PAN
ARCHIVES OF ACOUSTICS
Vol. **43**, No. 3, pp. 505–516 (2018)

DOI: 10.24425/123922

A Study on of Music Features Derived from Audio Recordings Examples – a Quantitative Analysis

Aleksandra DOROCHOWICZ[(1)], Bożena KOSTEK[(2)]

[(1)] *Multimedia Systems Department*
Faculty of Electronics, Telecommunications and Informatics
Gdansk University of Technology
Narutowicza 11/12, 80-233 Gdańsk, Poland; e-mail: aleksandra@multimed.org

[(2)] *Audio Acoustics Laboratory*
Faculty of Electronics, Telecommunications and Informatics
Gdansk University of Technology
Narutowicza 11/12, 80-233 Gdańsk, Poland; e-mail: bokostek@audioakustyka.org

(*received June 19, 2017; accepted March 21, 2018*)

The paper presents a comparative study of music features derived from audio recordings, i.e. the same music pieces but representing different music genres, excerpts performed by different musicians, and songs performed by a musician, whose style evolved over time. Firstly, the origin and the background of the division of music genres were shortly presented. Then, several objective parameters of an audio signal were recalled that have an easy interpretation in the context of perceptual relevance. Within the study parameter values were extracted from music excerpts, gathered and compared to determine to what extent they are similar within the songs of the same performer or samples representing the same piece.

Keywords: music genres; audio parametrization; music features.

1. Introduction

There are many formal systems describing differences between music pieces, which allow for defining its assignment to a category or type of music. Many features influence the final categorization of the song, some of them can be described or measured as is done within the Automatic music genre classification (AMGC) area (Bergstra *et al.*, 2006; Bhalke, 2017; Burred, Lerch, 2014; Hoffmann, Kostek, 2015; Kalliris *et al.*, 2016; Kotsakis *et al.*, 2012; Kostek, 2005; Ntalampiras, 2013; Schedl *et al.*, 2014; Silla *et al.*, 2007; Sturm, 2014; Tzanetakis *et al.*, 2002; Rosner *et al.*, 2014; Rosner, Kostek, 2018). One of the ways the songs can be described is by using their origin (J-rock, Brit pop), time when they are composed or performed, performance (Pluta *et al.*, 2017), mood of music (Barthet *et al.*, 2017; Plewa, Kostek, 2015), instruments used (symphonic, acoustic, rock), music techniques (riff, rap) or function that music plays (film score, religious or orchestral music) (Benward, 2003). Frequently, to be able to conduct a proper analysis of a piece, one may need to take into account a number of feature types at the same time. Music genres are also divided into smaller sub-groups (punk rock, new wave, post grunge) and they can be mixed (symphonic metal, pop rock).

Throughout the history, many types of classifications were created. They were based on music styles, forms and genres. Over time, the topologies become more and more complex and divided. Moreover, the pieces could belong to more than one category, so it is necessary to identify their definitions and significance. These include musical form, style, genre and descriptive features such as meter, tempo, structure or origin. The definition of the music form is based on the vocal and/or music instruments used, texture type or number of passages it consists of (one-passaged, cyclic). Moreover, the *Small Music Encyclopedia* describes music forms as the typical for a specific group of music piece schemes, designated by the analysis based on the specific pieces.

At the beginning of 17th century, the concept of the style in the music theory changed its meaning (Pascall, 2001; Seidel, Leisinger, 1998). Division based on antique tradition: religious, chamber and scenic (theatrical), within subtypes were proposed by Marc Scacchi in 1649 (Palisca, 1998). Athanasius Kircher proposed to expand it into church, madrigal and theatrical style, noting that it is not the place of performance that specifies the style, but the effects triggered. His proposal also predicts styles diffusions (Helman, 2016). Although the definition of the music style has changed through the centuries, according to the *Small Music Encyclopedia*, the music style is the term describing the common features of the compositional technique typical for the specific piece, author, for the national music, historical period (Dziębowska, 1998).

A music genre identifies some pieces of music as belonging to a shared tradition or set of conventions. Music genres are created based on many types of divisions: age of the recipients (e.g. adult contemporary, teen pop, music for children), artists' origin (e.g. Americana, Afro-Cuban, Brit pop, Italo disco, K-pop, etc.), time of the origin (e.g. baroque, classical, contemporary, etc.), artists'/music ambitions (classical music, ambient, country, etc.), ideology (e.g. rock, hip-hop, rap, metal, blues, New Age, etc.), instrumentation and treatment of musical instruments within the given genre (e.g. jazz, country, folk, electronic, blues, rock, etc.).

The American music magazines *Alternative Press* (Alternative Press, 2016) and *Rock Sound* (Rocksound, 2016) at the beginning of the 2016 drew attention to the problem of lack of the unequivocal music genres definition and divisions in the research study entitled "What is punk?". It was carried out by the Converse company in associate with the Polygraph's analyst (Daniels, 2016). Their bases were playlists tagged as punk at Spotify and YouTube's playlists. Over half of them (51%) contain such music bands as: Green Day, Blink-182 (50%), The Offspring (44%), Sum 41 (39%), Rise Against (38%), Fall Out Boy (38%), My Chemical Romance (35%), Bad Religion (30%), Nofx (28%), All Time Low (28%) and A Day to Remember (27%). All the bands belong to genres closely related to punk (pop-punk, emo, post-hardcore, metalcore), although they are not strictly punk. This research shows that it is not obvious to which category the performer or the song belongs, which makes defining it in some cases almost impossible.

It should also be noted that within the Music Information Retrieval area there exists a considerable disagreement among the researchers over music genre classification, as assigned by a human or by a machine learning algorithm, i.e. whether such an assignment is practical (Silla *et al.*, 2017; Tekman, Hortacsu, 2002). At the same time there are a lot of examples of research studies, as well as commercial applications (e.g. music social networking, music cataloguing tools for applications, etc.) related to automatic genre recognition and classification (e.g. Bergsta *et al.*, 2006; Tzanetakis *et al.*, 2002; Holzapfel, Stylianou, 2008; Kostek *et al.*, 2011; Kostek, Kaczmarek, 2013; Ntalampiras, 2013), where the user may choose a song belonging to the particular music genre (Tzanetakis *et al.*, 2002).

The aim of the study is to determine to what extent music feature values and characteristics of music excerpts are similar for the performer or for a music piece. Tracks gathered for the purpose of this study are described by their parametric representation, i.e. time and spectral parameters, such as: RMS (Root-Mean-Square) energy, zero-crossing rate, spectral centroid, spectral skewness, spectral kurtosis, spectral flatness, entropy of spectrum, brightness and roll-off, all extracted by the Matlab MIRtoolbox1.6.2 (MIRtoolbox). At the same time several descriptive features such as music genre, time of origin, country are identified. Music excerpts used in the experiment represent various music genres, assigned according to the artist's ambitions, origin, time of origin, they include various pieces of the same artist, and the same songs performed by a number of various artists.

2. Experiments

2.1. Building a database

In the study, 202 music excerpts, 0.5–1 minute long, were collected. They represent various music genres, i.e.: rock, pop-punk, new wave, glam metal, punk rock, Brit pop, pop, soft rock, blues, musical, rock and roll, psychedelic rock, soul, art rock, heavy metal, emo, post grunge, cabaret, pop rock, different time of origin (from 1943 to 2016), different countries and instruments used. Also, the research includes samples of one song performed by various artists and various songs performed by one artist.

Moreover, in our previous study we have performed listening tests to eventually correlate the subjective session outcome with the artificial intelligence algorithms results (Dorochowicz *et al.*, 2017). The task of the test participant was to assign a music excerpt to the specific music genre. The same goal was expected to be achieved by the machine learning approach. Even though the listening test results were not fully homogeneous due to the individual characteristics of music as well as the ambiguity of some tracks, causing that people attributed a given track to different music genres, still the results were more than encouraging (Dorochowicz *et al.*, 2017). We have obtained a good agreement between subjective and "objective" approaches, the latter one based on machine learning. This helps to determine which music excerpts can be considered as uniquely attributed to a particular genre, checked both by subjective tests and machine learning approach. Unambiguous cases obtained in that way were parametrized and further analysed.

2.2. Parametrization

As already mentioned the analysis performed is based on a quantitative comparison of parametric representation of music excerpts. To parametrize music excerpts several features have been chosen, as defined by the Matlab MIRtoolbox 1.6.2 (MIRtoolbox). The criterion for choosing them was a potential physical interpretation and perceptual relevance of a descriptor. They were as follows:

- RMS energy – Root-Mean-Square Energy; a parameter showing the global energy of the signal:

$$x_{\text{rms}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}}. \qquad (1)$$

- Zero-crossing rate – counts how many times the signal crosses the X-axis (Fig. 1), and it is related to the noisiness of the signal.

Also, several spectral distribution-based parameters were used that can be described by statistical moments: centroid, spread, skewness, kurtosis, flatness, as well as entropy:

- Spectral centroid – returns the first moment (mean), which is the geometric center (centroid) (Fig. 2). It is interpreted as a measure of Brightness (center of gravity; see Fig. 3):

$$\mu = \int x f(x)\,\mathrm{d}x. \qquad (2)$$

- Spectral skewness – is the third central moment, showing the asymmetry of the distribution around its mean. When the value is positive, the distribution has a longer tail to the right, the negative value means the opposite. When the value equals zero, the distribution is symmetrical (Fig. 4):

$$\mu_3 = \int (x - \mu_1)^3 f(x)\,\mathrm{d}x. \qquad (3)$$

- Spectral kurtosis – is defined as the fourth standardized moment minus 3 (to correct the kurtosis of the normal distribution equal to zero). Its interpretation is associated with flatness of the spectral distribution around its mean. Examples of the kurtosis figures are shown in Fig. 5.
- Spectral flatness – defined as the ratio of the geometric and arithmetic means of the coefficients

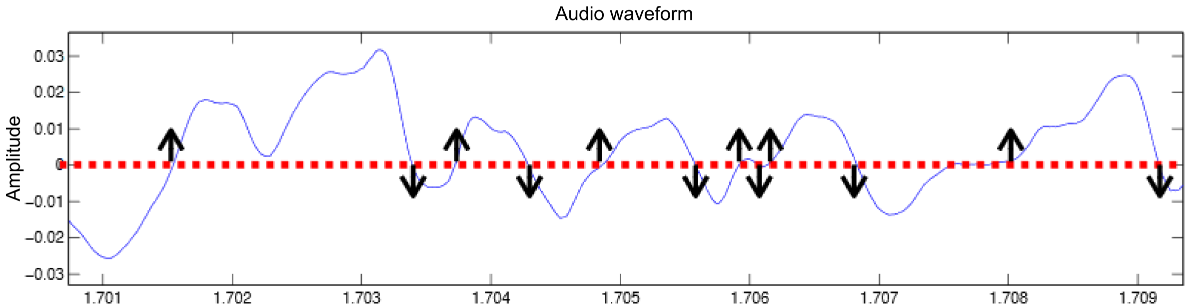

Fig. 1. Signal crossing X-axis (MIRtoolbox).

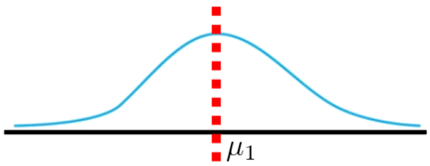

Fig. 2. Centroid (MIRtoolbox).

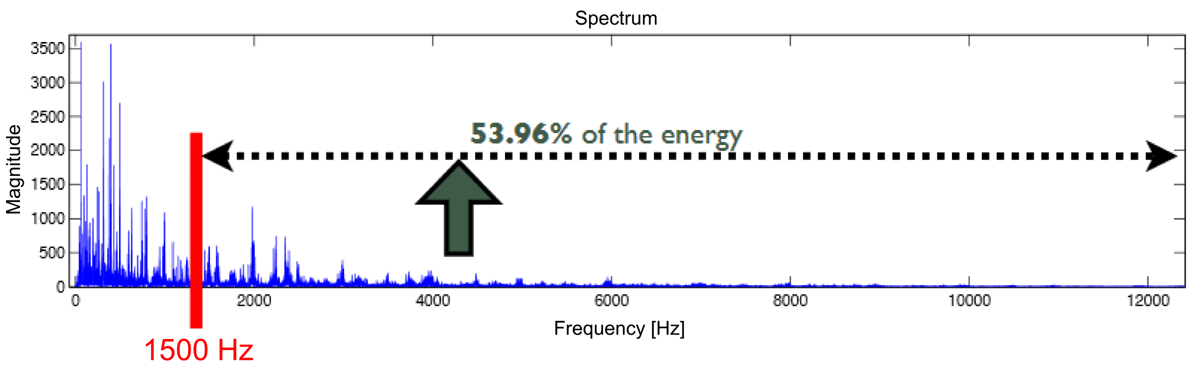

Fig. 3. Brightness (MIRtoolbox).

of the power density spectrum in every spectral bands (b) of the width of 1/4 octave (Eq. (4)). This feature is also called a tonality coefficient (Dubnov, 2004)

$$\mathrm{SFM} = 10\log_{10}\left(\frac{\left[\prod\limits_{k=1}^{N/2} P\left(\mathrm{e}^{j\frac{2\pi k}{N}}\right)\right]^{\frac{1}{N/2}}}{\frac{1}{N/2}\sum\limits_{k=1}^{N/2} P\left(\mathrm{e}^{j\frac{2\pi k}{N}}\right)}\right), \qquad (4)$$

where $P\left(\mathrm{e}^{j\frac{2\pi k}{N}}\right)$ is the PSD calculated on the basis of the N-point DFT.

- Entropy of spectrum – a measure that describes spectrum uniformity. Equation (5) returns the relative Shannon entropy:

$$H(X) = -\sum_{i=1}^{n} p\left(x_i\right)\log_b p\left(x_i\right). \qquad (5)$$

- Roll-off – estimates the amount of high frequency in the signal by finding the frequency below which 85% of the magnitude distribution is concentrated (Fig. 6).

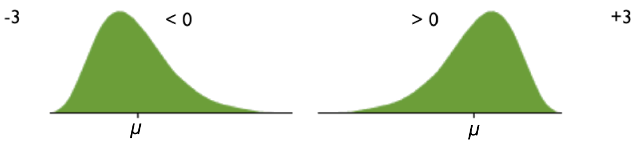

Fig. 4. The negative and positive value of the skewness (MIRtoolbox).

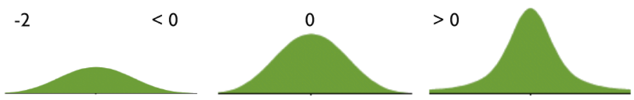

Fig. 5. Different values of the kurtosis (MIRtoolbox).

3. Analyses

Examples of the analysis results obtained are shown below. They concern a quantitative analysis of parameter values extracted from the music excerpts. The music excerpts analysed are divided into the following groups:

- **Artists' songs for whom parameters values are similar** – a thorough analysis of songs that are potentially similar in perception was performed, and among the music excerpts gathered such artists as listed below are to be mentioned:
  - Green Day (see Fig. 7) – punk rock, songs: *Basket Case*, *Blitzkrieg Bop* (cover of Ramones – punk rock), *Like a Rolling Stone* (cover of Bob Dylan – country), *Tired of Waiting* (cover of The Kinks – rock),
  - Elaine Paige (musical singer, songs: *Don't Cry For Me Argentina* (cover of Madonna – pop), *I Dreamed a Dream*, *Memory*),
  - The Blues Brothers – rock'n'roll, *Everybody Needs Somebody to Love*, *Jailhouse Rock* (cover of Elvis Presley – rock'n'roll), *Sweet Home Chicago*;
- **Artists' songs for whom parameter values differ**, examples of them are listed below:
  - Sarah Brightman (see Fig. 8) – musical singer, songs: *Don't Cry For Me Argentina* (cover of Madonna – pop), *Memory* (cover of Elaine Paige – musical singer), *My Heart Will Go On* (cover of Celine Dion – pop), *Time To Say Goodbye*,
  - The Police – new wave, songs: *Roxanne*, *Can't Stand Losing You*, *Spirits in the Material World*,
  - The Calling – post grunge, songs: *London Calling* (cover of The Clash – punk rock),

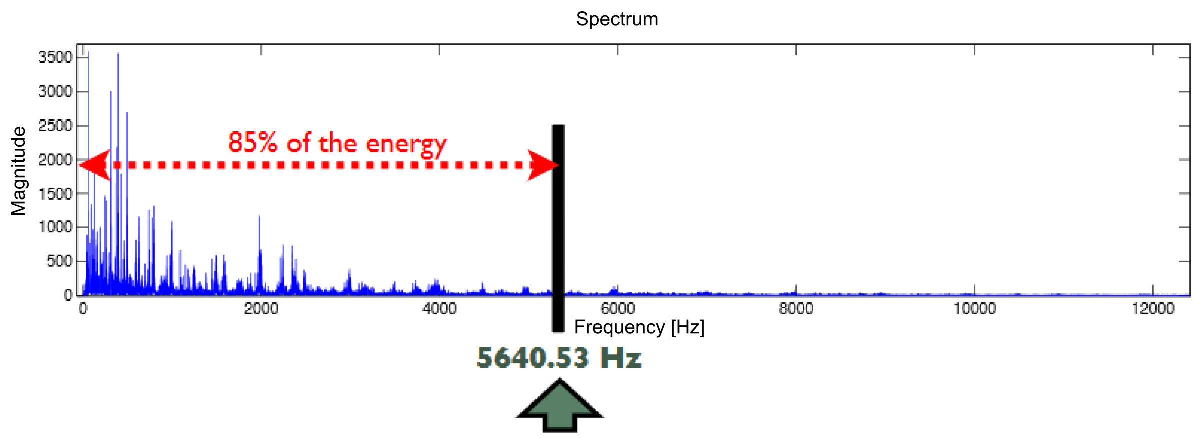

Fig. 6. Roll-off parameter illustration (MIRtoolbox).

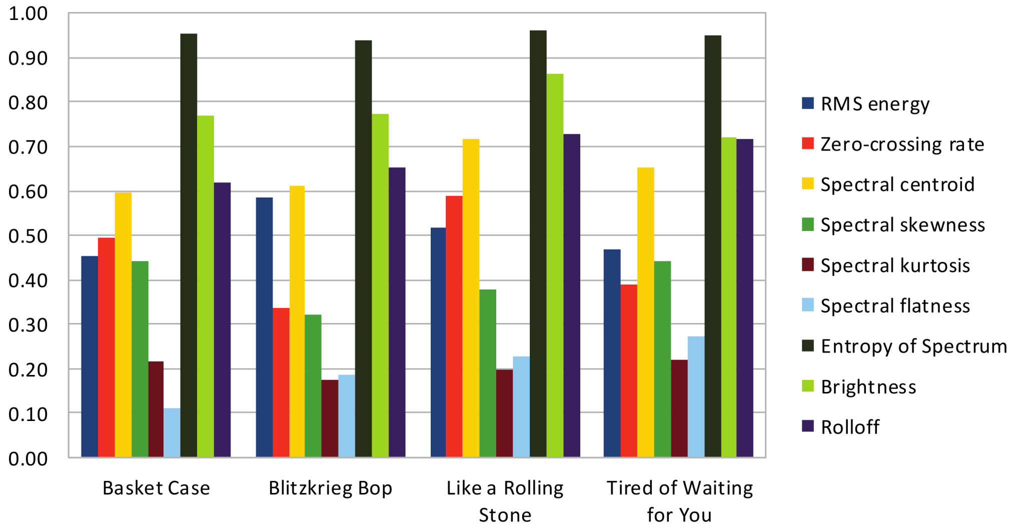

Fig. 7. Parameter values (normalized) similar for music excerpts of a singer (Green Day).

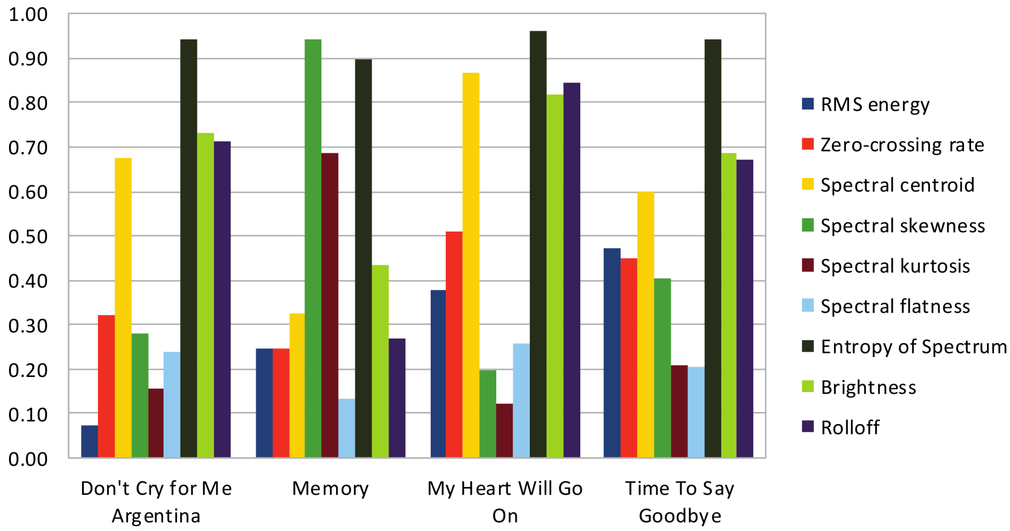

Fig. 8. Parameter values (normalized) different for music excerpts of the same singer (Sarah Brightman).

Our Lives, *Things Will Go My Way*, *When It All Falls Down*, *Wherever You Will Go*;

- **Artists, whose cover and their own songs have different characteristics**, such as:
  - Avril Lavigne – pop punk, songs: *All The Small Things* (cover of Blink-182 – punk rock), *Bad Reputation* (cover of Joan Jett – punk rock), *Basket Case* (cover of Green Day – punk rock), *Girlfriend*, *Here's To Never Growing Up*, *How You Remind Me* (cover of Nickelback – post grunge), *Knockin' On Heaven's Door* (cover of Bob Dylan – country), *Sk8er Boi*, *Stop Standing There*, *The Best Damn Thing*, *What The Hell*; shown in Table 1,
  - My Chemical Romance – emo, songs: *Desolation Row* (cover of Bob Dylan – country), Helena, *I'm Not Okay* (I Promise), *Na Na Na* (Na Na Na Na Na Na Na Na Na), *Song 2* (cover of Blur – Brit pop), *Welcome to the Black Parade*,
  - All Time Low – pop punk, songs: *Bad Reputation* (cover of Joan Jett – punk rock), *Blitzkrieg Bop* (cover of Ramones – punk rock), *For Baltimore*, *Kids in the Dark*, *Missing You*, *Poppin' Champagne*, *Runaways*, *Should I Stay or Should I Go* (cover of The Clash – punk rock);
- **Artists, who have an original song different from others contained in their discography**, such as e.g.:
  - Peggy Lee – jazz, songs: *He's a Tramp*, *I'm a Woman*, *It's a Good Day*, *Why Don't You Do Right* (shown in Fig. 9),

Table 1. Covers different than artist's own songs (Avril Lavigne) (covers are marked in grey colour).

Song	RMS energy	Zero-crossing rate	Spectral centroid	Spectral skewness	Spectral kurtosis	Spectral flatness	Entropy of spectrum	Brightness	Rolloff
Girlfriend	0.62	0.55	0.60	0.44	0.24	0.17	0.95	0.84	0.57
Here's To Never Growing Up	0.81	0.59	0.81	0.39	0.20	0.61	0.97	0.86	0.82
Sk8er Boi	0.64	0.62	0.66	0.49	0.28	0.23	0.97	0.88	0.61
Stop Standing There	0.69	0.57	0.68	0.47	0.25	0.41	0.96	0.85	0.70
The Best Damn Thing	0.83	0.66	0.60	0.45	0.25	0.22	0.96	0.84	0.59
What the Hell	0.79	0.70	0.70	0.52	0.28	0.42	0.96	0.87	0.69
All the Small Things	0.51	0.67	0.78	0.27	0.17	0.21	0.97	0.94	0.71
Bad Reputation	0.48	0.66	0.83	0.28	0.16	0.27	0.97	0.94	0.80
Basket Case	0.43	0.76	0.81	0.30	0.17	0.31	0.98	0.92	0.76
How You Remind Me	0.47	0.52	0.64	0.41	0.21	0.21	0.95	0.82	0.71
Knockin' on Heaven's Door	0.38	0.71	0.95	0.21	0.13	0.31	0.98	0.92	0.91
Standard Deviation	0.05	0.09	0.11	0.08	0.03	0.05	0.01	0.05	0.08

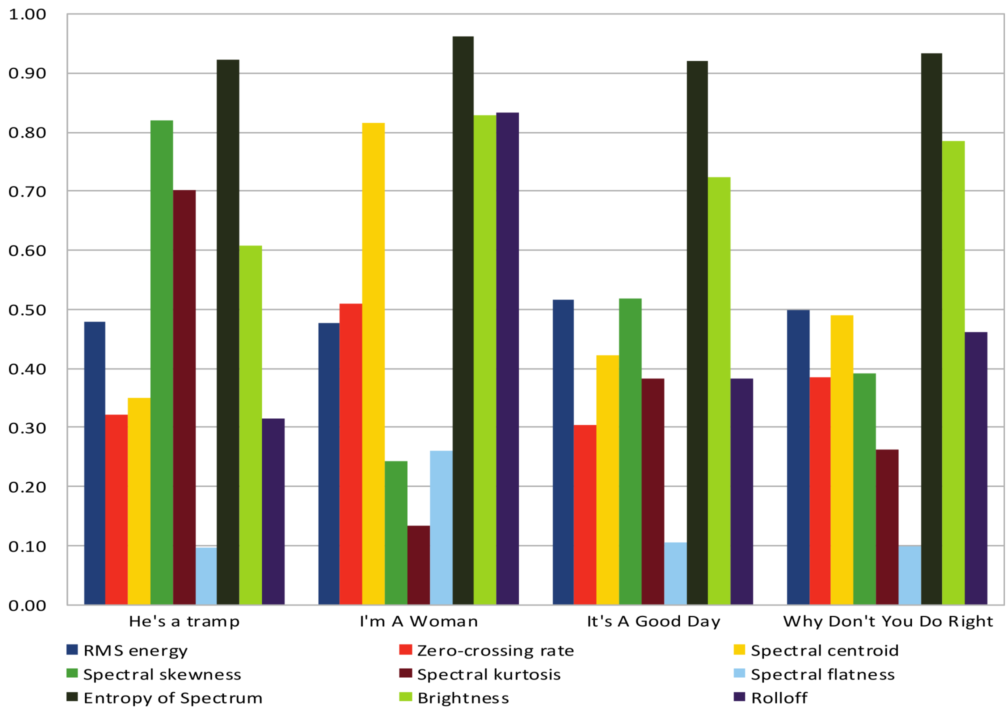

Fig. 9. One different song of Peggy Lee.

– or Madonna – pop: *Don't Cry For Me Argentina*, *Frozen*, *Material Girl*;

- **Artists, whose style evolved over time**, among them there are:

– Offspring – punk rock, songs: *Blitzkrieg Bop* (cover of Ramones – punk rock) (1984), *Pretty Fly* (For a White Guy) (1998), *The Kids Aren't Alright* (1998), *You're Gonna Go Far*, Kid (2008),

– or Meat Loaf (see Fig. 10) – art rock, songs: *I'd Do Anything For Love* (1993, Bat Out of Hell II), *Rock And Roll Dreams Come Through* (1993, Bat Out of Hell II), *It's All Coming Back To Me Now* (2006, Bat Out of Hell III), *Seize The Night* (2006, Bat Out of Hell III), *What About Love* (2006, Bat Out of Hell III).

Although not all the parameters of music samples of one group are very close to each other or may even manifest different tendency, their similarity/dissimilarity can be judged on the basis of the standard deviation. In Fig. 7, one can observe some parameter values that are spread out across the scale. For example, the zero-crossing rate has values from 0.34 up to 0.59, on the other hand, values of spectral kurtosis parameter are within the range of 0.18 to 0.22. The biggest value of the standard deviation is for the zero crossing parameter and is equal to 0.11, whereas the smallest one is nearly neglectful (i.e. 0.0095 for the entropy).

Parameters of music samples in Fig. 8 have in most cases different values, apart from entropy. However, there is no visible pattern of what values do the parameters take. Most of them differ from each other. The largest value of the standard deviation is 0.33 for spectral skewness.

In the case of Table 1 one may discern two groups of results, namely: samples of the songs owned by the artist (the first six songs), and the second one contains covers performed by this artist. Samples of the songs originally performed by the artist have higher RMS energy, lower spectral centroid, higher spectral skewness and kurtosis. Contrarily, covers have in most cases higher zero-crossing range and lower spectral flatness. It is an interesting observation as one may hypothesize that the artist has tried to imitate performance of the covers or did not use the whole potential when performing covers.

For this case a Test of Significance for Two Unknown Means and Unknown Standard Deviations (according to Eq. (6) was performed (Lane, 2017), which showed that for both significance levels, i.e. 5% and 1% (Soper, 2017), for three parameters: RMS-energy, spectral skewness and spectral kurtosis the differences are statistically significant.

$$t = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}, \tag{6}$$

where s_1^2, s_2^2 – variances of the two groups, n_1, n_2 – size of the groups.

Parameters of the first song contained in the chart in Fig. 9 differ as to their values comparing to other songs, this especially concerns Spectral skewness, Spectral kurtosis (both have higher values with the standard deviation approx. equals 0.24) and to some extent – Brightness (lower value). This song was written for different purpose, i.e. it comes from a movie, that is why it affected the artist's performance.

In the case of differences during the period of time (Fig. 10), two groups of the results coming from two albums: one from 1993 and the other from 2006 are presented for this particular artist (Meat Loaf). It can be observed that for the newer songs values are much higher for Spectral flatness, contrarily the older songs have higher Spectral centroid, skewness and kurtosis.

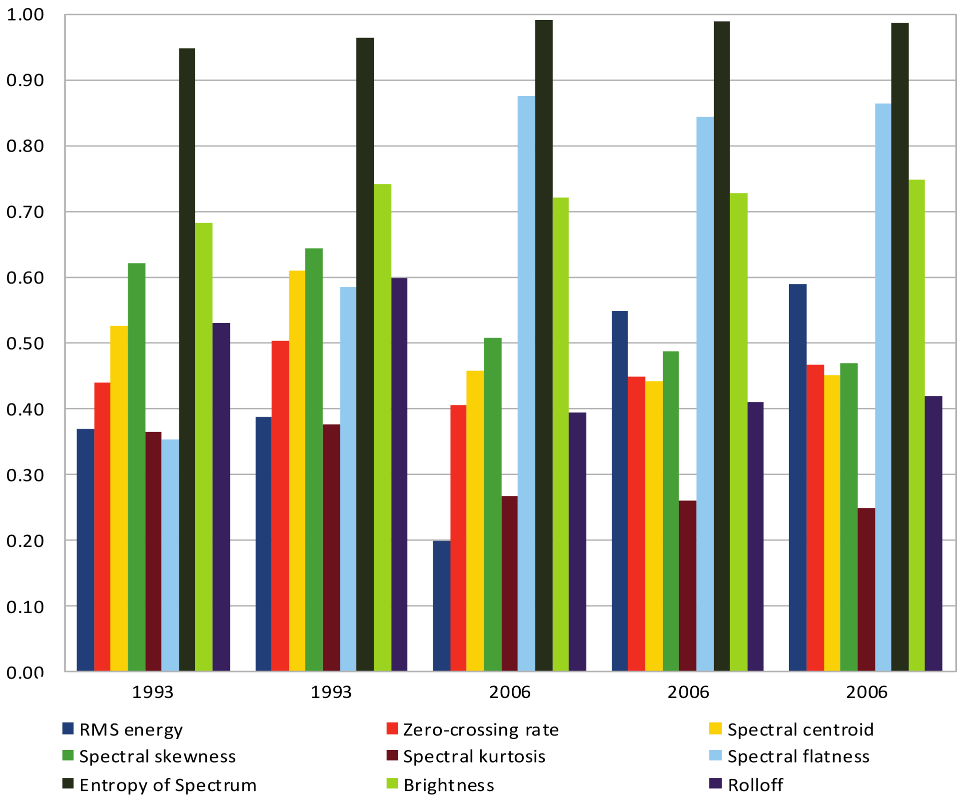

Fig. 10. Differences over the period of time of an artist's performance (Meat Loaf).

The highest value of the standard deviation was obtained for Spectral flatness. Relationships between different performances of the same song can also be found within the collection analysed. There are songs, for which regardless who the performer is, the results are similar (Fig. 11), such as Blitzkrieg Bop. However, differences may be observed between the original singer and the covers.

There are also songs where each performance returns different to some extent results (Fig. 12), such as for example Love Hurts. The largest differences may be observed for zero-crossing, spectral centroid, skewness and kurtosis.

Another group includes performances of covers which are similar, but they differ from the original performer, such as My Heart Will Go On by Celine Dion (Fig. 13). The parameters of music excerpts contained in Fig. 13 show, that two of the samples (the second and the third one) have similar results, while they differ from the results of the first sample. The largest values of the standard deviation were obtained for Spectral centroid and Rolloff.

There can also be found songs where the parameters depend on which fragment of the song is used for the analysis, and songs where the parameters do not depend on the fragment, they are constant for the whole track. An example of the first group of songs is a Green Day's song *Minority* (see Fig. 14), where parameter values are similar at the beginning and at the end, and they differ from the values of the parameters extracted for the middle fragment of the song. This is especially visible for Zero-crossing and Rolloff. The opposite example is Meat Loaf's song *What About Love* from 2006 (see Fig. 15). Parameter values for different fragments of the song are very close to each other.

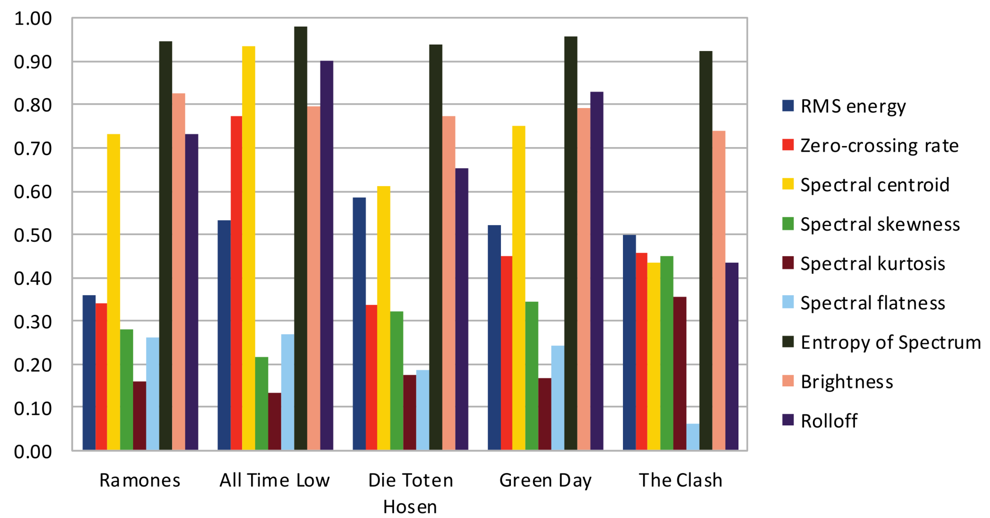

Fig. 11. Similar results between different performances of the same song *Blitzkrieg BOP* – song by Ramones (punk rock); covers by: All Time Low (pop punk), Die Toten Hosen (punk rock), Green Day (punk rock), The Clash (punk rock), The Offspring (punk rock).

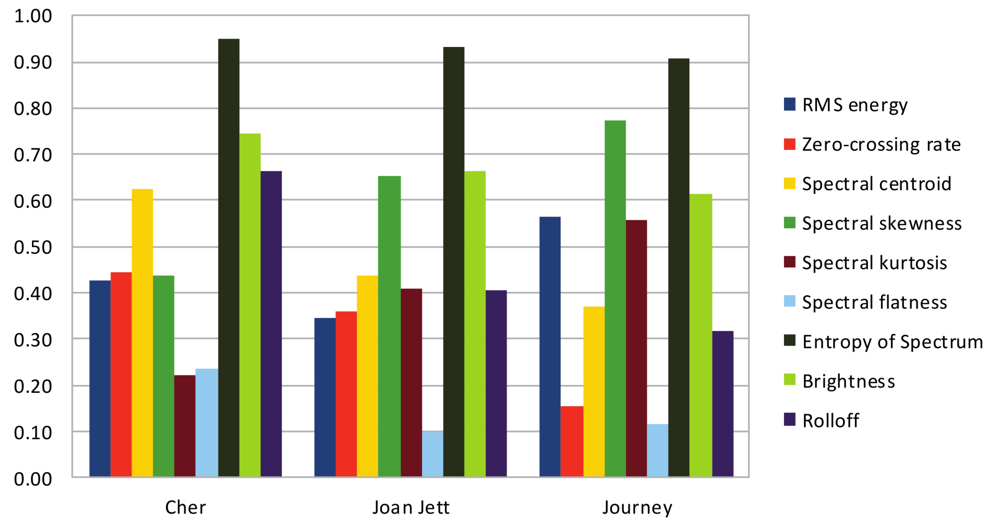

Fig. 12. Performance results of the song: *Love Hurts* – song by The Everly Brothers (rock'n'roll); covers by: Cher (pop), Joan Jett (punk rock), Journey (rock).

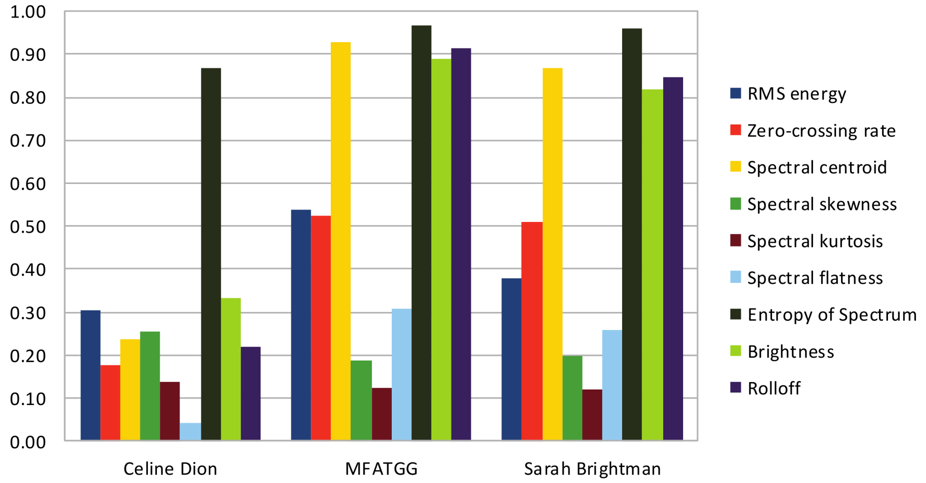

Fig. 13. Original performance different from covers: (*My Heart Will Go On* – song by: Celine Dion (pop); covers by: Me First and The Gimme Gimmes (punk rock), Sarah Brightman (musical singer).

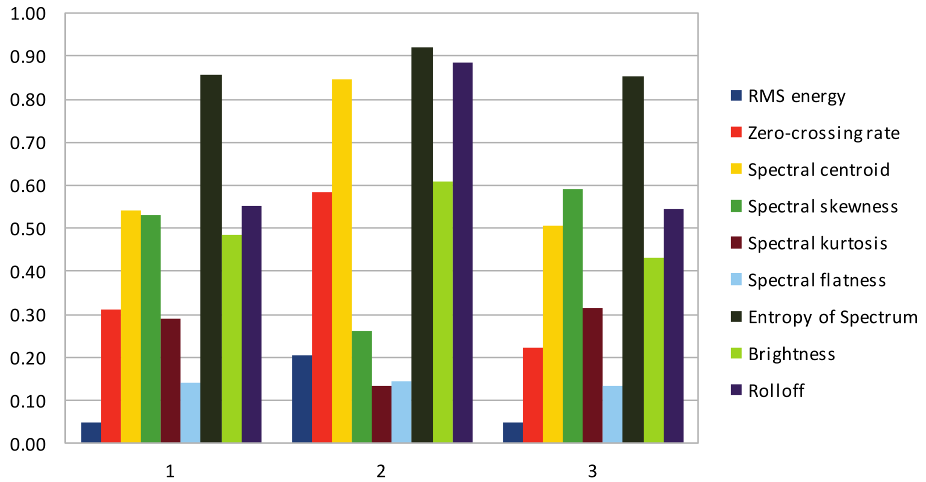

Fig. 14. Green Day's song *Minority* (1 – beginning, 2 – mid part of the song, 3 – end).

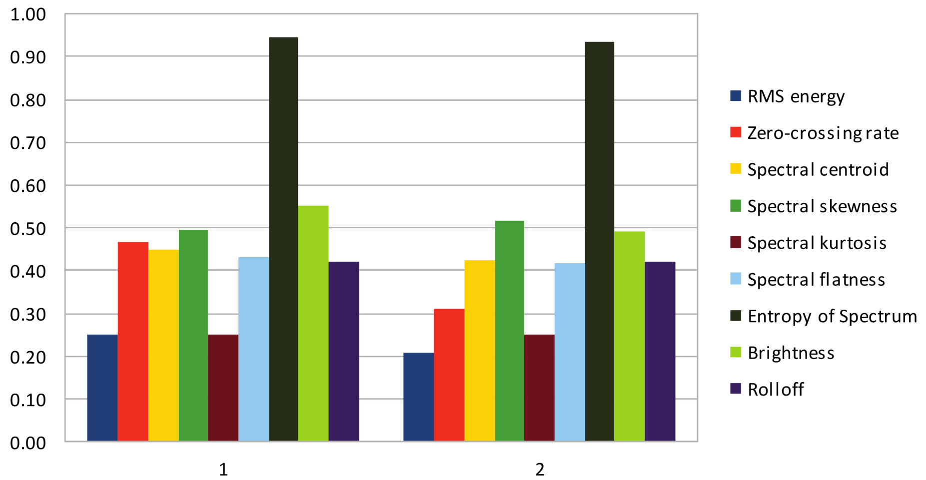

Fig. 15. Meat Loaf's song *What About Love* from 2006 (1 – Meat Loaf 1, 2 – Meat Loaf 2).

Finally, two interesting cases were analysed. In Fig. 16 parameter values for two versions of the same song are shown. The old version of Liebeslied (1988) is very punk-like. It is fast, garage, simply. On the other hand, the newer one (2009) is of a form of a ballad, more rock than punk. However, the instruments

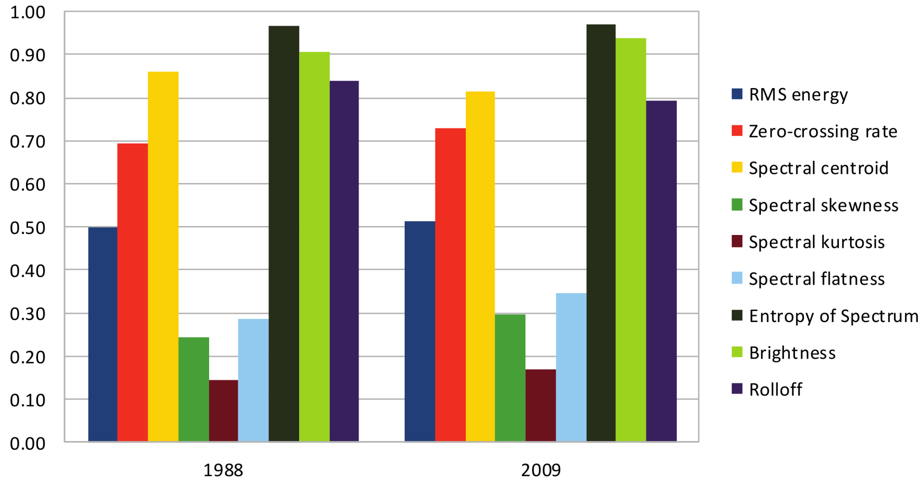

Fig. 16. No changes over time: Die Toten Hosen – Liebeslied – punk rock.

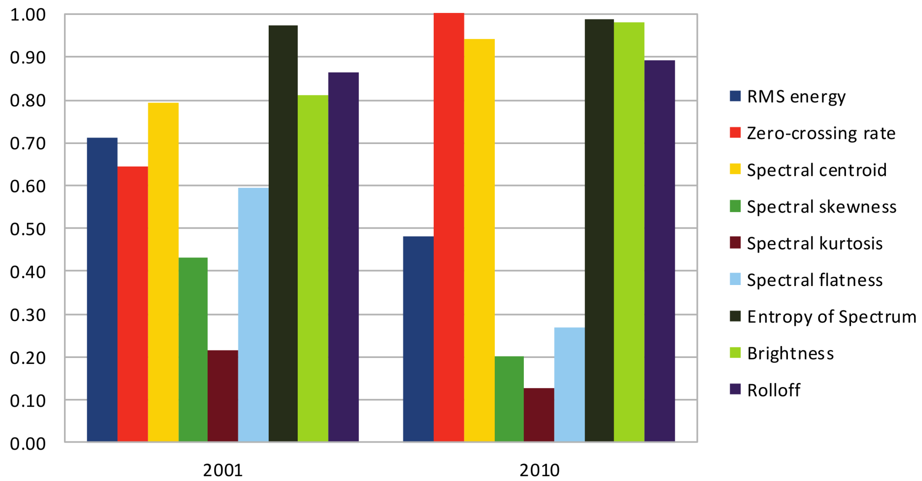

Fig. 17. Change of style: The Calling/Alex Band – *Wherever You Will Go* – post grunge/indie.

used are the same: vocal, two guitars, bass guitar and drums. The sound is different, but parameter values are similar.

Two versions of *Wherever You Will Go* (see Fig. 17), first recorded by the name of *The Calling* in 2001, then as Alex Band in 2010, sound similar, but parameter values differ much. These two versions are performed by the same band with the same instruments, it is a similar style. There are, however, some differences, the older version is more indie-like, also, when listening to the newer version, one can notice that musicians have more experience, but at the same time some health issues that the vocalist is struggling with.

4. Conclusions

The aim of the study was to compare values of parameters extracted from one song of many performances, representing a number of music genres, and various songs performed by one musician. The study shows that it is not obvious whether parameters are similar for an artist or for a song, even though they represent the same style or music genre. Contrarily, in some cases, the content of the songs is so distinctive, that a performance is not a significant factor. It is also possible to discern artists who developed their own characteristic features, which are visible in the song parametric representation. On the other hand, there are also artists, whose style evolved over time or they created just one album different from others in their repertoire.

This quantitative analysis shows that there are ambitious artists who use various techniques and instruments. For them, the analytical results are more diverse, and not so unified, but at the same time unique. Thus, such a performance may be difficult to be followed or imitated by others. At the same time, parameter extracted from songs performed by other more versatile artists are not always that diverse. In some cases, they are similar to the extend as if they were obtained from one song.

References

1. Alternative Press, 01.12.2016, http://www.altpress.com/index.php/news/entry/what_is_punk_this_new_infographic_can_tell_you
2. Barthet M., Fazekas G., Allik A., Thalmann F.B., Sandler M. (2016), *From interactive to adaptive mood-based music listening experiences in social or personal contexts*, Journal of the Audio Engineering Society, **64**, 9, 673–682, doi.org/10.17743/jaes.2016.0042.
3. Benward B., Saker M. (2003), *Music in theory and practice*, Vol. I, p. 12, 7th ed., McGraw-Hill.
4. Bergstra J., Casagrande N., Erhan D., Eck D., Kegl B. (2006), *Aggregate features and AdaBoost for music classification*, Machine Learning, **65**, 2/3, 473–484, http://dx.doi.org/10.1007/s10994-006-9019-7.
5. Bhalke D.G., Rajesh B., Bormane D.S. (2017), *Automatic genre classification using fractional fourier transform based mel frequency cepstral coefficient and timbral features*, Archives of Acoustics, **42**, 2, 213–222, doi: 10.1515/aoa-2017-002.
6. Burred J.J., Lerch A. (2014), *Hierarchical approach to automatic musical genre classification*, Journal of the Audio Engineering Society, **52**, 7/8, 724–739.
7. Daniels M. (2016), *Crowdsourcing the definition of "Punk"*, http://poly-graph.co/punk (retrieved 01.10.2016).
8. Dorochowicz A., Hoffmann H., Majdańczuk A., Kostek B. (2017), *Classification of musical genres by means of listening tests and decision algorithms*, ISMIS '2017, Springer Series Studies in Big Data, Intelligent Methods and Big Data in Industrial Applications, Bembenik R., Skonieczny L., Protaziuk G., Kryszkiewicz M., Rybinski H. [Eds.].
9. Dubnov S. (2004), *Generalization of Spectral Flatness Measure for Non-Gaussian Linear Processes*, IEEE Signal Processing Letters, **11**, 8, 698–701, doi: 10.1109/LSP.2004.831663.
10. Helman Z. (2016), *The concept of style and music of the twentieth century* [in Polish: *Pojęcie stylu a muzyka XX wieku*], Institute of Musicology, University of Warsaw, Warszawa, http://ksiegarnia.iknt.pl/uploads/files/PRM_2006_fragment.pdf.
11. Hoffmann P., Kostek B. (2015), *Bass enhancement settings in portable devices based on music genre recognition*, Journal of the Audio Engineering Society, **63**, 12, 980–989, http://dx.doi.org/10.17743/jaes.2015.0087.
12. Holzapfel A., Stylianou Y. (2008), *Musical genre classification using nonnegative matrix factorization-based features*, IEEE Transactions on Audio, Speech, and Language Processing, **16**, 2, 424–434, http://dx.doi.org/10.1109/TASL.2007.909434.
13. Kalliris G.M., Dimoulas C.A., Uhle C. (2016), *Guest Editors' note. Special issue on intelligent audio processing, semantics, and interaction*, Journal of the Audio Engineering Society, **64**, 7/8, 464–465.
14. Kostek B. (2005), *Perception-based data processing in acoustics. Applications to music information retrieval and psychophysiology of hearing*, Series on Cognitive Technologies, Springer Verlag, Berlin, Heidelberg, New York.
15. Kostek B., Kaczmarek A. (2013), *Music recommendation based on multidimensional description and similarity measures*, Fundamenta Informaticae, **127**, 1–4, 325–340, http://dx.doi.org/10.3233/FI-2013-912.
16. Kostek B., Kupryjanow A., Zwan P., Jiang W., Ras Z., Wojnarski M., Swietlicka J. (2011), *Report of the ISMIS 2011 contest: music information retrieval, foundations of intelligent systems*, ISMIS 2011, LNAI 6804, pp. 715–724, M. Kryszkiewicz *et al.* [Eds.], Springer Verlag, Berlin, Heidelberg.
17. Kotsakis R., Kalliris G., Dimoulas C. (2012), *Investigation of broadcast-audio semantic analysis scenarios employing radio-programme-adaptive pattern classification*, Speech Communication, **54**, 6, 743–762.
18. Lane D.M. (2017), *Difference between two means (Independent Groups)*, http://onlinestatbook.com/2/tests_of_means/difference_means.html
19. Matlab MIRtoolbox 1.6.2 Specification.
20. Ntalampiras S. (2013), *A novel holistic modeling approach for generalized sound recognition*, IEEE Signal Processing Letters, **20**, 2, 185–188, http://dx.doi.org/10.1109/LSP.2013.2237902.
21. Palisca C.V. (1998), *Marc Scacchi's defense of new music (1649)* [in Polish: *Marca Scacchiego obrona nowej muzyki (1649)*], Muzyka, **XLIII**, 2, 131–132.
22. Pascall R. (2001), *The new Grove dictionary of music and musicians*, S. Sadie, J. Tyrrell [Eds.], 24, 2/London, pp. 638–642.
23. Plewa M., Kostek B. (2015), *Music mood visualization using self-organizing maps*, Archives of Acoustics, **40**, 4, 513–525, doi: 10.1515/aoa-2015-0051.
24. Pluta M., Spalek L.J., Delekta R.J. (2017), *An automatic synthesis of musical phrases from multipitch samples*, Archives of Acoustics, **42**, 2, 235–247, doi: 10.1515/aoa-2017-0026.
25. Reljin N., Pokrajac D. (2017), *Music performers classification by using multifractal features: a case study*, Archives of Acoustics, **42**, 2, 223–233, doi: 10.1515/aoa-2017-0025.
26. RockSound (2016), http://www.rocksound.tv/news/read/study-green-day-blink-182-are-punk-my-chemical-romance-are-emo (retrieved 01.13.2016).
27. Rosner A., Kostek B. (2018), *Automatic music genre classification based on musical instrument track separation*, Journal of Intelligent Information Systems, **50**, 363–384 doi: 10.1007/s10844-017-0464-5.
28. Rosner A., Schuller B., Kostek B. (2014), *Classification of music genres based on music separation into harmonic and drum components*, Archives of Acoustics, **39**, 4, 629–638, doi: 10.2478/aoa-2014-0068.
29. Schedl M., Gómez E., Urba J. (2014), *Music information retrieval: recent developments and applications*, Foundations and Trends® in Information Retrieval, **8**, 2–3, 127–261, doi: 10.1561/1500000042.

30. Seidel W., Leisinger U. (1998), *Music in History and the Present* [in German: *Die Musik in Geschichte und Gegenwart*], article: *Stil*, [Ed.] L. Finscher, pp. 1740–1759, Bärenreiter and Metzler, Kassel.

31. Silla C.N., Kaestner C.A., Koerich A.L. (2007), *Automatic music genre classification using ensemble of classifiers*, IEEE International Conference on Systems, Man and Cybernetics, pp. 1687–1692.

32. *Small Music Encyclopedia* (1968), [in Polish: *Mała Encyklopedia Muzyki*], Śledziński S. [Ed.], PWN, Warszawa.

33. Soper D.S. (2017), *Student t-value calculator* [Software], http://www.danielsoper.com/statcalc/calculator.aspx?id=10 (retrieval 02.18.2017).

34. Tekman H.G., Hortacsu N. (2002), *Aspects of stylistic knowledge: what are different styles like and why do we listen to them?*, Psychology of Music, **30**, 1, 28–47.

35. Tzanetakis G., Essl G., Cook P. (2002), *Automatic musical genre classification of audio signals*, IEEE Transactions on Speech and Audio Processing, **10**, 5, 293–302.