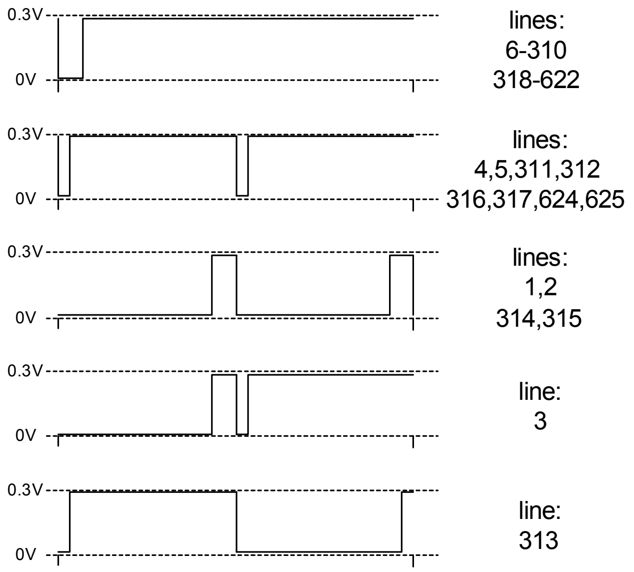

Fig. 8. Examples of synchronization signals

Every pattern is stored using 26 words (26x24bits) with an additional 6 words padding and transmitted using single block transfer of time slots.

The second one table is used for selection of particular synchronization pattern. This table has 625 entries corresponding to the appropriate video lines. This table is quite large but calculation of the line number and selection of the appropriate synchronization pattern is very simple due to addressing with incrementation with hardware supported modulo operation.

The last one table is the video buffer that has 26x625 word size. Theoretical resolution (pixel resolution) is equal to the 629x625 pixels but some lines are not destined for video data, and the beginning part of every line can not be used due to horizontal synchronization pulses.

Vertical resolution depends on PAL video standard but horizontal is pixel clock dependent so it is ESSI constraint.

DSP main clock is set by multiplier (M) and divider (D) registers of PLL and depends on the input clock $f_{crystal}$

$$f_{core} = \frac{M}{D} f_{crystal} \qquad (3)$$

For particular system values: M=8; D=1; $f_{crystal}$=12,288MHz (available on DSP56303EVM board) sets a maximum allowable core clock frequency: f_{core}=98,304MHz.

ESSI clock should be set as high as possible for high horizontal resolution but with two limitations. The first one limitation is related to the allowable bandwidth of the SDTV PAL signal and for digital systems a maximum clock frequency is 13.5MHz [7,8]. An additional analog low-pass filtering is necessary for limitation of high-frequency components of video analog signal. Pixel should be square or rectangular with specified proportions [7] but for non-broadcast purposes it is not important. The second one limitation is related to the number of slots because single TV line should be generated using single ESSI Network Mode transfer (e.g. 24 bits per 30 Time Slots = 720 pixels). The number of time slots should be between 2 and 32. Settings of the pixel clock frequency (ESSI clock) are possible by two ways: using internal f_{core} clock or using external clock generator connected to the ESSI clock input. The first one method is assumed in particular application, but the second one method is more flexible.

Pixel clock is dependent on the core frequency and ESSI divider (E).

$$f_{pixel} = \frac{f_{core}}{E} \quad (4)$$

There are three dividers: the fixed one (by two), and two programmable. The optimal time slot block size should be integer value B_{size} The period of the horizontal line T_{line} is defined by PAL specification.

$$T_{line} = \frac{f_{pixel}}{24 B_{size}} = 64[\mu s] \quad (5)$$

The optimal values are: E=10 and B=26, and for such settings a mentioned theoretical resolution 629x625 could be obtained.

The PAL video system reduces number of horizontal lines so lines from 24 to 320 and from 336 to 622 could be used only for visible image. Two half-lines no.23 and 623 are not considered.

Visible part of horizontal line is limited to the 511 pixels due to synchronization pulses (4.7µs).

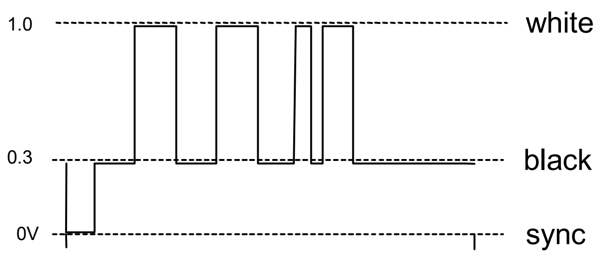

Fig. 9. Examples of synchronization pulse and video signals (1-bit coded) from 2-bit DAC

Real resolution depends greatly on monitor. A low cost LCD monitor support only single field display so vertical resolution is reduced by two. This video signal generator support interlaced video signal but for the flicker free (non-interlaced) mode both fields should be identical. Cheap LCD monitors does not support high horizontal resolutions. This is physical limitation of a panel and very often horizontal resolution is specified as number of subpixels (R, G, and B cells) not a real pixel.

Video image data are stored in the external memory of DSP56303EVM module. Synchronization patterns and line table is stored in internal memories of DSP.

5. Conclusions

Proposed and tested system is very convenient for multiple-camera applications. Estimation of the internal parameters is the time consuming task but for the proposed system it takes about a few minutes for every camera typically (but it depends on focal length). Designed software and electronic parts gives low power consumption (~1.5A, 14.4V DC) sufficient for mobile operation.

Calibration process is based on OpenCV algorithm routines used by specially created software tools for this application. This tool integrates libdc for real-time work with OpenCV.

In further work compression will be considered for improving refresh rate of transmitted image. Image synthesis based on detected position of features (corners) is possible by this DSP due to reduced software requirements related to the video image generation with application of Network Mode of ESSI. DMA channels could be also used for further improvements of system and reduction of number of interrupts.

Acknowledgements

This work is supported by the MNiSW grant N514 004 32/0434 (Poland).

This work is supported by the UE EFRR ZPORR project Z/2.32/I/1.3.1/267/05 "Szczecin University of Technology - Research and Education Center of Modern Multimedia Technologies" (Poland).

Literature

[1] Aghajan H., Cavallaro A., *Multi-Camera Networks. Principles and Applications*, Elsevier, 2009

[2] Bradski G., Kaehler A., *OpenCV. Computer Vision with the OpenCV library*. O'Reilly, 2008

[3] Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc

[4] Dersch, H.: Panorama Tools. Open Source Software for Immersive Imaging, International VR Photography Conference, Berkeley, 2007, http://webuser.fhfurtwangen.de/~dersch/IVRPA.pdf

[5] Freescale Semiconductor DSP56303 Technical Data, 2005

[6] Funk N., Yang Y-H., "Using a Raster Display for Photometric Stereo" *Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*, 2007, pp.201-207

[7] Intersil, BT.656 Video Interface for ICs, Application Note AN9728.2, Intersil, 2002, www.intersilsemi.com/data/an/an9728.pdf

[8] ITU-R Recommendation BT.656

[9] Kitagawa M., Windsor B., *MoCap for Artists. Workflow and Techniques for Motion Capture*, Focal Press, 2008

[10] Komagata H., Ishii I., Takahaski A., Wakatsuki S., Imai H., "A geometrical method for calibration of internal camera parameters of fish-eye lenses" *Systems and computers in Japan* vol. 38 no. 12, 2007, pp. 55-65

[11] Mazurek P., "Lens holder influence on camera calibration" *31. Międzynarodowa Konferencja z Podstaw Elektrotechniki i Teorii Obwodów IC-SPETO'2008*, pp.151-152 +CD, 2008

[12] Mazurek P., "Real-time correction of cameras' geometric distortion using GPGPU", *14-th IEEE/IFAC International Conference on Methods and Models in Automation and Robotics MMAR'2009*, 2009

[13] OpenCV library http://sourceforge.net/projects/opencvlibrary

[14] Redheendran T. M., Freescale Semiconductor DSP56303 Enhanced Synchronous Serial Interface (ESSI) Programming, Application Note AN1784, 2005

[15] Russ, J.C.: The Image Processing Handbook, CRC, 2006

[16] Salomon, D.: Transformations and Projections in Computer Graphics, Springer, 2006

[17] Zhan Z., "Camera Calibration based on Liquid Crystal Display (LCD)" *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII, Part B3b. 2008

dr inż. Przemysław Mazurek
West-Pomeranian University of Technology, Szczecin
Faculty of Electrical Engineering
Department of Signal Processing and Multimedia Engineering
26. Kwietnia 10 Str.,
71-126 Szczecin, Poland
+48 91 4495311
przemyslaw.mazurek@zut.edu.pl