

A NOVEL VISUAL REPRESENTATION OF THE SIGNALS IN THE TIME-FREQUENCY DOMAIN

Hristo ZHIVOMIROV ¹

The work presented in this paper describes a novel way for visual representation of the results obtained by Short-Time Fourier Transform (STFT) on a given signal in the time-frequency domain (TFD), named “iris-spectrogram” (irisgram) for its similarity to a human iris.

In the beginning of the paper, a brief introduction to the TFD is made, along with the basics of the STFT-analysis and the classical spectrogram visualization. Further, the newly proposed irisgram is described, and the considerations concerning the attaining of better visual perception of the image are exposed. A few examples are given in order to clarify the irisgram visualization of different kind of signals (both stationary and non-stationary) and to make comparison with the classic spectrogram representation. Finally, conclusions are made including the pros and cons of the newly proposed method of TFD-visualization and its possible applications.

Keywords: time-frequency domain, short-time Fourier transform, spectrogram, visual representation.

1. Introduction

Revealing the time and frequency dependence of a signal can be achieved by performing a time-frequency analysis and plotting the results into the time-frequency domain (TFD), providing a distribution of the signal energy over both time t and frequency f simultaneously. For nonstationary signals, this type of representation is more adequate and easily interpretable by the user [1]. For instance, if one deal with a music signal, the TFD-visualization could be done using a sonogram or even by musical notation.

The time-frequency analysis and representation are used for analysis of signals and data (speech, music, vibrations, time series, etc.), fundamental frequency detection, instantaneous frequency determination, etc.; in the fields of the signal and image processing, data analysis, measurements, acoustics and vibration, machinery diagnosis, seismology, etc.

In the present work, a novel way for visual representation of the results obtained by short-time Fourier transform (STFT) on a given signal is described.

¹ Assist. Prof., Ph.D., Department of Theory of Electrical Engineering and Measurements, Technical University of Varna, Varna, Bulgaria, e-mail: hristo_car@abv.bg

2. Background

A. Short-Time Fourier Transform

Let $x[n]$ be a given real-world (*i.e.*, mathematically real and time limited) discrete signal. Let's one extract signal frames (segments), denoted as $x_l[m]$ at regular intervals using a time-limited window function $w[m]$, expressed as [2]

$$x_l[m] = w[m]x[m + lH], \quad (1)$$

where $m \in \{1, \dots, M\}$ is the local time index (*i.e.*, an index relative to the start of the sliding extraction window), $M \in \mathbb{N}$ is the window length, $l \in \mathbb{N}$ is the frame index, $H \in \mathbb{N}$ is the hop size (*i.e.*, the time advance, expressed in samples, from one signal frame to the next). Further, a Discrete Fourier Transform (DFT) is perform on every frame $x_l[m]$, given a local two-sided spectrum [2]

$$\begin{aligned} \dot{X}[k, l] &\triangleq \frac{1}{M} \sum_{m=1}^K x_l[m] e^{-j2\pi \frac{mk}{K}} = \\ &= \frac{1}{M} \sum_{m=1}^K w[m] x[m + lH] e^{-j2\pi \frac{mk}{K}}, \end{aligned} \quad (2)$$

where $k \in \{1, \dots, K\}$ is the frequency bin index and $K \in \mathbb{N}$ is the DFT size; usually $K = M$ (for $K > M$ the sequence $x_l[m]$ is zero padded and the spectrum is interpolated).

The term $\dot{X}[k, l]$ is called STFT of $x[n]$ and corresponds to the local time-frequency behavior of the signal around the time index lH and the frequency bin k . If the sampling frequency is denoted as f_s , the above indices correspond to continuous time $t = lH/f_s$ and linear frequency $f = kf_s/K$ [2].

B. Classic Spectrogram Visualization

The STFT $\dot{X}[k, l]$ is always complex-valued sequence (for real-world signals), containing information about the localized amplitude and phase spectrum of the accessible time segments of the signal $x[n]$. For most of the applications, the phase spectrum is not significant, so one is interested only in the amplitude spectrum. For visualization purposes the so-called spectrogram is used, which in this paper is defined as

$$\text{Spectrogram} \triangleq \frac{2}{\sum_{m=1}^M w[m]} |\dot{X}[k, l]|, k \in \left\{1, \dots, \left\lceil \frac{K+1}{2} \right\rceil\right\}. \quad (3)$$

The correction factor in Eq. (3) is used so that the spectrogram to be in form of one-sided amplitude spectrum [3].

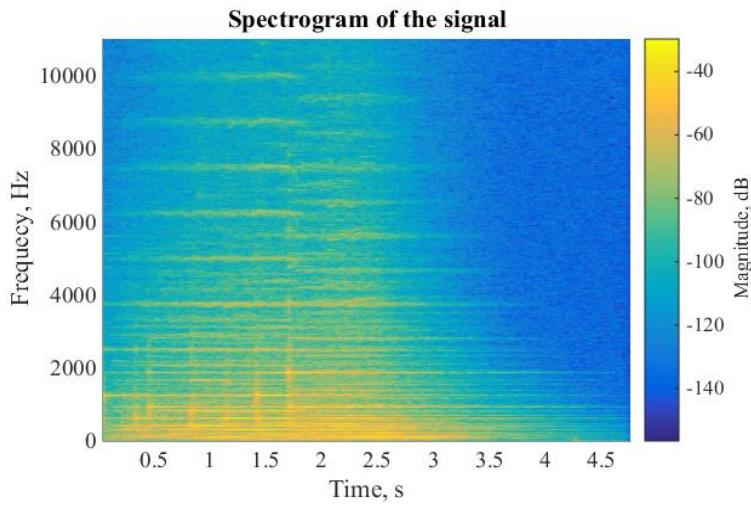


Fig. 1. An example of time-frequency representation of a signal using classic spectrogram.

One usually plots the spectrogram as a rectangular plot – Fig. 1, where the horizontal axis represents time and the vertical axis – frequency. The spectral magnitude usually is given in dB with color, which helps to visualize a wide range of signal values (in the vicinity of 10^6 times).

3. Proposition Description

A. Novel Irisgram Visualization

The author proposes a novel way for visual representation of the results obtained by time-frequency analysis named “iris-spectrogram” or “irisgram” for its similarity to a human iris – Fig. 2.

The irisgram is a circular plot of the classic spectrogram where the time increases azimuthally (circumferentially) clockwise, the frequency increases radially and the signal level is given axially with a color.



Fig. 2. A human iris (left) and an iris spectrogram (right). Note the similarity of the two images giving the name of the irisgram.

One can think about the irisgram like an ordinary spectrogram plot that is bent in a way so that its right side touches the left one. This action transforms the rectangular spectrogram plot into a circular one, so the instances $t=T$ and $t=0$ become merged (cf. Fig. 3). This allows a better understanding of the concept that all digital signals (even nonstationary ones) have a periodic nature (with period equal to the duration of the signal).

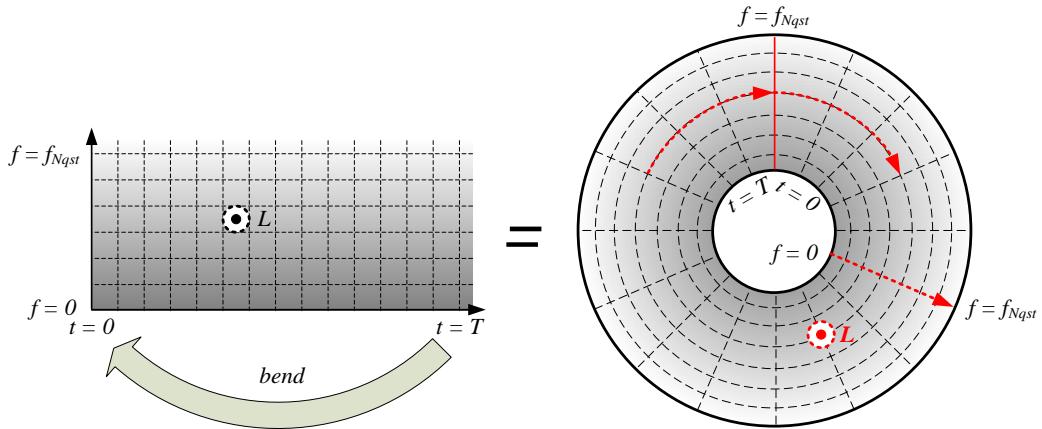


Fig. 3. The irisgram is a circular plot of the spectrogram where the time increases azimuthally (circumferentially) clockwise, the frequency increases radially and the signal level is given axially with color.

The way of arrangement of the plot is based on mnemonics considerations. The time-flow in the irisgram follows the well-known movement of the clock hands. Similarly, there is an association between the rays of light radiating from the sun and the frequency direction on the irisgram plot.

The circular form of the irisgram do not allow annotations to be placed on the time and frequency axes, so a Data Cursor instrument must be used (cf. Fig. 6) in order to determine the exact location of a given point into the TFD .

B. Improving the Perception

Several measures to improve the visual perception of the irisgram have been taken, including ones oriented to people with vision impairment.

On the first place, the dynamic range of the visualized spectrum is limited to 120 dB (corresponding to 20-bit ADC), which is well enough to represent the reasonable difference in the spectral magnitudes (*e.g.*, from volts down to microvolts). The noise floor is also restricted to at least -120 dB, since the lower levels are physically unrealizable in the common practice (*cf.* Fig. 4a). These arrangements remove the useless visual information and map a smaller range of data values to the color map of the plot.

The color map itself must satisfy the following conditions [4], [5]:

- to represent intuitively the magnitude of the signal without need to refer to the color key too often;
- to represent truly the magnitude differences of the signal, even if they are not large;
- to be well distinguished by users with vision deficiencies (*e.g.*, color blindness);
- to be printer-friendly, since in many cases the analysis must be represented on paper.

To address the above conditions the Matlab[®] monochrome color map “bone” is applied (*cf.* Fig. 4b). The great advantages of the monochrome color maps are well described in [4, 5, 6].

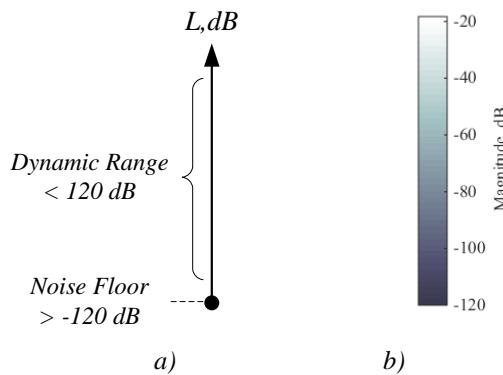


Fig. 4. a) an algorithm is implemented to restrict the visualized dynamic range and noise floor to 120 dB and -120 dB, respectively in order to remove useless information; b) a monochrome color map is used instead of the “rainbow” type, yielding a good, natural perception of the signal level even for color-blind users.

Modification of the color map is done and a hue of the color named “Charcoal” (according to the ISCC-NBS System of Color Designation; arithmetic RGB triplet

[0.22, 0.22, 0.31]) is used as a color corresponding to the lowest level of the signal, instead of the default black color (arithmetic RGB triplet [0, 0, 0]). This gives more brightness of the image, so it became more printer-friendly (*cf.* Fig. 5).



Fig. 5. The colors corresponding to the lowest signal level of the Matlab® color map ‘bone’ (left) and of the modified color map used in the irisgram (right). The last gives more brightness of the image and hence is more printer-friendly.

Further, the specific look of the irisgram allows humans with macular degeneration (or other diseases causing loss of central vision) to have a complete look at the whole time-frequency pattern if they focus their vision into the ‘pupil’ (the central white circle) of the irisgram.

4. Examples

Computer simulations were performed in the Matlab® software environment in order to compare the classic spectrogram and the novel irisgram visualization of the TFD.

Signals of different types (both stationary and non-stationary) were synthesized with duration $T = 60$ s and length $N = 2.646 \cdot 10^6$ samples at sampling frequency $f_s = 44100$ Hz. The mathematical descriptions of the generated signals are listed in Table 1. The signals TS1–TS4 were additionally mixed with Gaussian white noise WN (*cf.* Tab. 1). Also, two real-world signals were analyzed – samples of human speech and music.

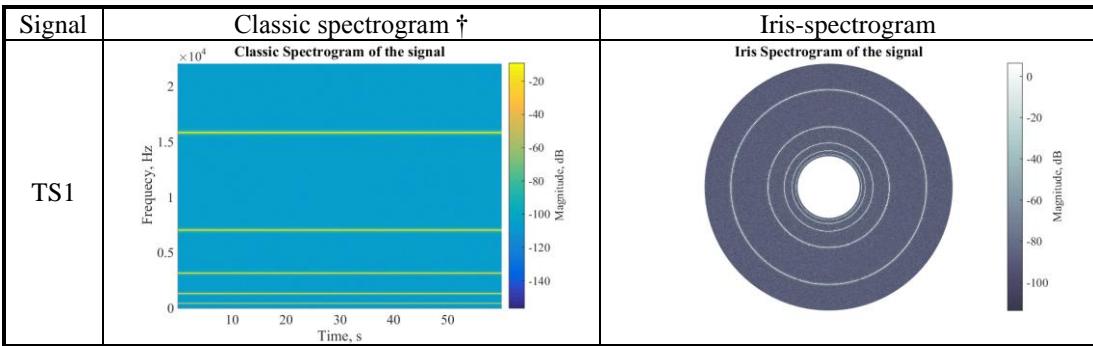
The STFT analysis is performed using the Blackman window [7] with length $win = 2048$, overlapping $ovrlp = \frac{3}{4} \cdot win$, and number of FFT points $nfft = 2 \cdot win$. The analysis results are shown in Table 2.

The classic spectrogram is plotted using the in-build Matlab® function *spectrogram* [7], which computes the PSD of the signal segments, while the irisgram is plotted with the author’s function *irisgram*, accessed in the Matlab® Central File Exchange [10]. The last computes the amplitude spectrum of the analyzed segments, according to Eq. (3).

Table 1
Mathematical description of the test signals used in the examples

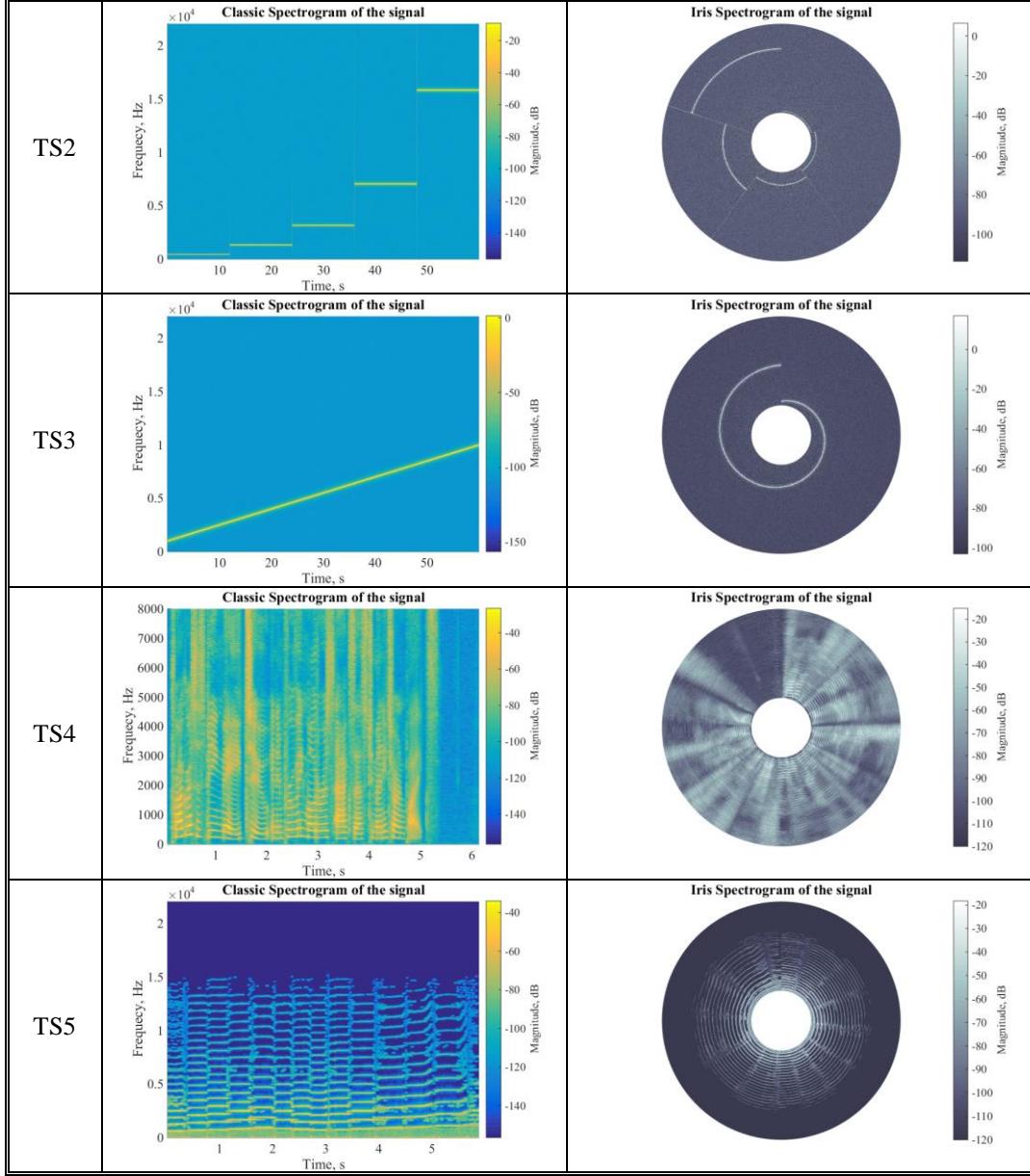
Signal	Signal type	Signal description
TS1	Sum of sine-waves (stationary) $x[n] = \sum_k U_{mk} \sin(2\pi f_k t)$	$x[n] = 0.1 \sin(2\pi 440t) + 0.5 \sin(2\pi 1320t) + 1.0 \sin(2\pi 3168t) + 2.0 \sin(2\pi 7040t) + 3.0 \sin(2\pi 15840t)$
TS2	Sequence of sine-waves (non-stationary) $x[n] = \begin{cases} U_{m1} \sin(2\pi f_1 t), & 0 \leq t < t_1 \\ U_{m2} \sin(2\pi f_2 t), & t_1 \leq t < t_2 \\ \dots \end{cases}$	$x[n] = \begin{cases} 0.1 \sin(2\pi 440t), & 0 \leq t < 12 \\ 0.5 \sin(2\pi 1320t), & 12 \leq t < 24 \\ 1.0 \sin(2\pi 3168t), & 24 \leq t < 36 \\ 2.0 \sin(2\pi 7040t), & 36 \leq t < 48 \\ 3.0 \sin(2\pi 15840t), & 48 \leq t < 60 \end{cases}$
TS3	Sweep signal (linear chirp) $x[n] = U_m \sin\left(2\pi f_1 t + \pi \frac{f_2}{T} t^2\right)$	$x[n] = 10 \sin\left(2\pi 1000t + \pi \frac{10000}{60} t^2\right)$
TS4	Human speech	Record "DR2_FRAM1_SI522" from the TIMIT database [8]
TS5	Music	Violin sound sample [9]
WN	Additive white Gaussian noise $\varepsilon[n] = \mathcal{N}(0, \sigma^2)$	$\varepsilon[n] = \mathcal{N}(0, 10^{-6})$

Table 2
Results from the Matlab visualizations of the classic and iris-spectrograms



\dagger – the magnitudes of the classic spectrogram and the iris-spectrogram differ since the first one represents the power spectrum density, while the second – the amplitude spectrum of the signal.

Results from the Matlab visualizations of the classic and iris-spectrograms (cont.)



A zoomed portion of the irisgram of the signal TS5 – violin sound sample, is shown in Fig. 6. A few data points are marked via data tips using the Matlab[®] Data Cursor instrument. Three different draws of the bow and the differences in the tones can be distinguished on the plot. The first area, denoted as A in Fig. 6, containing a three data points representing a reference tone ~ 700 Hz and two

overtones (harmonics) at ~ 1400 Hz and ~ 2100 Hz. The next two areas – *B* and *C* are given for indicative purposes – better understanding of the irisgram conception. They show a pair of data tips that could be interpret as a time-localized spectrum (points at one and the same time instance) at time ~ 0.91 s and a time-shifting filter tuned at ~ 732 Hz (points with one and the same frequency), respectively.

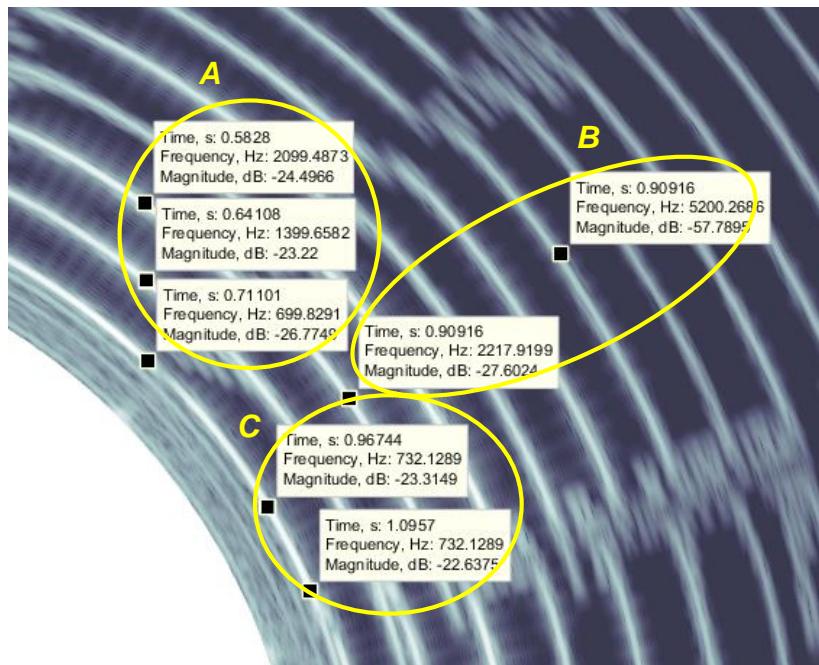


Fig. 6. The irisgram of the signal TS5 – violin sound sample, zoomed and marked with data tips.

5. Conclusions

A novel way for TFD data visualization is presented in the paper, named “irisgram” – a circular plot where the time increases azimuthally (circumferentially) clockwise, the frequency increases radially and the signal level is given axially, coded with a color.

The irisgram visualization allows:

- better perception of the plot by users with vision impairment (*e.g.*, color blindness or macular degeneration);
- performing a kind of “irisodiagnosis” of the signal, examining specific patterns and colors of the “signal iris” to determine information about the signal itself, if the personnel are trained enough to work with the irisgram.

The accompanying monochrome color map contributes better reproduction on paper, when black and white printing is used.

Besides, there are two possible inconveniences that could arise when one uses the irisgram – the need of a proper Data Cursor instrument and the possible psychological inertia, when the user is too accustomed with the classical spectrogram visualization. Nevertheless, the irisgram representation is easy to be apprehending, since the way of arrangement of the plot is intuitive enough.

The possible application of this contribution is where the visual representation of the results obtained by STFT is a need. The corresponding Matlab®-function is accessible at [10].

R E F E R E N C E S

- [1]. *B. Boashash*, Time Frequency Signal Analysis and Processing: A Comprehensive Reference, Oxford, Elsevier, 2015.
- [2]. *J. Benesty, M. Sondhi, Y. Huang*, Springer Handbook of Speech Processing, Berlin, Springer, 2008.
- [3]. *G. Heinzel, A. Rüdiger, R. Schilling*, Spectrum and spectral density estimation by the Discrete Fourier transform (DFT), including a comprehensive list of window functions and some new flat-top windows, Germany, Hannover, Max-Planck-Institut für Gravitationsphysik, 2002.
- [4]. *S. Eddins*, Rainbow Color Map Critiques, Natick, The MathWorks Inc., 2014, Online at: http://www.mathworks.com/tagteam/81137_92238v00_RainbowColorMap_57312.pdf, Last accessed on Mar. 5th, 2018.
- [5]. *K. Moreland*, “Diverging Color Maps for Scientific Visualization”, ISVC '09 Proceedings of the 5th International Symposium on Advances in Visual Computing: Part II, pp. 92-103, 2009, DOI: 10.1007/978-3-642-10520-3_9.
- [6]. *A. Telea*, Data Visualization: Principles and Practice, Boca Raton, CRC Press, 2015.
- [7]. Signal Processing Toolbox Reference, Natick, The MathWorks Inc., 2017, Online at: https://www.mathworks.com/help/pdf_doc/signal/signal_ref.pdf, Last accessed on Mar. 5th, 2018.
- [8]. *J. Garofolo et al.*, TIMIT Acoustic-Phonetic Continuous Speech Corpus LDC93S1, Philadelphia, Linguistic Data Consortium, 1993.
- [9]. Wikimedia Commons, https://commons.wikimedia.org/wiki/File:Spectrogram_of_violin.png, Last accessed on Mar. 5th, 2018.
- [10]. *H. Zhivomirov*, “Spectrogram Visualization with Matlab Implementation, version 1.1”, Online at: <https://www.mathworks.com/matlabcentral/fileexchange/64882-spectrogram-visualization-with-matlab-implementation>, Last accessed on Mar. 5th, 2018.