

Songs Recognition Using Audio Information Fusion

Paweł Biernacki

Abstract—The article presents information fusion approach for song classification with use of acoustic signal. Many acoustic features can contribute to correct identification of a song. Taking into consideration only one set of features may result in omission of relevant information. It is possible to improve the accuracy of identification process by means of the information fusion technique, in which various aspects of acoustic fingerprint are taken into consideration. Two sets of signal features were distinguished: one were based on frequency analysis (harmonic elements) and the other were based on multidimensional correlation ratios. An identification of a commercial was made with use of SVM and k-NN classifiers. The music audio signal database was used for assessing the effectiveness of the proposed solution. Results show an improved effectiveness of identification in relation to applying only one set of song features.

Keywords—acoustic identification, signal parametrization, information fusion, Schur coefficients, Bayes rule

I. INTRODUCTION

QUERIES pertaining to music, especially the MP3 format, are asked more and more frequently by the average Internet user [1]. Google is forming a new manner for searching data. For example, if one would like to learn the name of the author of a given image, he can enter a piece of the image into the search engine in a digital form (Google+Photos), or in case one would like to know the author and the title of a song, he only has to play a fragment of the song to the microphone built in his computer or phone (Google Ears).

Basing on the said fragment, the Google system is then able to find the song title and other crucial data. The song recognition process of audio fingerprinting is a developing field and certainly will find multiple application. A great interest in this field is exhibited by record companies that care for legal use of the songs they publish. The solution discussed herein would enable such companies a certain control over songs, for instance, by means of constant media monitoring. The use of monitoring makes it possible to determine how often and on which broadcasting stations a given musical piece is played. Moreover, in this way one can determine even the total broadcast time of the song.

At present, there are several websites that offer song recognition with use of the above described method, though most of those I have encountered are still a work-in-progress. The existing systems either require improvement due to unsatisfactory effectiveness, or could be substantially perfected, which is often announced by their very authors. Acoustic sensors (microphones) can be used to collect signals in order to determine the type and the title of a given song. Application of such sensors is attractive owing to fast installation and the lowest cost [2].

P. Biernacki is with University of Technology Wrocław, Signal Processing Section, 50-350 Wrocław, Poland (e-mail: pawel.biernacki@pwr.wroc.pl).

Selecting the relevant classification algorithm plays a critical role in song identification [1,7]. In order to classify a given musical piece effectively, characteristic features of the song or a combination of such characteristics need to be determined. The said features ought to assign the recorded music to one class explicitly and exhibit resistance to external interference (e.g. wind, noise). A small demand for data stream must be enabled in real time and a relative assignment must be ensured.

Common classification features include: moment measurement [7], eigenvectors [8], linear prediction coefficients [12], cepstral coefficients [6], levels of various harmonics [1, 4], means, variances, zero cross rates [5], etc. As can be seen, there are many types of features and choosing one of them does not guarantee successful identification.

Therefore, employing more than one group of features combined in order to improve classification seems to be the right solution. The method for combining the information (information fusion) [9] should allow for a more complete depiction of the acoustic signal model. Song recognition based on, e.g., two classifiers gives a smaller chance of misidentifying the song or not identifying any.

In the study depicted in the article I have decided on employing two types of features of an acoustic signal from a musical piece: harmonic content (spectrum analysis) and signal moments (correlation analysis).

The former feature is related to the recorded sound. The latter is based on a description of the signal with Volterra series [10], in which the signal is parametrized with use of correlation coefficients. Basing on the obtained signal parameters, a separate song recognition process is conducted; subsequently, a final decision on identification of the song is made (the song is assigned to one of the classes specified in the database) with use of the Bayesian decision rule (based on the conditional probability theorem).

II. HASHING FUNCTION IN ESTIMATING HARMONIC PARAMETERS

Hashing functions are used universally in the world of cryptographies for verifying and converting large amounts of data. Nevertheless, it is impossible to make a direct use of the version applied in cryptography in song identification, since this function is too sensitive to any alternations made to the musical work, resulting in files of thoroughly different parameters. The function described in the article is intended to imitate human hearing (Human Auditory System, Has). It aims to provide similar results for songs altered or modified to some extent (e.g., due to quantization or compression).

A schematic diagram for the algorithm is presented in Fig. 1. As can be seen, the input signal is first divided into overlapping windows, then, each of them is windowed with a Hanning window. Thus formed fragments of the signal are subjected to

the Fourier transformation. Further calculations include only the real part of the spectrum. The spectrum is divided into subbands of varying width. The widths of subsequent subbands are not equal, as their values increase in a non-linear way with the rise in frequency. In [9] a division into 32 subbands in the range of 300-3000 Hz is assumed, with the increase in the width by a coefficient of $2(1/12) = 1.06$ to every subband, which correspond to subsequent musical whole tones.

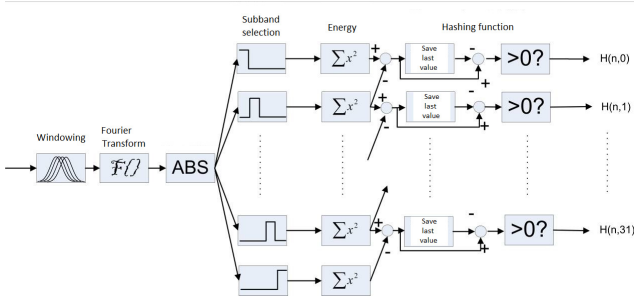


Fig. 1. Hashing function scheme.[9]

Studies confirm, that human hearing operates on a similar principle. Most often, the window length of 400 ms is adopted, with the overlap ratio of adjacent windows of 31/32. If the number of the subband is marked as m , the number of the temporary frame is marked as n and the energy of the m subband and of the n frame are marked as $EB(n, m)$, then the pixel value that corresponds to the m subband and the n frame can be determined basing on the following equation:

$$H(n, m) = \begin{cases} 1 & \text{when } EB(n, m) - EB(n, m+1) - \\ & +(EB(n-1, m) - EB(n-1, m+1)) > 0 \\ 0 & \text{when } EB(n, m) - EB(n, m+1) - \\ & +(EB(n-1, m) - EB(n-1, m+1)) \leq 0 \end{cases} \quad (1)$$

The result of this function is a sequence of binary values forming patterns dependent on the song being parametrized. The outcome of the feature extraction (1) can be seen in the picture below.



Fig. 2. Signal parametrization using hashing function.

III. REPRESENTATIVES OF THE SIGNAL WITH VOLTERRA SERIES

In order to determine the vector of parameters (features) that describe a given musical work, a multidimensional orthogonal signal representation was employed [10,11].

Estimate x_0 of the X signal can be written in the form of the multidimensional Fourier serie:

$$\hat{x}_0 = \sum_{i_1=0}^N x \rho^{i_1} x_0^{i_1} + \sum_{i_1=0}^N \sum_{i_2=i_1}^N x \rho^{i_1, i_2} x_0^{i_1, i_2} + \dots \\ \dots + \sum_{i_1=0}^N \sum_{i_2=i_1}^N \dots \sum_{i_M=i_{M-1}}^N x \rho^{i_1, \dots, i_M} x_0^{i_1, \dots, i_M} \quad (2)$$

where N is the estimation order, M its nonlinearity degree and

$$x \rho^{i_1, \dots, i_M} = (x_0, r_0^{i_1, \dots, i_M}) \quad (3)$$

are the generalized (multi-dimensional) Fourier [7] (i.e. Schur-type) coefficients. These may be interpreted as multidimensional coefficients of correlation between linear and non-linear observation of the X signal.

A. Multidimensional Audio Signal Parametrization

Given a vector $|x \rangle_T$ of samples $\{x_0, \dots, x_T\}$ of a time-series (a sound signal), observed on a finite time interval $\{0, \dots, T\}$, the signal parametrization problem can be stated as follows (see Figure 3). The estimate of the desired signal

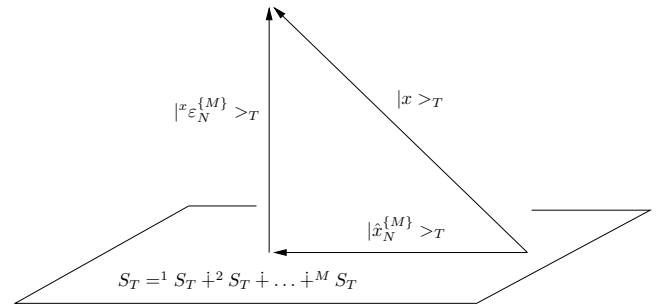


Fig. 3. The estimate $|\hat{x} \rangle_T$ of the desired signal

$$|\hat{x}_N^{\{M\}} \rangle_T \triangleq \mathbf{P}(S_T) |x \rangle_T \quad (4)$$

is the orthogonal projection of the element $|x \rangle_T$ on the space S_T spanned by the following set of the linear and nonlinear observations

$$|X \rangle_T = [|^1 X \rangle_T \quad |^2 X \rangle_T \quad \dots \quad |^M X \rangle_T] \quad (5)$$

where

$$|^m X \rangle_T = [|x_{i_1} \dots x_{i_m} \rangle_T; \quad i_1 = 0, \dots, N, \\ i_2 = i_1, \dots, N, \dots, i_m = i_{m-1}, \dots, N] \quad (6)$$

for $m = 1, \dots, M$. The orthogonal projection operator on $|X \rangle_T$ is defined as

$$\mathbf{P}(S_T) \triangleq |X \rangle_T \langle X | X \rangle_T^{-1} \langle X | T \quad (7)$$

If an ON (generalized; i.e., multidimensional) basis of the space S_T is known, the projection operator on S_T can be decomposed as

$$\mathbf{P}(S_T) = \sum_{j_1=0}^N \mathbf{P}(|r_0^{j_1} \rangle_T) + \dots + \\ + \sum_{j_1=0}^N \dots \sum_{j_M=j_{M-1}}^N \mathbf{P}(|r_0^{j_1, \dots, j_M} \rangle_T) \quad (8)$$

where $\mathbf{P}(|r_0^{j_1, \dots, j_m} \rangle_T)$ stands for the orthogonal projection operator on the one-dimensional subspace spanned by the element $r_0^{j_1, \dots, j_m}$, $m = 1, \dots, M$ of an ON basis of the space S_T . Since

$$\mathbf{P}(|r_0^{j_1, \dots, j_w} \rangle_T) = |r_0^{j_1, \dots, j_w} \rangle_T \langle r_0^{j_1, \dots, j_w} |_T \quad (9)$$

the orthogonal expansion of the estimate of the desired signal can be written as

$$\begin{aligned} |\hat{x}_N^{\{M\}} \rangle_T &= \mathbf{P}(S_T)|x \rangle_T = \sum_{j_1=0}^N |r_0^{j_1} \rangle_T \langle r_0^{j_1} |x \rangle_T + \\ &+ \dots + \sum_{j_1=0}^N \dots \sum_{j_M=j_M-1}^N |r_0^{j_1, \dots, j_M} \rangle_T \langle r_0^{j_1, \dots, j_M} |x \rangle_T \quad (10) \end{aligned}$$

The estimation error associated with the element $|\hat{x}_N^{\{M\}} \rangle_T$ is then

$$|x \varepsilon_N^{\{M\}} \rangle_T \triangleq \mathbf{P}(S_T^\perp)|x \rangle_T = |x \rangle_T - |\hat{x}_N^{\{M\}} \rangle_T \perp S_T \quad (11)$$

The estimate (4) will be called optimal (in the least-squares sense) if the norm

$$\| |x \varepsilon_N^{\{M\}} \rangle_T \| = \langle x \varepsilon_N^{\{M\}} | x \varepsilon_N^{\{M\}} \rangle_T > \frac{1}{2} \quad (12)$$

of the estimation error vector (11) is minimized for each $T = 0, 1, 2, \dots$

The multidimensional signal parametrization problem can be solved by the derivation of a (generalized) ON basis of the estimation space S_T (i.e. calculation of the orthogonal representation (the generalized Fourier coefficients) of the vector $|x \rangle_T$ in the orthogonal expansion (10)).

In order to derive the desired ON basis of the estimation space S_T , we employ (consult [7][11]) the following

Theorem 1: The partial orthogonalization step results from the recurrence relations

$$\begin{aligned} |e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T &= [|e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T + \\ &+ |r_{i_1+1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w}] (1 - (\rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w})^2)^{-\frac{1}{2}} \quad (13) \end{aligned}$$

$$\begin{aligned} |r_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T &= [|e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w} + \\ &+ |r_{i_1+1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T] (1 - (\rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w})^2)^{-\frac{1}{2}} \quad (14) \end{aligned}$$

where

$$\rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w} = - \langle e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} | r_{i_1+1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \quad (15)$$

Proof.

Let see that $|y_{i_1} \dots y_{i_q} \rangle_T \in S_{T; i_1, \dots, i_q}^{j_1, \dots, j_w}$, so (13) can be rewritten as

$$|e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T = |y_{i_1} \dots y_{i_q} \rangle_T - \mathbf{P}(S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w}) |y_{i_1} \dots y_{i_q} \rangle_T \quad (16)$$

Because $|r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \perp S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w}$ and $|r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \in S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w}$ then

$$\mathbf{P}(S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w}) = \mathbf{P}(S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w-1}) \oplus \mathbf{P}(|r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T) \quad (17)$$

So the equation (16) can be rewritten as

$$\begin{aligned} |e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T &= |e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T - \\ &+ |r_{i_1+1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \langle r_{i_1+1, \dots, i_q+1}^{j_1, \dots, j_w} | y_{i_1} \dots y_{i_q} \rangle_T \quad (18) \end{aligned}$$

Because $|\widehat{e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1}} \rangle_T = |y_{i_1} \dots y_{i_q} \rangle_T - |\widehat{y_{i_1} \dots y_{i_q}} \rangle_T$ and $|\widehat{y_{i_1} \dots y_{i_q}} \rangle_T \in S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w-1}$ and $|r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \perp S_{T; i_1, \dots, i_q+1}^{j_1, \dots, j_w-1}$ then

$$\begin{aligned} \langle r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} | y_{i_1} \dots y_{i_q} \rangle_T &= \langle y_{i_1} \dots y_{i_q} | r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T = \\ &= \langle y_{i_1} \dots y_{i_q} | r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T - \langle \widehat{y_{i_1} \dots y_{i_q}} | r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T = \\ &= \langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} | r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T = \\ &= - \langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} | \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T > \frac{1}{2} \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w} \quad (19) \end{aligned}$$

Hence the conclusion that

$$\begin{aligned} |e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T &= [|e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T + |r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w}] \\ &\langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} | \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T > \frac{1}{2} \\ &\cdot \langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w} | \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T^{-\frac{1}{2}} \quad (20) \end{aligned}$$

Exploiting the identity $\langle e_{i_1, \dots, i_q}^{j_1, \dots, j_w} | e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T = 1$ we receive

$$\begin{aligned} \langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} | \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T &< \frac{1}{2} \langle \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w} | \varepsilon_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T^{-\frac{1}{2}} = \\ &(1 - \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w})^{-\frac{1}{2}} \quad (21) \end{aligned}$$

Finally

$$\begin{aligned} |e_{i_1, \dots, i_q}^{j_1, \dots, j_w} \rangle_T &= [|e_{i_1, \dots, i_q}^{j_1, \dots, j_w-1} \rangle_T + |r_{i_1, \dots, i_q+1}^{j_1, \dots, j_w} \rangle_T \rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w}] \\ &\cdot (1 - (\rho_{i_1, \dots, i_q; T}^{j_1, \dots, j_w})^2)^{-\frac{1}{2}} \quad (22) \end{aligned}$$

Similarly, the second relation in (14) can be found [7]. The relations above make it possible to construct an orthogonal parametrization (de-correlation) filter, operating directly on the signal samples. The above recurrence relations (14) solve the problem of the real-time derivation of the (generalized) ON basis of the estimation space. The diagram of the signal parametrization filter is presented in Figure 4.

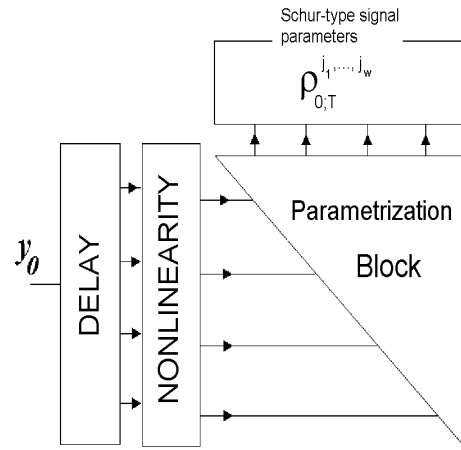


Fig. 4. Nonlinear orthogonal signal parametrization filter

The Schur coefficients (3) can be used for audio signal parametrization. Entire song audio signal parametrization procedure can be done in the tree steps:

TABLE I
SIGNAL IDENTIFICATION EFFECTIVENESS FOR SINGLE
FEATURES

	Harmonic features	Correlation features
K-NN (k=3)	73.91	75.72
K-NN (k=5)	72.45	77.39
MGC	73.28	75.35
SVM linear kernel	74.61	76.08
SVM Gauss kernel	74.06	75.55

- 1) dividing audio signal into the one second long pieces

$$y_{T_k} = \{y_T(n); n = k * f_p, k * f_p + 1, \dots, k * 2 * f_p - 1\} \quad (23)$$

where k means k -th signal section and $k = 0, 2, \dots, K - 1$.

- 2) computing Schur coefficients (3) vector for each piece
3) construct ID matrix for the audio signal (using Schur coefficients vectors as its rows)

$$YM = \begin{pmatrix} \rho_{0;T_0}^0 & \dots & \rho_{0;T_1}^{j_1, \dots, j_w} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{0;T_{K-1}} & \dots & \rho_{0;T_K}^{j_1, \dots, j_w} & \dots \end{pmatrix} \quad (24)$$

Matrix (21) can be used for creating the base of the songs and for classification.

IV. SONGS CLASSIFICATION

The calculated sets of features (4), (21) constitute a signal database for classifying a given musical work to one of the formerly specified classes of songs (standard classes are in the database created with use of former recordings). The study used information fusion on the decision-making level. Conclusive ranking of the object (song) is based on connecting identification decisions for individual features of the given object. It aims to improve effectiveness of the diagnosis and to limit possible miscategorization. In order to obtain good classification results on the level of individual features, different classifiers were examined: k-NN (the k-nearest neighbors algorithm), MGC (multivariate Gaussian classifier), and SVM (support vector machines). The classifiers were examined in order to determine the one with the highest song identification effectiveness for individual groups of features.

Table I shows the obtained identification results for single groups of features and examined sorters. Finally, SVM classifier was selected for harmonic features and k-NN classifier was chosen for correlation features.

In order to produce the final result of the song data classification, the Bayes rule was applied.

$$P(Y|X_1, X_2) = P(Y) \sum_{i=1}^2 P(X_i|Y) \quad (25)$$

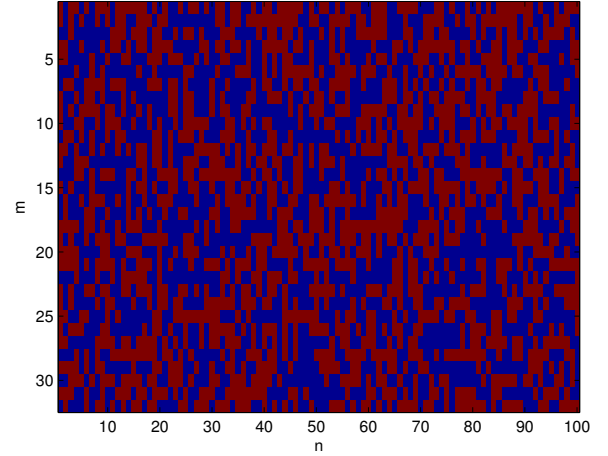


Fig. 5. Hash 'fingerprint' for song 1

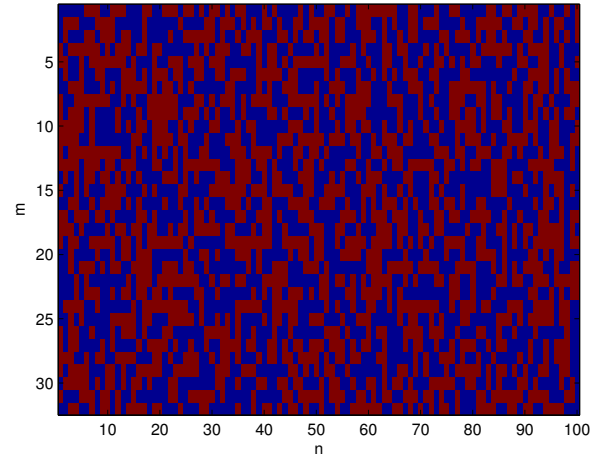


Fig. 6. Hash 'fingerprint' for song 2

where X_1 is a set of harmonic features, X_2 is a set of correlation features, Y is an outcome of the ranking. The elements X_1 and X_2 were proven to be independent.

V. SIMULATIONS

The formerly created database of musical pieces was used for simulation (100 musical works in the database). The songs selected for the parameterization were of the best possible quality: sampling rate: 44100Hz, quantization: 16 bit, file format: wav in the mono version. Every song forms an object class in the database (i.e., one song stands for several realizations of this piece, characterized by a different volume, disruptions, recording from the radio, the TV, etc.)

One-second long pieces of acoustic signal recordings were used for the identification process. The signal was divided into fragments with 2048 samples ($f_p = 16$ kHz, 16-bit word), windowed (Hamming window) and processed according to the above schemes. In pictures 4 and 5 one can see examples of hash fingerprints for two different songs.

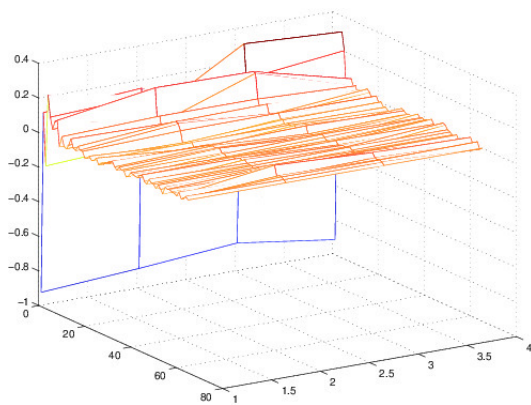


Fig. 7. Matrix of Schur coefficients - song1.

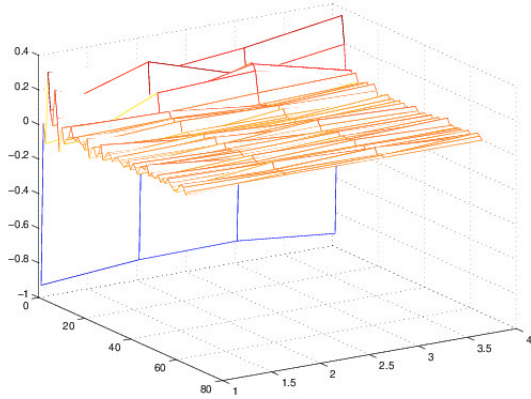


Fig. 8. Matrix of Schur coefficients - song2.

TABLE II
EFFECTIVENESS OF THE RANKING OF THE SONGS
DEPENDING ON THE APPLIED HARVEST OF FEATURES

	Effectiveness
Harmonic features	74.61
Correlation features	77.39
Information fusion	84.02

In pictures 6 and 7 the matrices of Schur coefficients (6) describing correlation ratios of acoustic signals can be seen.

It was observed, that within one class these matrices are not different as much as in the case of signal spectra. Here, the impact of varying conditions of the recording is smaller, which should naturally improve the effectiveness of the final song identification. Table II shows results of the ranking of musical works depending on the applied harvest of features.

The following conclusions can be drawn basing on the data provided in Table 2:

- due to the large variance of harmonic features within a single class of songs, the effectiveness of identification based on these parameters only is the lowest. It would be possible to improve the result by distinguishing certain signal subbands with greater inter-class differences .
- Identifying signals with use of multidimensional correlation features is more effective than with harmonic features. This is mainly due to the smaller variance

of correlation parameters within one class. It could be possible to improve the result by selecting proper Schur coefficients that would best describe a given class of signals. It is possible to apply the mutual information technique [3] in order to select the most characteristic parameters of a given class of signals

- The use of combined information from the diagnoses of both features resulted in the expected improvement in the effectiveness of song classifications. An almost 6% rise is a very good result. Should the findings specified above be applied, the results ought to surpass those diagnosed above. Of course, one should take into consideration the extra computing power needed for vital calculations, particularly when the suggested solution is supposed to take place in real time.

VI. CONCLUSION

In the article an application of an approach based on the information fusion theory in classifying musical works was proposed. Two types of parameters describing the acoustic signal were chosen: harmonic features (with hashing function) and multidimensional correlation features. On this basis a fragmentary identification of a given class of songs was made. Next, the final identification of the object was performed, using the Bayes rule for classification. The findings proved a significant improvement in the effectiveness of identification with information fusion in relation to the inference based only on a single feature of a musical piece.

REFERENCES

- [1] T. Damarla, T. Pham, D. Lake , "An algorithm for classifying multiple targets using acoustic signature" , *Proceedings of SPIE signal processing, sensor fusion and target recognition*, pp 421427, 2004
- [2] B. Guo, M. Nixon, "Gait feature subset selection by mutual information" , *IEEE Trans Syst Man Cybern Part A* 39(1):3646, 2009
- [3] M. Munich, "Bayesian subspace methods for acoustic signature recognition" , *Proceedings of the 12th European signal processing conference*, pp 14, 2004
- [4] G.D. Whitaker, "An Overview of Information Fusion", DERA, British Crown Copyright 2000
- [5] S. Schetzen, "The Volterra and Wiener Theories of nonlinear systems", John Wiley & Sons, New York, 1980
- [6] P. Biernacki, "Orthogonal Schur-type solution of the nonlinear echo-cancelling problem", *IEEE ICECS 2007*, Piscataway, NJ, 2007
- [7] P. Biernacki, J. Zarzycki, "Multidimensional Nonlinear Noise-Cancelling Filters of the Volterra-Wiener Class", *Proc. 2-Nd Int. Workshop on Multidimensional (nD) Systems (NDS-2000)*, ed. Inst. of Control and Comp. Eng. TU of Zielona Gora Press, Czochka Castle, pp. 255-261, 2000
- [8] P. Biernacki, "Strategies for adaptive nonlinear noise reduction Volterra-Wiener filter structure selection", *Proc. ICECS 2006*, Nice, France, December 10-13 2006, cd-rom
- [9] S. Haykin, "Adaptive filter theory", Upper Saddle River, Prentice Hall 1996
- [10] D.T.L. Lee, M. Morf, B. Friedlander, "Recursive Least-Squares Ladder Estimation Algorithms", *IEEE Trans. on CAS*, Vol. 28, pp. 467-481, 1981
- [11] J. Zarzycki, "Multidimensional Nonlinear Schur Parametrization of NonGaussian Stochastic Signals - Part One: Statement of the Problem", *Multidimensional Systems and Signal Processing (MDSSP) Journ.* (Kluwer Academic Publishers), Vol. 15, No. 3, pp. 217-241, 2004.