

# Rough Sets Applied to Mood of Music Recognition

Bożena Kostek

Gdansk University of Technology, Faculty of  
Electronics, Telecommunications and Informatics,  
Audio Acoustics Laboratory, Narutowicza 11/12,  
80-233 Gdańsk, Poland  
Email: bokostek@audioakustyka.org

Magdalena Plewa

Gdansk University of Technology, Faculty of  
Electronics, Telecommunications and Informatics,  
Multimedia Systems Department and Audio Acoustics  
Laboratory, Narutowicza 11/12, 80-233 Gdańsk, PL  
Email: mplewa@sound.eti.pg.gda.pl

□  
**Abstract**—Mood of music is considered as one of the most intuitive criteria for listeners, thus this work is focused on the emotional content of music and its automatic recognition. The research study presented in this work contains an attempt to music emotion recognition including audio parameterization and rough sets. A music set consisting of 154 excerpts from 10 music genres was evaluated in the listening experiment. This may be treated as a ground truth. The results achieved indicated a strong correlation between subjective results and objective descriptors and on that basis a vector of parameters related to mood of music was created. On the other hand, rough set-based processing was applied to derive reducts containing the most promising features in the context of mood recognition, as well as confusion matrices of the mood recognition. Both approaches indicate strong relationship between objective descriptors and subjective evaluation of mood of music.

## INTRODUCTION

In music perception studies many different classifications and systems that describe music components are defined. Levitin [1] observed that from the listener's perspective there are seven major elements of music: loudness, pitch, melody, harmony, rhythm, tempo, and meter. These components are significant for discussion related to emotions included in music. The conventional approach to studying music emotion perception or to assigning music genre consists in subjective tests, in which a number of listeners evaluate a given music excerpt, and then these results are analyzed using statistical processing. This process is very lengthy and arduous, and does not always return reliable results, as mood and emotion are often treated by the listeners as interchangeable terms. Even such a notion as "music genre" may not precisely be defined, but still it plays an essential role in music appreciation and cognition. Mood plays an important role in music interpretation and annotation. Still, even though mood is an intuitive way of music describing, it is very difficult to find an exact correlation between physical features and perceived mood, which is necessary to make the annotation process automatic. This way, mood could automatically be added to music recommendation systems [2], [3].

Music as a form of art is perceived and interpreted in many different ways. It contains layers of music composition

elements, emotions that may elicit different meanings, references to other pieces and other elements that are difficult to define and interpret. On the other hand, music can be treated as an audio signal and parameterized according to its temporal and spectral characteristics. The relationship between music described by music features and parameters derived from the signal is very difficult to find. The researchers' task is to identify parameters that are related to those features.

As already mentioned, with the growth of accessible digital music libraries over the past decade, there is a need for research into automated systems for searching, organizing and recommending music. Therefore, the aim of this study is to find a correlation between subjective evaluation of music mood and objective data processing. For that purpose, the rough set-based [4], [5] analysis is carried out, employing the Rough Set Exploration (RSES) system [6], [7].

The paper is organized as follows; Section II presents the mood model used in this study, list of labels and the principles of the listening tests carried out by the authors, as well as results of mood perception-based evaluation. In Section III correlation analysis is performed for data gathered in the listening tests and in feature vectors that are assigned to the set of 150 music excerpts. As a result, a table is created that contains the most significant parameters correlated to two mood-related features, namely: valence and arousal. Next, the same data are analyzed using the rough set-based processing and then results obtained compared to those presented in Section III. Finally, a short summary is included in Conclusion Section.

## SUBJECTIVE EVALUATION

### A. Model of Mood of Music

This section includes information with regard to the main experiment that aimed for subjective emotional content evaluation of larger set music. Outcomes from previous experiments [8]–[12] were taken into consideration and affected the final form of this experiment including the model of emotions, music dataset and the experimental procedure.

The test was executed to collect subjective mood evaluation results of a large set of songs using specially

---

This work was supported by Gdansk University of Technology

designed graphical interface (Fig. 1, for translation see Table I). The model of mood of music used in the experiment was proposed by the authors. The main assumptions were that the model has to be intuitive for users and compatible with dimensional models consisting of two dimensions as proposed by Thayer [13].



Fig. 1 Graphical interface dedicated for mood of music evaluation

Since previous stages of the research showed that 2-dimensional model is not very intuitive for listeners, an alternative solution was proposed [10]. The set of mood labels was selected from the Mood Dictionary Associated with Music proposed by the authors [10] Mood labels (originally in Polish) along with their translation can be found in Table I.

Mood descriptors were placed on a 2-dimensional plane, with regards to dimensions retrieved from Multidimensional Scaling (MDS) experiment [11]. This placement is coherent with Thayer's model [13] and Russel's [14] emotion representation. It is also consistent with findings of Brinker *et al.*, [15] and Hevner [16], who placed mood descriptors on Valence/Arousal (VA) plane.

TABLE I.

LIST OF MOOD LABELS USED IN GRAPHICAL INTERFACE DESIGNED FOR MOOD OF MUSIC REPRESENTATION

No.	Mood label (Polish)	Mood label (English)
1	Agresywny	Aggressive
2	Depresyjny	Depressive
3	Ekscytujący	Exciting
4	Energiczny	Energetic
5	Neutralny	Neutral
6	Relaksujący	Relaxing
7	Smutny	Sad
8	Spokojny	Calm
9	Wesoły	Joyful

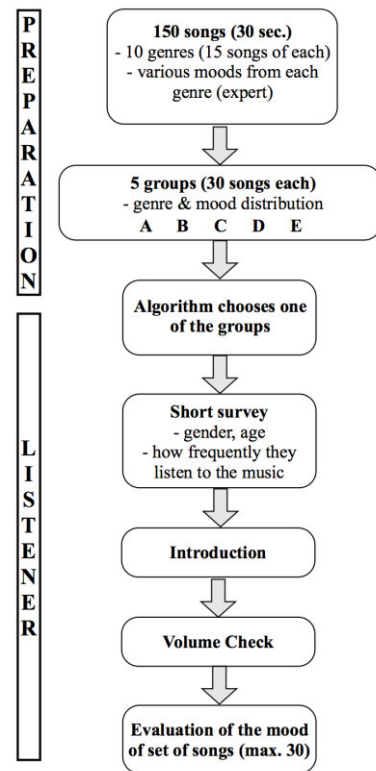


Fig. 2 Listening test arrangement related to music mood evaluation

The intensity of color corresponds to the intensity of the particular emotion contained in music. The "white" area placed in the center is considered as a neutral, where no emotional content is included. This concept is also strongly related to fuzzy logic and the concept of "degrees of truth". It includes various states of truth in between 0 (false) and 1 (truth), which is very intuitive, when it comes to evaluation of such a subtle substance as emotions.

### B. Listening Procedure

The stages of the test are presented in Fig. 2. The main part of the test consisted of a series of musical excerpts presented one after another, where listeners were asked to evaluate the mood of music by clicking at the graphical mood representation. It was preceded by a short survey and the level check. The experiment was performed in Polish and its total average duration was approximately 12 minutes. The test was performed using a WEB-based interface. 112 listeners (57 women and 55 man) within age from 16 to 56 (average age 28) participated in the experiment. Majority of the audience reported that they listen to music everyday.

150 tracks were chosen from 10 different music genres, to obtain a diversified set. Chosen music styles were as follows: Blues, Classical, Country, Dance & DJ, Hard Rock & Metal, Jazz, Pop, R&B, Rap & Hip-Hop, Rock. It is worth noting, that styles chosen are easy to distinguish between each other and cover various music material. 30-sec music excerpts used in the experiment came from the SYNAT music database

[17], [18]. A detailed list of music files used in the experiments is available in [9].

### C. Results

Answers provided by subjects were pre-processed and only valid entries were included in the result analysis. A submission was considered as valid, if consisted of 5 to 30 evaluated songs.

Results were analyzed with both dimensional and label approaches. In the dimensional approach answers were analyzed in the polar coordinates. To each field on the graphical interface, the number value was assigned according to the mood intensity (from 0 to 3) and the angle was assigned according to the position of the label. This allowed mapping onto 2D Energy/Arousal plane (Fig. 3). Achieved results were used as polar and Euclidean coordinates depending on the employed method. Selected results are listed in Table II.

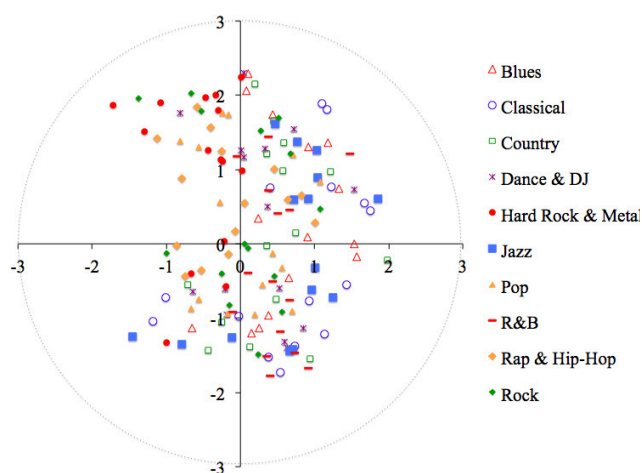


Fig. 3 Mapping of 150 songs (divided by the genre) onto mood plane based on the listening test results

Averaged ratings are mostly within range [0,2]. Only for aggressive mood a stronger result was obtained. In this case mood of music might be strongly related to the perceived emotion. Also distribution of music pieces sorted by genre is very interesting.

For pop and rock music, excerpts are distributed in all quadrants of the AV plane. There is no jazz songs in III quadrant. Blues was placed by listeners mostly on the right half of the VA plane (positive) as well as classical, where only two pieces were placed in quadrant III. A similar tendency is observed for country excerpts. Rap & Hip-Hop was considered mostly as music with high arousal and appeared mostly in quadrants I and II but some excerpts are placed in quadrant III.

Distribution of Blues and R&B was quite similar (mostly quadrants I and IV), which might be related to musical similarities and common roots of both genres. Although blues is often considered as very sad music, it was placed on the "positive" half of VA plane.

What is interesting, Dance & DJ excerpts are distributed among all quadrants as well as Pop and Rock. These music genres are very frequent in popular culture, therefore are not strongly related to one and only esthetics or topic but are mixtures of different trends. That is also reflected in values of the standard deviation, which are highest for these genres. Averaged results for each genre are listed in Table II. This table was used in the rough set-based analysis further on. Detailed results of the listening test are presented in [9].

TABLE II.  
AVERAGED RESULTS FOR VARIOUS MUSIC GENRES

Music genre	Averaged Valence	Averaged Arousal	St. Dev. Valence	St. Dev. Arousal
Blues	0.60	0.33	1.11	1.16
Classical	0.68	-0.25	1.28	1.25
Country	0.44	0.10	1.05	1.02
Dance & DJ	0.25	0.26	1.06	1.15
Hard Rock & Metal	-0.55	1.03	0.99	1.09
Jazz	0.60	-0.10	1.19	1.33
Pop	0.04	0.22	1.04	1.20
R&B	0.49	-0.32	0.99	1.12
Rap & Hip-Hop	-0.17	0.62	1.10	1.12
Rock	0.00	0.43	1.24	1.13

All observations listed above are important cues for conducting experiments with decision systems. Although they are related to the specific music set, they might represent more general trends due to carefully selecting excerpts for the performed evaluation.

Contrarily, the label analysis was performed, where the number of occurrences of each label was calculated for every song. As a result, each song is described with a 9-element vector, where each position refers to the mood label (Table I). The value describes the percentage of occurrences of each label. Some songs are described by all listeners with mainly one label, while for other evaluations is spread among the labels. Examples of songs along with their label description are shown in Fig. 4.

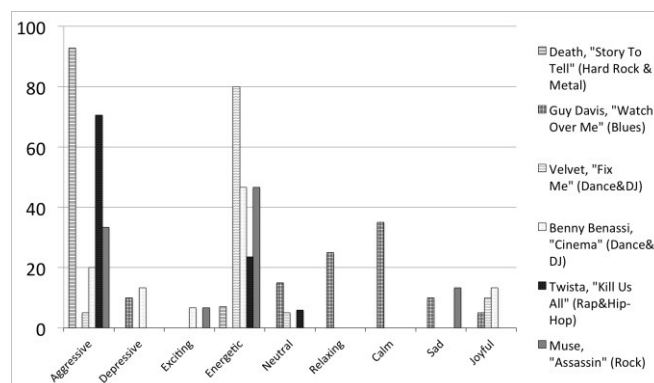


Fig. 4 Example of results of mood labels assigned to particular songs. The vertical axis describes the percentage of occurrences of each label

## CORRELATION ANALYSIS

The starting point of the feature vector (FV) content creation for the purpose of automatic mood recognition was examination of previous studies performed in MIR by the authors and their collaborators. Resulted from them was FV applied to two databases, namely ISMIS [18] and SYNAT [9], [19]–[25], thus its content may be treated as very thoroughly analyzed. Moreover, the same FV was used in the ISMIS'2011 conference in music competition [18], in which more than 100 teams participated, thus it may also be treated as a kind of benchmarking. ISMIS is a database of approx. 1300 music excerpts of high quality audio excerpts, collected and divided into six music genres. On the other hand, the SYNAT database is a collection of 52532 pieces of music described with a set of descriptors obtained through the analysis of mp3-quality recordings. For the SYNAT database, the analysis band is limited to 8kHz. The database stores 173-feature vectors, which in majority are the MPEG-7 standard parameters (109). The vector has additionally been supplemented with 20 Mel-Frequency Cepstral Coefficients (MFCC), 20 MFCC variances and 24 time-related ‘dedicated’ parameters. The SYNAT database was realized by the Gdansk University of Technology (GUT) [19] and music was collected from the Internet. The vector includes parameters associated with the MPEG-7 standard, melcepstral coefficients (MFCC) and is supplemented by the so-called dedicated parameters which refer to temporal characteristic of the analyzed music excerpt. Full list of parameters was shown in the earlier study [12], [20].

In addition, parameters from the MIR Toolbox [26] were calculated. This set of parameters contains a lot of information but not necessarily related to mood of music. At previous stages of this study a selection of parameters based on correlation analysis was performed and returned good results, therefore this approach was also applied. Correlation between subjective values of Valence and Arousal and parameters was calculated and therefore a set of features strongly related to mood was created. Only parameters with correlation coefficient higher than 0.50 were included in the final feature vector. Eventually, the feature vector describing mood of music consisted of 16 parameters from SYNAT, listed in Table III and 16 parameters from MIR Toolbox Table IV. It is worth noting that correlation is slightly stronger with parameters based on music characteristic than from SYNAT, which describe general properties of an audio signal. Description of parameters included in SYNAT can be found in [9], [12] and features from MIR Toolbox in the MIR manual [26].

TABLE III.

PARAMETERS CORRELATED WITH SUBJECTIVE MOOD OF MUSIC EVALUATION SELECTED FROM 173-SYNAT FVS

Valence		
No.	Parameter	Corr.
1	ASE2	-0.72

2	MFCC7	-0.62
3	PEAK_RMS10FR_MEAN	0.51
4	ASE_M	-0.50
5	ASE26	-0.50

## Arousal

6	MFCC1	-0.79
7	MFCC2	-0.78
8	SFM13	-0.63
9	SFM12	-0.61
10	SFM14	-0.56
11	SFM15	-0.56
12	SFM10	-0.53
13	1RMS_TCD	0.52
14	SFM_M	-0.51
15	SFM11	-0.51
16	SFM16	-0.50

TABLE IV.

PARAMETERS CORRELATED WITH SUBJECTIVE MOOD OF MUSIC EVALUATION SELECTED FROM THE MIR TOOLBOX

Valence		
No.	Parameter	Corr.
1	Spectral irregularity	0.73
2	MTBF2	0.54
3	Spectral roughness	0.52
Arousal		
4	Brightness	0.83
5	Entropy of Spectrum	0.79
6	Timbre Zerocross.	0.64
7	Tonal Harmonic Change Detection	0.64
8	Harmonic Change Detection Function	0.62
9	Spectral Centroid Mean	0.57
10	Key Clarity	0.55
11	Tempo	0.55
12	Spectral Flux Period	0.55
13	Spectral Irregularity	0.53
14	Spectral Rolloff 185	0.52
15	Spectral Kurtosis	0.51
16	Spectral roughness	0.50

## ROUGH SETS-BASED ANALYSIS

## A. Input Data for Rough Sets

As mentioned earlier, the RSES was used in this study. The system generates interpretable results in the form of reducts,

classification measures, etc., thus it is very useful for the analysis and classification of data.

Data for the rough set-based analysis were prepared in various configurations. Table V gathers datasets in a form of decision tables, in which: "Mood label I", "Mood label II", and "Music Genre" are the decision attributes. Averaged values of Valence and Arousal were calculated, and averaged position on the Valence/Arousal plane indicated Mood label I. Mood label II was chosen according to the label analysis, where the number of occurrences of each label was calculated for every song and the label with maximum number of occurrences was chosen. The last column shows music genres. As seen from Table V, not always labels from Mood labels I and II are consistent between each other. Moreover, it is interesting that the position in VA plane does not condition music genre assigned.

Datasets from Table V were analyzed by the RSES system, the returned results are shown in a form of confusion matrices (Tables VI-VIII), containing true positives (TP), as well as coverage and accuracy measures. Data were randomly split into 60% and 40% subsets for training and testing the classifier. For every dataset a separate set of rules was created by the following procedure:

1. Every parameter value range was discretized by a local variant of Maximum Discernibility discretization method.
2. Reducts were calculated from training subset by the genetic method available in RSES software.
3. Reducts were tested against training subset to create decision rules, based on every object in the subset.
4. Testing subset was classified and accuracy was expressed by calculating a confusion matrix.

TABLE V.  
LISTENING EXPERIMENT RESULTS USED IN THE RSES-BASED PROCESSING

Song No.	Valence	Arousal	Mood Label I	Mood Label II	Music Genre
1	0.09	2.06	Energetic	Energetic	Blues
2	0.43	1.74	Energetic	Energetic	Blues
...	...	...	...	...	...
15	0.92	1.31	Exciting	Energetic	Blues
16	0.41	0.75	Exciting	Exciting	Classical
17	1.23	0.77	Exciting	Exciting	Classical
25	-1.00	-0.73	Depressive	Sad	Classical
...	...	...	...	...	...
30	1.67	0.55	Joyful	Exciting	Classical
31	0.36	1.21	Energetic	Energetic	Country
32	0.19	2.15	Energetic	Energetic	Country
...	...	...	...	...	...
40	-0.71	-0.55	Depressive	Depressive	Country
45	1.21	0.97	Exciting	Energetic	Country
46	-0.81	1.77	Aggressive	Aggressive	Dance & DJ
47	-0.21	-0.61	Calm	Relaxing	Dance & DJ
...	...	...	...	...	...

60	0.72	1.54	Exciting	Energetic	Dance & DJ
61	-1.30	1.51	Aggressive	Aggressive	Hard Rock & Metal
62	-1.72	1.86	Aggressive	Aggressive	Hard Rock & Metal
...	...	...	...	...	...
71	-0.19	-0.57	Calm	Sad	Hard Rock & Metal
...	...	...	...	...	...
75	-0.26	1.14	Energetic	Aggressive	Hard Rock & Metal
76	0.77	1.37	Exciting	Energetic	Jazz
77	0.72	0.59	Exciting	Energetic	Jazz
...	...	...	...	...	...
90	0.92	0.61	Exciting	Energetic	Jazz
91	-0.56	1.30	Energetic	Energetic	Pop
92	-0.16	1.74	Energetic	Energetic	Pop
...	...	...	...	...	...
105	0.69	-0.90	Relaxing	Calm	Pop
106	-0.05	1.17	Energetic	Energetic	R&B
107	0.38	0.72	Exciting	Exciting	R&B
...	...	...	...	...	...
120	0.43	-0.51	Relaxing	Neutral	R&B
121	-1.12	1.42	Aggressive	Aggressive	Rap & Hip-Hop
122	-0.59	1.84	Energetic	Aggressive	Rap & Hip-Hop
...	...	...	...	...	...
135	-0.26	1.25	Energetic	Aggressive	Rock
136	-1.37	1.96	Aggressive	Aggressive	Rock
...	...	...	...	...	...
149	0.24	-1.49	Calm	Calm	Rock
150	1.08	0.47	Exciting	Joyful	Rock

### B. Results

As seen from Tables VI-VIII, the rough set-based processing brought satisfying results with a high accuracy and coverage.

TABLE VI.  
CONFUSION MATRIX FOR CLASSIFICATION OF MOOD LABEL I, WITH ATTRIBUTES "VALENCE", "AROUSAL", "GENRE"

	Energ.	Exc.	Joy.	Calm	Relax.	Depr.	Aggr.	Sad	No. of obj.	Acc.	Cov.
Energ.	16	0	0	0	0	0	0	0	16	1	1
Exc.	0	11	0	0	0	0	0	0	11	1	1
Joy.	0	0	4	0	0	0	0	2	6	0.667	1
Calm	1	0	0	9	0	0	0	0	10	0.9	1
Relax.	1	0	0	0	1	0	0	0	2	0.5	1
Depr.	0	0	1	0	0	7	0	0	9	0.875	0.889
Aggr.	1	0	0	0	0	0	3	0	4	0.75	1
Sad	0	0	0	0	1	0	0	0	2	0	0.5
TP	0.84	1	0.8	1	0.5	1	1	0			

Total number of tested objects: 60, Total accuracy: 0.879,  
Total coverage: 0.967.

TABLE VII.

CONFUSION MATRIX FOR CLASSIFICATION OF MOOD, WITH FEATURES “VALENCE”, “AROUSAL”, CLASSIFICATION GOAL IS (MOOD LABEL I)

	Energ.	Exc.	Joy.	Calm	Relax.	Depr.	Aggr.	Sad	No. of obj.	Acc.	Cov.
Energ.	10	0	1	1	0	0	0	0	13	0.833	0.923
Exc.	0	4	0	0	0	0	0	0	5	1	0.8
Joy.	0	0	13	1	0	0	0	0	14	0.929	1
Calm	0	0	0	2	0	0	0	0	8	1	0.25
Relax.	0	0	0	0	7	1	0	0	8	0.875	1
Depr.	0	0	0	0	0	9	0	0	10	1	0.9
Aggr.	0	0	0	0	0	0	2	0	2	1	1
Sad	0	0	0	0	0	0	0	0	0	0	0
TP	1	1	0.93	0.5	1	0.9	1	0			

Total number of tested objects: 60, Total accuracy: 0.922,  
Total coverage: 0.85.

TABLE VIII.

CONFUSION MATRIX FOR CLASSIFICATION OF MOOD, WITH FEATURES “VALENCE”, “AROUSAL”, CLASSIFICATION GOAL IS MOOD LABEL II

	Energ.	Exc.	Joy.	Calm	Relax.	Depr.	Aggr.	Sad	No. of obj.	Acc.	Cov.
Energ.	12	0	0	0	0	2	0	0	15	0.857	0.933
Exc.	0	9	0	0	0	0	0	0	9	1	1
Joy.	0	0	4	0	0	0	0	0	5	1	0.8
Calm	0	0	0	9	1	0	0	0	12	0.9	0.833
Relax.	0	0	1	0	6	0	0	0	8	0.857	0.875
Depr.	0	0	1	0	0	5	1	0	7	0.714	1
Aggr.	0	0	0	0	0	0	2	0	2	1	1
Sad	0	0	1	0	0	0	0	0	2	0	0.5
TP	1	1	0.57	1	0.86	0.71	0.67	0			

Total number of tested objects: 60, Total accuracy: 0.87,  
Total coverage: 0.9.

Another set of data containing FVs and assigned Mood label I, Mood label II and music genre for the same 150 music excerpts was called FVs. These data returned confusion matrix (see Table IX). This dataset was classified by the following 10 reducts, with sizes ranging from 8 features to 11 features, each with positive region of 1.0 (i.e. percentage of objects from training set accurately classified), see Table X.

It should be of interest that reducts contain rhythm-related features belonging to “dedicated” features, such e.g. 1RMS\_TCD\_10FR\_VAR, 2RMS\_TCD\_10FR\_VAR, etc.

This confirms a notion that music genre and related mood rely heavily on rhythmic features of music. Moreover, when comparing the correlation analysis performed (Table III), it may be seen that similar parameters were derived from this analysis for Valence and Arousal, which are features corresponding to mood of music.

TABLE IX.

CONFUSION MATRIX FOR FVs-BASED SET

	Energ.	Exc.	Joy.	Calm	Relax.	Depr.	Aggr.	Sad	No. of obj.	Acc.	Cov.
Energ.	3	0	1	0	3	0	6	1	14	0.214	1
Exc.	1	2	4	0	2	1	1	2	13	0.154	1
Joy.	0	1	1	1	0	0	0	0	3	0.333	1
Calm	1	0	0	1	0	2	0	1	5	0.2	1
Relax.	0	0	2	1	7	0	1	1	12	0.583	1
Depr.	0	0	3	2	2	1	1	0	9	0.111	1
Aggr.	0	0	0	0	0	0	3	0	3	1	1
Sad	0	0	0	1	0	0	0	0	1	0	1
TP	0.6	0.67	0.09	0.17	0.5	0.25	0.25	0			

Total number of tested objects: 60, Total accuracy: 0.3,  
Total coverage: 1.

As mentioned before other datasets combinations were tested. Two more attempts were made to classify songs:  
- mood based on “Valence”, “Arousal”, and music genre features,  
- genre based on “Valence”, “Arousal”, and “Mood Label I” features.

Even though, confusion matrices did not bring very satisfying results, i.e. the total accuracy was at level of 0.25 and the total coverage was approx. 0.56, still, the classification performed, based on a majority voting, returned adequate results.

TABLE X.

REDUCTS DERIVED FROM THE RSES SYSTEM FOR FVs

No.	Parameters
1	1RMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, ASC_V, ASE12, ASE27, ASEV14, MFCC20, SFMV6
2	1RMS_TCD, 1RMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, 3RMS_TCD, ASC_V, ASE_MV, ASE10, ASE17, ASE27
3	1RMS_TCD, 1RMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, 3RMS_TCD, ASC_V, ASE_MV, ASE10, ASE27, SFM7
4	1RMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, 3RMS_TCD_10FR_MEAN, ASE_MV, ASE10, ASE12, ASE20, ASEV22, MFCC20, SFM8
5	1RMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, 3RMS_TCD_10FR_MEAN, ASC_V, ASE_MV, ASE10, ASE5, ASEV14, ASEV22, MFCC20

6	IRMS_TCD, IRMS_TCD_10FR_VAR, ASE10, ASE17, ASE20, ASE5, ASEV1, MFCC20, MFCC7, SFM5, SFMV6
7	IRMS_TCD, IRMS_TCD_10FR_VAR, 3RMS_TCD, ASE_MV, ASE10, ASE17, ASE20, MFCC7, SFM5, SFM7, SFMV6
8	IRMS_TCD, IRMS_TCD_10FR_VAR, 2RMS_TCD_10FR_VAR, ASC_V, ASE10, ASE12, ASEV14, MFCC7, SFM15, SFM8, SFMV6
9	IRMS_TCD, 2RMS_TCD_10FR_VAR, 3RMS_TCD_10FR_MEAN, ASC_V, ASE12, ASE22, ASE27, ASEV14, ASEV17, SFM5, SFM7
10	IRMS_TCD, 2RMS_TCD_10FR_VAR, 3RMS_TCD, ASC_V, ASE_MV, ASE12, ASE17, ASE27, ASE9, ASEV1, ASEV17

Finally, the last experiment used a decision table consisted of the joint Synat and MIR Toolbox parameters (derived from the correlation analysis as the ones correlated with mood of music). The decision attribute was Mood Label I. The confusion matrix for this dataset is presented in Table XI. 44 reducts resulted from the processing, each containing 6 to 8 attributes (Fig. 5).

TABLE XI. CONFUSION MATRIX FOR CLASSIFICATION OF MOOD

	Energ	Depr.	Exc.	Joy.	Relax.	Calm	Aggr.	Sad	No. of obj.	Acc.	Cov.
Energ.	5	0	0	1	0	0	1	2	9	0.556	1
Depr.	0	1	0	2	1	2	0	1	7	0.143	1
Exc.	4	0	5	2	1	1	1	0	14	0.357	1
Joy.	0	0	0	3	1	1	0	0	5	0.6	1
Relax.	0	2	0	0	4	1	1	2	10	0.4	1
Calm	1	1	1	1	1	3	1	2	11	0.273	1
Aggr.	0	0	0	0	0	0	2	0	2	1	1
Sad	0	0	0	1	0	0	1	0	2	0	1
TP	0.5	0.25	0.83	0.3	0.5	0.38	0.29	0			

Total number of tested objects: 60, Total accuracy: 0.383, Total coverage: 1.

Omitted attributes:

Tonal\_hcdf1\_HarmonicChangeDetectionFunction\_Mean, Spectral\_irregularity1\_Spectral\_irregularity\_Mean2, Spectral\_kurtosis1\_Spectral\_kurtosis\_Mean.

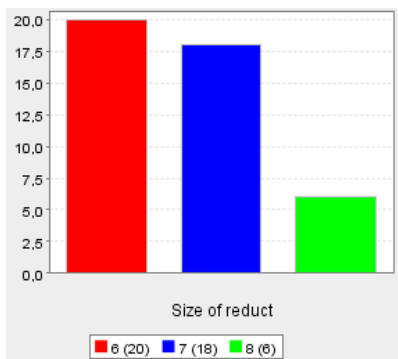


Fig. 5 Histogram of reduct sizes

An attribute called Spectral irregularity\_mean was the most often used (see Fig. 6), also Spectral brightness is among those most often occurring, which confirm the results derived from the correlation analysis (see Table IV).

Based on these attributes and all training cases a set of rules was generated, containing 2583 rules (Fig. 7), aimed at precise classification of all 90 training objects. Such a disproportion indicates low quality of approximations, and low generalization, as the test set is classified with total accuracy of only 0.383.

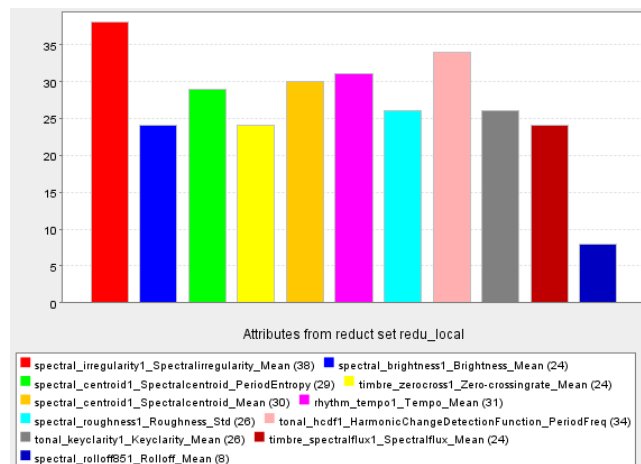


Fig. 6 Histogram of occurrence of attributes in reducts

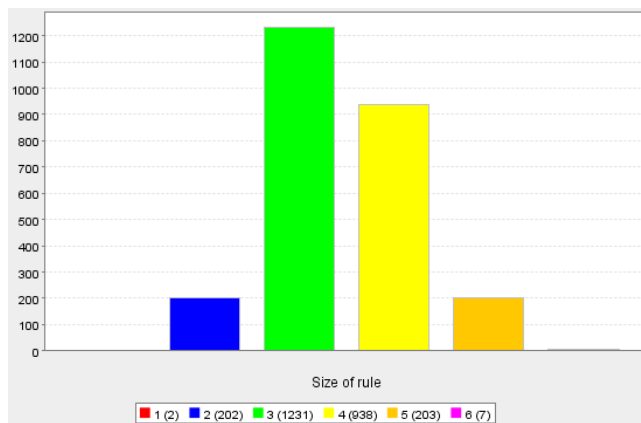


Fig. 7 Histogram of rule lengths

CONCLUSION

Several subsets were created for mood classification within the study presented. Gathered data contain results of subjective evaluation, which may be treated as ground truth in this type of classification. This means that we treat listeners' opinions, derived from a meaningful statistical analysis, as the starting point in such analyses. However, it should be remembered that listeners may impose their own emotions on those intended by the composer. That's why it would be preferable if the composer's perspective on the emotion content is known. However from the results obtained, distribution of music objects on the VA plane is coherent with findings available in the literature of the subject

[27], reflected by the prior concepts on musical expressiveness.

Data were carefully analyzed using correlation analysis, as well as the rough set-based processing. In the experiments conducted, the authors used the RSES application for testing the effectiveness of recognizing music genres using the rough set theory. It should be observed that reducts containing only a small number of parameters brought very similar results to those obtained in the correlation analysis, among them spectral irregularity mean and spectral brightness features may be found. The latter analysis was performed for two important features describing mood of music, namely valence and arousal. High accuracy and coverage achieved in rough sets processing confirm consistency between dimensional and label interpretation of the proposed model of mood of music.

#### REFERENCES

- [1] D. J. Levitin, *This Is Your Brain on Music: The Science of a Human Obsession*, London, Grove/Atlantic, 2008.
- [2] M. A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, M. Slaney, *Content-Based Music Information Retrieval: Current Directions and Future Challenges*, Proceedings of the IEEE, vol. 96, 4, pp. 668-696, April 2008.
- [3] M. Purgina, A. Kuznetsov, E. Pyshkin, *An Approach for Developing a Mobile Accessed Music Search Integration Platform*, Proceedings of the 2013 Federated Conference on Computer Science and Information Systems, pp. 267-273, 2013.
- [4] Z. Pawlak, "Rough sets", *International Journal of Computer & Information Sciences*, vol. 11, no. 5, 341-356, 1982.
- [5] A. Skowron, L. Polkowski (ed.), "Rough Sets" in *Knowledge Discovery vol. 1 and 2*, Physica Verlag, Heidelberg, 1998.
- [6] J. G. Bazan, M. S. Szczuka, J. Wróblewski, "A new version of rough set exploration system", *Third International Conference on Rough Sets and Current Trends in Computing RSCTC*, volume 2475, Lecture Notes in Artificial Intelligence, Malvern, PA, Springer-Verlag, pp. 397-404, October 14-16, 2002.
- [7] J. Wróblewski, "Covering with Reducts – A Fast Algorithm for Rule Generation", *Proceeding of RSCTC'98, LNAI 1424*, Springer Verlag, Berlin, pp. 402-407, 1998.
- [8] B. Kostek, M. Plewa, "Parametrization and correlation analysis applied to music mood classification", *International J. of Computational Intelligence Studies*, vol. 2, no. 1, pp. 4-25, 2013.
- [9] M. Plewa, "Automatic mood indexing of music excerpts based on correlation between subjective evaluation and feature vector", *Doctoral Thesis*, Gdańsk University of Technology. Faculty of Electronics, Telecommunications and Informatics, March 2016.
- [10] M. Plewa, B. Kostek, "Creating Mood Dictionary Associated with Music, 132 Audio Eng. Soc. Convention", Paper no. 8607, Budapest, April 26-2, 2012.
- [11] M. Plewa, B. Kostek, "Multidimensional Scaling Analysis Applied to Music Mood Recognition", 134th Audio Eng. Soc. Convention, Paper no. 8876, Rome, May 4-7, 2013.
- [12] M. Plewa, B. Kostek, "Music Mood Visualization Using Self-Organizing Maps", *Archives of Acoustics*, vol. 50, no. 4, 2015, <http://dx.doi.org/10.1515/aoa-2015-0051>
- [13] R. E. Thayer, *The Biopsychology of Mood and Arousal*, Oxford University Press, 1989.
- [14] J. A. Russel, "A circumplex model of affects", *Journal of personality and Social Psychology*, 39, pp. 1161-1178, 1980.
- [15] B. Brinker, Dinther R., Skowronek J., "Expressed music mood classification compared with valence and arousal ratings", *EURASIP J. Audio, Speech, and Music Processing*, 1, 2012, <http://link.springer.com/journal/13636/2012/1/page/1>, access 6.10.2015.
- [16] K. Hevner, "The affective value of pitch and tempo in music", *American Journal of Psychology*, vol. 49, pp. 621-630, 1937.
- [17] B. Kostek, "Content-Based Approach to Automatic Recommendation of Music", 131st Audio Eng. Soc. Convention, Paper No: 8506, New York, October 21-23, 2011.
- [18] B. Kostek, A. Kupryjanow, P. Żwan, W. Jiang, Z. Ras, M. Wojnarski, J. Swietlicka, "Report of the ISMIS 2011 Contest: Music Information Retrieval", *Foundations of Intelligent Systems, ISMIS 2011*.
- [19] P. Hoffmann, B. Kostek, "Music Data Processing and Mining in Large Databases for Active Media", *Active Media Technology, LNCS*, vol. 8610, pp. 85 – 95. Springer, 2014.
- [20] Hoffmann P., Kostek B., *Bass Enhancement Settings in Portable Devices Based on Music Genre Recognition*; *J. Audio Eng. Soc.*, vol. 63, no. 12, pp. 980 - 989, 12.2015, <http://dx.doi.org/10.17743/jaes.2015.0087>.
- [21] B. Kostek, A. Kaczmarek, "Music Recommendation Based on Multidimensional Description and Similarity Measures, *Fundamenta Informaticae*, pp. 1001-1017, DOI 10.3233/FI-2012-0000, 2013.
- [22] B. Kostek, P. Hoffmann, A. Kaczmarek, P. Spaleniak, "Creating a Reliable Music Discovery and Recommendation System", *Intelligent Tools for Building a Scientific Information Platform: From Research to Implementation*, Springer Verlag, 2013.
- [23] B. Kostek, "Music Information Retrieval in Music Repositories", Chapter 17, in: *Rough Sets and Intelligent Systems* (Skowron A., Suraj Z., Eds.), vol. 1, ISRL, 42, pp. 463-489, Springer Verlag, Berlin Heidelberg, 2013.
- [24] A. Rosner, F. Weninger, B. Schuller, M. Michalak, B. Kostek, "Influence of Low-Level Features Extracted from Rhythmic and Harmonic Sections on Music Genre Classification", *International Conference on Man-Machine Interactions*, pp. 467-473, 2013.
- [25] A. Rosner, B. Schuller, and B. Kostek, "Classification of Music Genres Based on Music Separation into Harmonic and Drum Components," *Archives of Acoustics*, vol. 39, no. 4, pp. 629-638 (2014), <http://dx.doi.org/10.2478/aoa-2014-0068>.
- [26] O. Lartillot, "MIRtoolbox 1.4: User's Manual", *Finnish Centre of Excellence in Interdisciplinary Music Research Swiss Center for Affective Sciences*, 2012.
- [27] A. Rauber, M. Frühwirth, "Automatically Analyzing and Organizing Music Archives", *5th European Conference on Research and Advanced Technology for Digital Libraries*, Springer, London 2001.