THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC.

SIGNAL PROCESSING
ALGORITHMS, ARCHITECTURES, ARRANGEMENTS, AND APPLICATIONS
SPA 2015
September 23-25th, 2015, Poznań, POLAND

Application of Fast Cameras to String Vibrations Recording

J. Kotus, P. Szczuko, M. Szczodrak, A. Czyżewski
Faculty of Electronics, Telecommunications, and Informatics, Gdańsk University of Technology
Narutowicza 11/12, Gdańsk, PL80233, POLAND

***Abstract*- A hardware and software solution for guitar string vibration measurement by fast cameras is described. Orthogonal setup for 3D image acquisition is proposed capable to capture several thousand image frames per second. Dedicated image processing algorithm was developed and described in the paper, aimed at tracking the movement of some selected points along the string. Fast and accurate tracking results provided a detailed information about vibrations, that was transformed into sound samples. Described sound processing methods were applied in order to enable a comparison of captured string vibrations with the sound recorded using a microphone. Analysis of obtained results, conclusions, and future work plans are included.**

I. INTRODUCTION

The main aim of this research is to obtain a detailed information about picked string movement and its influence on the evoked sound, and compare recreated sound with the reference microphone recording. A hardware and software solution for guitar string vibration measurement by fast cameras was designed and applied for this purpose. Detailed information about the recording setup and image processing algorithm used for image-to-sound transformation are presented in this paper.

A string movement is a complex process, starting with introduction of an energy by pulling the string, releasing it and then propagating energy onto all string segments. This induces guitar body vibrations, as a guitar nut and saddle are in contact with string endpoints. All mentioned elements influence air molecules disturbances inside the body and around it, and final sound is emitted through a sound hole and all elements of the body [18].

The guitar acoustics can be studied by mathematical modeling and simulation (e.g. finite elements), by sound recording and analysis, and finally by direct measurements of vibrations: electromechanical by piezo-electrics [2], electromagnetic pickup [13], an analog position sensing detector [1], laser vibrometry [4] and high speed cameras [10][15]. An interesting idea was presented by Pakarinen and Karjalainen involving string vibrations measurements in two polarizations and in multiple points along a metal string using electric field sensing [14]. Askenfelt and Jansson presented an experimental study of wave motion on a piano string and a corresponding spectra. In this research string motion was measured using an electrodynamic method [2].

As it was mentioned above, string vibrations can be analysed in many different ways. Application of fast cameras for analyzing string vibration is the main purpose of this work. Similar approach can be found in Kartofelev et. al. research. They applied high-speed line scan cameras to capture piano and bass guitar strings motion [7][10].

Our approach involves synchronous visual and acoustical recording: high speed cameras register string vibration, and the microphone records result sound. Thus a detailed observation is possible: movement phases are visualized on a plane graph, the location is transformed into sound samples and comparison with the real sound by means of time and frequency domains analysis is performed.

II. METHOD DESCRIPTION

Visual tracking of vibrating object movement is performed by utilizing synchronized high speed cameras, able to acquire more than 4000 frames per second, for observation of the object from two orthogonal directions. As a result high resolution spatio-temporal measurements of displacement, speed, and acceleration of the vibrating object are obtained.

Using macro lens and high resolution sensor assures precision of the object localization, reaching up to 16 microns. Previous Authors' experience involves research and development of video analysis methods published previously [5][6][17], able to automatically localize and track object of interest and key points visible in the image. Vibrating element (e.g. a single point on the string) can be tracked without interfering its motion. Registered sequence of the object movement is analyzed offline and provides detailed data of high spatio-temporal resolution. A dedicated software for video analysis of object vibrations in three dimensions was developed by the Authors.

A. Image and sound acquisition

Image acquisition system was design to capture spatial vibrations of the string. Two fast cameras were applied for this purpose. Cameras were oriented at the right angle (90 degrees) between them. In this configuration it is possible to record the vibration of a desired point on the string in two views simultaneously. In Fig. 1 the configuration of recording setup is depicted. The whole system was based on two fast cameras: Basler acA2000-340kc. Each camera was equipped with a lens SV-2514V with spacer ring Goyo Optical EXT-ULSET with a 10 mm length and was connected to NI PXIe-1435 Camera Link frame-grabber. During the recording additional external light source was applied. For image capture the StreamPix v5.0 Multi Camera software by NorPix Inc. [11] was used. The region of interest was limited to: 2000x16 pixels. The applied frame rate was equal to 4200 frames per second and camera exposure time was equal to 200 µs.

For the acoustic signal acquisition B&K 4189-A-021 microphone was used, connected into the PXIe 4492 card. Measurement microphone was placed between fast cameras, in front of the observed string vibrating point.

In Fig. 2. the configuration setup is shown. To avoid cameras vibration they were mounted on heavy wood box, and the guitar body was fixed as well. During recordings the string was plucked manually.

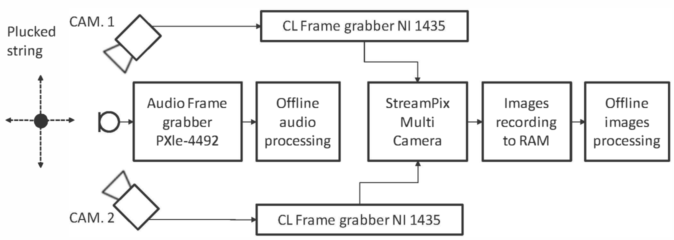

Fig. 1. The recording setup of the string vibrations

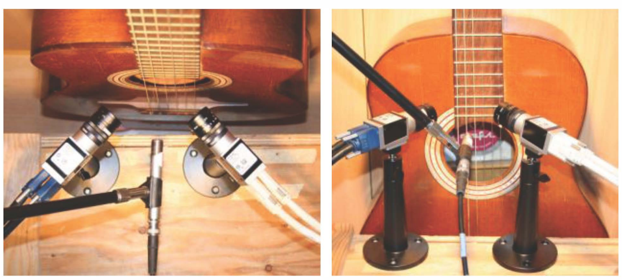

Fig. 2. The overview of recording setup

B. *Image processing algorithm*

Video recording from each camera is processed to estimate location of the observed string element, expressed in pixels. It is assumed that the camera does not introduce non-linear distortions, and each observed part is uniformly projected onto image pixels, with a scale ratio of 62 pixels per 1 mm. Therefore, with an unknown but uniform ratio the displacement in pixels can be transformed into local sound pressure change, and thus as a sound sample value.

The string segment displacement is determined by the following procedure: color to grayscale image conversion, image binarization by adaptive thresholding, image vertical projection, string segment localization and center value calculation, and fusion of results from two cameras. These phases are described and discussed below.

Color to grayscale conversion

The acquired image is a raw Bayer image [3] and initially conversion to RGB is applied using Gradient-corrected bilinear interpolation algorithm with filter pattern 'GBRG' [8]. The converted image is RGB, 8 bits per channel. As the sought object bright color is distinct in the image a color information can be reduced to a grayscale, by applying for every pixel component transformation (1):

$$k = 0.5 \cdot (-r + g + b) \qquad (1)$$

where:

r, g, b – red, green, blue channel values respectively,

k – result gray value.

Subtracting the red component (copper bronze wound red color) from the image results in increased separation of the visible light glint on the string segment from other elements (Fig. 3). The 0.5 factor in (1) is used for normalization only, as the expected maximum value of $(-r+g+b)$ is $(-0+255+255)$ for 8-bit per channel, thus the $0 \leq k \leq 255$.

Fig. 3. Comparison of light glint visibility and contrast in: a) original color image, b) standard grayscale image (brightest area value is 189), c) proposed conversion method (brightest areas value is 242)

Image binarization by multilevel thresholding

The 8-bit grayscale image still contains noise and not precise shape of the object of interest. Binarization procedure is aimed at zeroing pixels determined as background and setting foreground pixels values to 1 (Fig. 4). Various methods are present in the literature, and an adaptive thresholding is proposed in this work: Otsu method based on image histogram analysis [12].

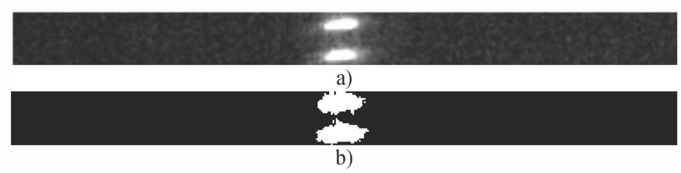

Fig. 4. Sample results of image binarization: a) original image, b) result of Otsu thresholding

This method assumes a histogram of pixel values is bimodal: two distinct classes of foreground and background pixels are observed. Then, the threshold is adjusted for maximum separability of classes. The best threshold provides minimum intra-class variance and maximum inter-class variance.

Image vertical projection and filtration

Elliptical shapes of light glints on the string wound are analyzed only by means of location, because the string element is considered as ideally rigid, thus the size and eccentricity are neglected. A vertical projection, i.e. sum of pixels in every column, was used to determine the location of the object (2).

$$p(i)=\Sigma_j I(i,j) \qquad (2)$$

where:

$I(i,j)$ – image pixel at (i,j) coordinates

$p(i)$ – vertical projection result

At this stage still some visual noise was expected, therefore a moving average of length 15 was applied to $p(i)$ to reduce the impact of noise on the analyzed shape (Fig. 5). Finally a threshold of 0.9 of $max(p(i))$ was used, to determine columns containing definitive parts of the object.

String segment localization and center value calculation

The first and the last columns were taken as the object left and right bound, and their mean value as the object center location (Fig. 6). The obtained result is a one-dimensional signal, representing the movement of the string segment in one direction only (Fig. 7).

Results fusion

Two synchronized cameras were used in the recordings. Obtained displacements were measured along planes parallel to images planes. The fusion of both results allowed acquiring two dimensional movement of the string segment.

Based on the assumption of cameras axes orthogonality, a local coordinate system was proposed, with x axis consistent with x in the first camera, and y axis with x on the second camera (Fig. 8). Therefore by superposition of x and y a top view on the string movement was obtained (Fig. 9).

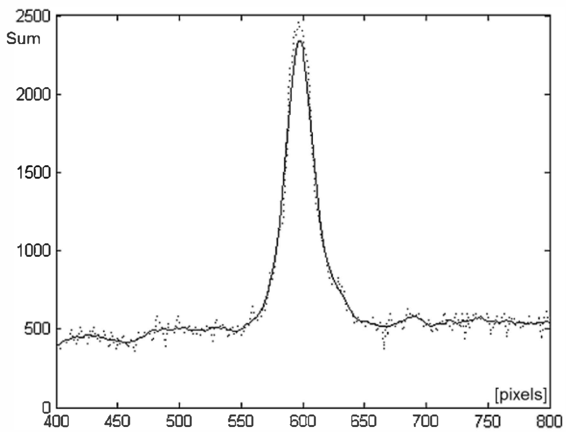

Fig. 5. Sample vertical projection (dotted) and smoothed profile obtained by moving average (line)

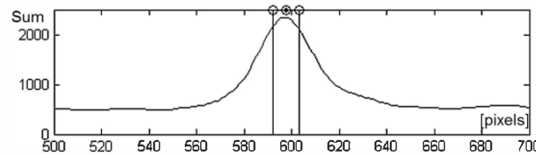

Fig. 6. Result string localization: vertical lines denote area of values exceeding the $p_{max}\cdot$ 0.9 threshold, and the final location is middle point

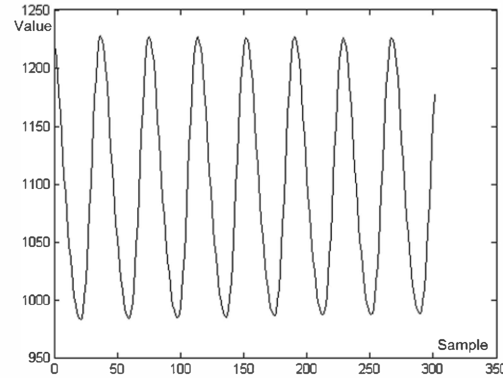

Fig. 7. Samples 16400:16700 from recording of the A string, $x(t)$ signal, left camera

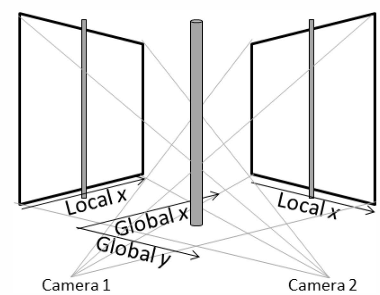

Fig. 8. The concept of local coordinate system of two orthogonal cameras and global coordinate system

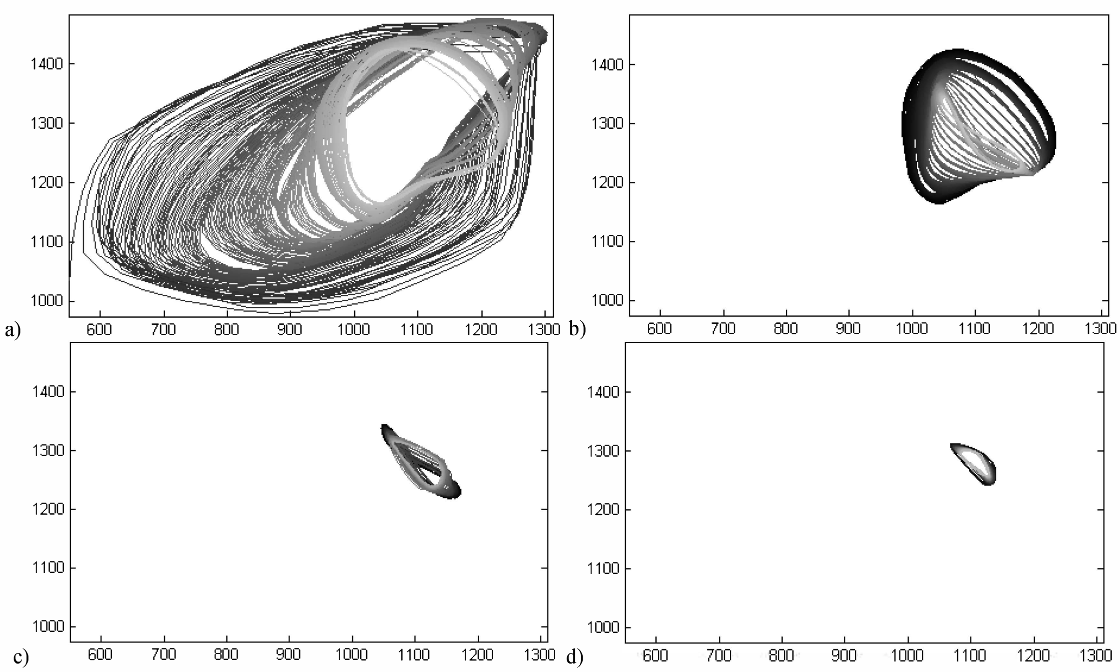

Fig. 9. Phases of the registered motion: a) samples 0:1k, b) samples 1k:2k, c) samples 2k:3k, d) samples 3k:4k. Older samples are represented by darker color, newer by brighter color. Axes scaled in pixels

The real sound perceived at any point in space is a result of air pressure disturbances. For simplification it was assumed here, that the moving string directly influences air compression and rarefaction, and pressure changes depend on changes of distance between the string element and point of reception. The Euclidean distance between each position x(t), y(t) over time and the receiver (x_{rec},y_{rec}) (Fig. 10) was calculated as:

$$r(t)=[\ (x_{rec}-x(t))^2 + (y_{rec}-y(t))^2\ ]^{\ 0.5} \tag{3}$$

The process involving receiver point was a simplified approach mimicking the principle of operation of pickup in electric (or electro-acoustic) guitars, where the movement of steel string in the electromagnetic field disturbs the current flow in the pickup, what is directly interpreted as a sound signal.

Receiver location has partial influence on the obtained signal. Particular trajectory observed from various points result in slightly changed relations of the recreated harmonics (Fig. 11).

The obtained $r(t)$ is normalized (Fig. 12), to eliminate bias and limit the signal to a range <–1; 1>:

$$r_n(t)=[\ r(t) - \dot{r}\ ]\ /\ [\ \max(r(t)) - \dot{r}\ ] \tag{4}$$

where:

$\dot{r}$ – mean value of the signal $r(t)$.

III. Sound Analysis Results

The video recording was made with 4200 frames per second, thus resulting in the same number of sound samples per second. Based on the Nyquist-Shannon sampling theorem the maximum sound frequency component reconstructed correctly is 0.5·4200=2100 [Hz]. All higher frequencies are distorted due to aliasing. For sound recordings an antialiasing filter is applied, preceding the sampling. Similar approach to limit the harmonics in visual sampling during the observation of string vibration is not feasible, therefore every component higher than Nyquist frequency is aliased. The recorded guitar A string is 110Hz, thus, up to 2090Hz (19th harmonic) is correctly recreated.

Acoustic signal recorded by the measurement microphone was compared with the signal obtained using described video processing algorithm. For this purpose the frequency analysis was performed. Both signals were synchronized in the time domain, aligning signals beginnings. In Fig. 13-15 three different parts of both signals are shown. It can be observed that for the first part the first harmonic (110 Hz) is dominant (Fig. 13). In the second part of the signal (Fig. 14) in microphone sound the second harmonic (220 Hz) starts to dominate (the period of the top curve is two times shorter than for bottom curve). Another difference is related to release phase (Fig. 15), as the amplitude of video-based signal decreases slower than for microphone signal.

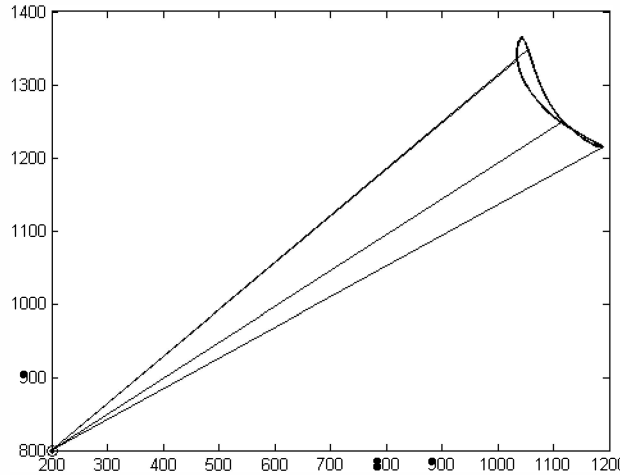

Fig. 10. Superposition of signals x and y interpreted as top view on the string movement. Sample observation point located at (200,800) and measured distances to string localization are illustrated

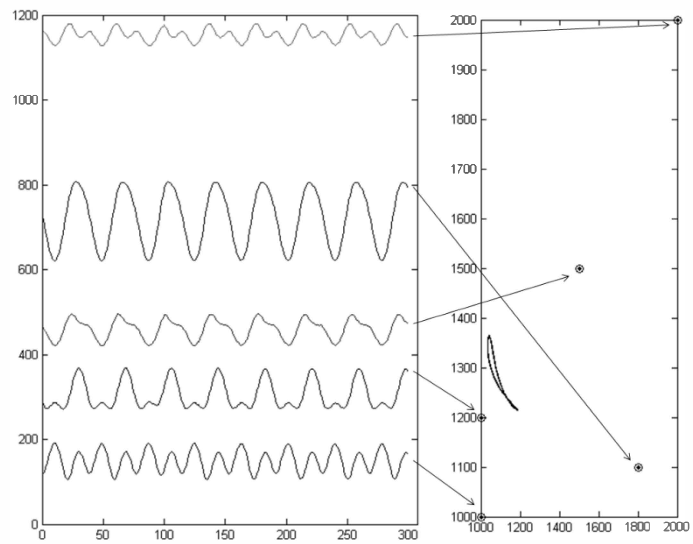

Fig. 11. Result signals (left) obtained by calculating distances from testing receiver points (right) to part of the trajectory (samples 16400:16700)

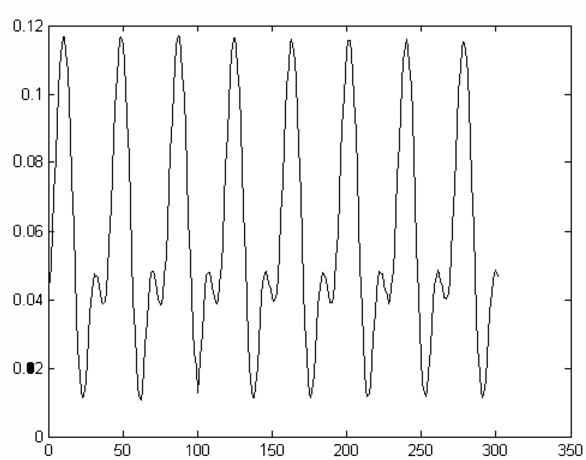

Fig. 12. Samples 16400:16700 from result sound $r(t)$ of the A string, for receiver (200,800)

Additional information about differences between considered signals were found using analysis performed in frequency domain. For this purpose average spectra for both signals were calculated (Fig. 16). Frequency analysis resolution was 1024 FFT points, Hanning window was applied, 30000 Sa of signals were analyzed, sampling frequency of signals was equal to 4200 Sa/s. The original sampling frequency for microphone signal was 48000 Sa/s, therefore it was resampled down to 4200 for comparison.

Level of f_0=110Hz from video analysis was selected as the reference and real sound spectrum was calibrated accordingly to match this level.

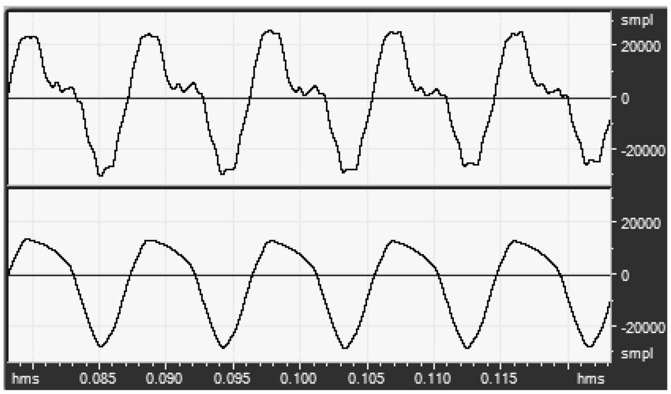

Fig. 13. Acoustic signal recorded by the measurement microphone (top) and the signal obtained using described video processing algorithm (bottom) – beginning of the signal

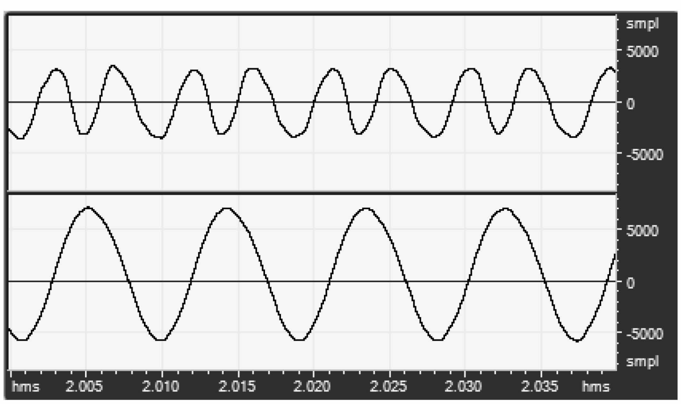

Fig. 14. Second harmonic became dominant in acoustic signal recorded by the measurement microphone – middle part of the signal, t=2 seconds

Higher harmonics for real sound are relatively greater than for sound calculated on the basis on image processing, nevertheless up to 12 frequency components can be observed, matching the ones present in real acoustic signal. Several reasons of such fact can be indicated. First, different means of acquisition produced different results, as the microphone registers acoustic pressure evoked by the whole vibrating string together with the response of the instrument body, thus spectrum is very wide and contains many harmonics. For the sound reconstructed from videos a movement of only single point of vibrating string is observed, and the response of the instrument body is neglected. The “source” signal which causes the instrument body vibrations is observed rather than the “result”. Lastly, assumed receiver point has partial influence on harmonics relations. These are main reasons of differences between considered signals.

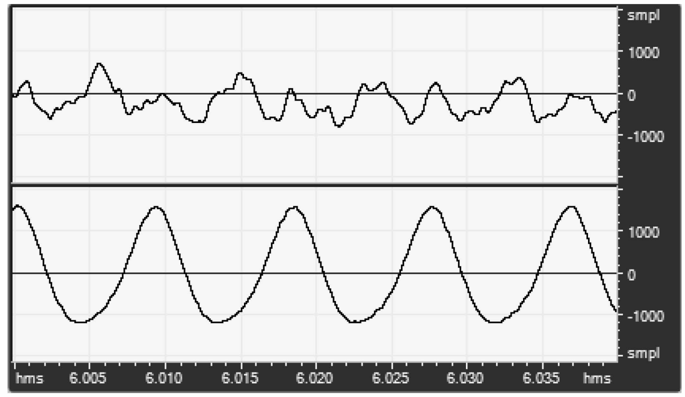

Fig. 15. Amplitude of signal obtained from video frames decreases slower than for signal from microphone – last part of the signal, t=6 seconds

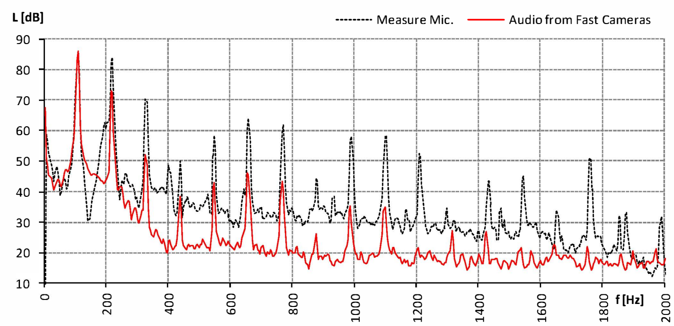

Fig. 16. Average spectra for sound recorded by the measurement microphone (black line) and calculated on the basis on image processing (red line)

IV. Conclusion and Future Work

A hardware and software solution for guitar string vibration measurement by fast cameras was described in the paper. On the basis of performed video and audio recordings comparison of captured string vibrations with the sound recorded using a microphone was successfully performed both in time and frequency domain. The conducted research revealed complexity of acoustic and mechanic phenomena of guitar sound.

In the future work Authors will focus on diagnosing differences between particular harmonics amplitudes for both kinds of signals. This work will involve comparison of several types of string instruments.

Acknowledgements

The project was funded by the National Science Centre allocated on the basis of the decision number DEC-2012/05/B/ST7/02151

REFERENCES

[1] Achkire Y., Preumont A., *Optical measurement of cable and string vibration*. Shock and Vibration 5(3) (1998) 171-179

[2] Askenfelt A., Jansson E. V., *From touch to string vibration. III: String motion and spectra*. J. Acoust. Soc. Am. 93(4) (1993) 2181-2196

[3] Bayer B. E., Color imaging array, US Patent No. 3971065

[4] Czajkowska M., *Analysis of Classical Guitars' Vibrational Behavior based on Scanning Laser Vibrometer Measurements*, AIP Conf. Proc. 1457, 336, 27–29 June 2012, Ancona, Italy, 2012

[5] Czyżewski A., Szwoch G., Dalka P., Szczuko P., Ciarkowski A., Ellwart D., Merta T., Łopatka K., Kulasek Ł., Wolski J., *Multi-stage video analysis framework*, In: Weiyao Lin (Ed.) Video Surveillance, 145 – 171, ISBN 978-953-307-436-8, Intech 2011

[6] Dalka P., Szwoch G., Szczuko P., Czyżewski A., *Video Content Analysis in the Urban Area Telemonitoring System*, In: G.A. Tsihrintzis et al. (Eds.): Multimedia Services in Intelligent Environments, 241 – 261, Springer-Verlag Berlin Heidelberg 2010

[7] Kartofelev D., Mustonen M., Stulov A., Valimaki V., *Application of high-speed line scan camera for string vibration measurements*. In: Proceedings of International Symposium on Musical Acoustics (ISMA 2014), Le Mans, France (2014) [1-6]

[8] Leroy N., Flety E., Bevilacqua F., *Reflective optical pickup for violin*. In: Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06), Paris, France (2006) 204-207

[9] Malvar H. S., Li wei He, Cutler R., High-quality linear interpolation for demosaicing of bayer-patterned color images. In Proceedings of the IEEE International Conference on Speech, Acoustics, and Signal Processing, 2004

[10] Mustonen M., Kartofelev D., Stulov A., *Application of high-speed line scan camera for acoustic measurements of vibrating objects*. Forum Acusticum, European Acoustics Association 2014

[11] NorPix, *StreamPix*, www.norpix.com/products/streampix/streampix.php

[12] Otsu, N., *A Threshold Selection Method from Gray-Level Histograms*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, 1979, pp. 62-66.

[13] Paiva R. C. D., Pakarinen J., Valimaki V., *Acoustics and modeling of pickups*. Journal of the Audio Engineering Society 60(10) (2012) 768-782

[14] Pakarinen J.,. Karjalainen M., *An apparatus for measuring string vibration using electric field sensing*. In Proceedings of the Stockholm Music Acoustics Conference (SMAC03), Stockholm, Sweden (2003) 739-742

[15] Plath N., *High-speed camera displacement measurement (HCDM) technique of string vibration*. In Proceedings of the Stockholm Music Acoustics Conference (SMAC 2013), Stockholm, Sweden (2013) 188-192

[16] Podlesak M.. Lee A. R., *A photovoltaic detector for string vibration measurement*. J. Acoust. Soc. Am. 79(6) (1986) 2092-2093

[17] Szczodrak M., Dalka P., Czyzewski A., *Performance evaluation of video object tracking algorithm in autonomous surveillance system*, Proc. IEEE ICIT 2010, 31-34, 2010

[18] Wolfe, J. *Guitar Acoustics*, The University New South Wales, 2006. http://newt.phys.unsw.edu.au/music/guitar/