# Rule-based classification of musical genres from a global cultural background

Jakob Abeßer°, Hanna Lukashevich°, Christian Dittmar°,
Paul Bräuer•, and Fabienne Krause•

°Fraunhofer IDMT
Ehrenbergstr. 31, 98693 Ilmenau, Germany
{abr,lkh,dmr}@idmt.fraunhofer.de
http://www.idmt.fraunhofer.de

• Piranha Musik & IT
Bergmannstr. 102, 10961 Berlin, Germany
{paul.braeuer,fabienne.krause}@piranha.de
http://www.piranha.de

**Abstract.** In this paper, we present a rule-based classification framework that allows to describe musical genres by means of different musical properties. We compare the presented approach with a state-of-the-art genre classification system including feature extraction as well as different feature selection and feature space transformation methods. A set of 6 music genres from a global cultural background was selected to investigate the applicability of the presented method for computational ethnomusicology (CE). Four different experiments were carried out to investigate different influence factors on the achievable classification accuracy.

**Key words:** genre classification, rule-based, expert system, computational ethnomusicology

## 1 Introduction

Automatic classification of music is a common need. Today, more music than ever is accessible and in fact listened to. Automated classification of music is an interdisciplinary research discipline which now exists for several decades [23] [25]. Music classification has become more necessary and difficult at the same time. Nowadays, it has to meet the challenge of a growing and diverse global range of music. The constant merging of different musical influences into new styles, which sometimes are more than just hybrids, has to be taken into account. While rough classification is achieved quickly and easily [19], detailed metadata about music recordings is much more problematic to obtain, as lies in the symbolic nature of music as a cultural phenomenon [5] [34]. This paper is organized as follows. After outlining the goals and challenges of this publication in Sec. 2, we provide an extensive overview of related work in Sec. 3. Then, we explain the rule-based classification approach that we follow in this paper in Sec. 4. In Sec. 5, we explain the performed experiments and discuss the obtained results. Finally, Sec. 6 provides conclusions of this work and Sec. 7 gives a short outlook.

## 2 Goals & challenges

To achieve a valid genre classification, the growing and diverse range of music can not be neglected. Especially in the global context of what formerly was called "world music", clearly separable genres have given way to a wide variety of different forms of musical fusion. Globalized "western" styles are combined with local traditional styles, different traditions are blended, new cross-cultural styles emerge[1]. Automatic classification has to meet increasingly well defined needs, be it in the area of private listening, public events or functional music (muzak, movie soundtracks, advertisement etc.).

Since the description of genres is no longer only a matter of marketing, an appropriate classification system has to mirror the structure of tastes, peer groups and musical styles, fragmented as they are. At the same time it has to be concise to not produce an overload of classes, which ultimately do not provide useful information. In this paper, we will suggest a meaningful form of a single-label classification. By testing its limitations on fusion styles however, we also want to promote and display the need of a multi-label approach as a complementary in a truly encompassing framework of music classification that provides refined, perceptible information.

## 3 Previous approaches

Commonly, automatic genre classification is approached by statistical methods to evaluate previously extracted audio features [36] or by expert systems [31]. In most publications, the focus is on western music genres such as Rock, Jazz or Classical music. Only a few experiments have been conducted so far with regional genre selections such as traditional Malay Music [12]. [37] gives a broad overview over publications dealing with the application of Music Information Retrieval (MIR) techniques for the analysis of different regional music styles in the context of CE. While dealing with music from worldwide origins, multiple stylistic influences on music pieces have to be taken into account. The inevitable problem of multiple genre labels for single songs have been addressed amongst others in [27]. The authors introduced the multi-domain labeling approach from genre classification that takes both the temporal structure as well as three different musical dimensions into account.

The use of expert systems for music genre classification was rarely covered so far. Guidelines for the design of hierarchical genre taxonomies as well as the use

[1] In fact, a lot of music, we consider original today, is a blending of very different traditions who only over time became "authentic". Take Samba as an example, a combination of Christian and military music, percussion styles of African slaves and South American instrument handcraft.

of musical descriptors such as instrument presence (e.g. brass, strings), voice type (e.g. soft), rhythm or tempo to formulate differences and relationships between musical genres were presented in [31]. It was pointed out that musical high-level features such as instrumentation could not yet be extracted reliably [31] [4] [33]. This is an important requirement for an automatic evaluation of a compiled set of rules. Furthermore, expert systems are considered as being expensive to implement and difficult to maintain in face of the potential merging and splitting of genres into multiple sub-genres within future developments of music [4].

Some work has been conducted in automatic genre estimation in symbolic music data and MIDI files [28], the encouraging classification rates of 98% have been obtained for three root genres. In [30] a library for the extraction of high-level features from MIDI files has been presented. Additionally the authors proved the importance of instrument identification for genre recognition tasks [29]. They showed that the instrumentation took about 40% of weighting assigned to different feature groups. Unfortunately, within the music genres from a global cultural background, providing MIDI files is hardly possible, especially regarding instrumentation.

Humans are used to describe musical genres using sets of simple rules applied to high-level musical properties. One might think of the characteristic off-beat accentuation in Reggae music, the typical distorted guitar for Rock music or the presence of numerous percussion instruments for different Latin American genres. Rule-based expert systems have been used for generating music in the style of certain classical composers by using generative grammars related to melody and harmony [8]. The authors presented an algorithm that uses a given knowledge base in terms of melody-related grammar rules related to compare the works of four Italian composers of the 17th century based on musical properties such as interval structure, the occurrence of melodic figures and rhythmic metrics. Anglade et al. [2] presented an automatic genre classification approach based on frequent chord sequences. To obtain the rules characterizing the genres they utilized a first-order logic decision three induction algorithm TILDE [7].

The majority of the above mentioned studies work within a confined frame of e.g. western art music of one century. In a global context however, not only the values of musical properties, but the properties themselves differ from style to style. Music of Sub-Saharan Africa for example, can not be described in terms of chord structure. According to [22] in the field of Afro-Caribean music, the western misunderstanding of defining global styles by melodic terms is as common as it is wrong. Questions of genre description are constantly disturbed by conflicting interests between musicology and the music industry [18]. Even in scientific circles it is an ongoing discussion, how to systematize the selection of the appropriate features for the classification of music genres. With cultural and musicological questions arising, it is questionable whether a widespread consensus will ever be reached in this matter [15].

## 4 New approach

### 4.1 Rule-based framework

In this paper, we present an extension of the rule-based classification approach previously introduced in [1]. The basic terms *concept*, *class*, *observable* and *property* are introduced in the following subsections. Afterwards the algorithmic steps of assigning a relevance value to each defined class for a given piece of music are explained in detail.

**Concepts & classes** The term concept represents a general approach to categorize music. A number of classes represent the semantic entities of a concept. In this paper, we investigate 6 classes of the concept "music genre".

**Observables** In this paper, songs and song segments are attributed regarding different musical properties using so-called *observables*. Most often, they correspond to musical high-level features that are commonly used in Music Information Retrieval (MIR) literature. Examples are the "tonal complexity of the melody", the "variety of rhythm" or the "presence of guitar". In contrast to low-level and mid-level audio features, observables refer to musical segments or complete songs instead of fixed temporal frames of equal lengths. Furthermore, they do not have an infinite range of values but a limited number of possible values that have been defined before and that are exclusive to each other.

The assortment of *observable categories* and of appropriate observable values in this paper have been initially developed by musicologists, who proved the significance of each category for the investigated genres in various academic user studies in cooperation with musicology students at the Humboldt University of Berlin. The selected categories contain superior structures of global music dimensions such as instrumentation, vocals, rhythm, and melody/harmony. Some of the observables rely on more or less objective facts like number of instruments, bar measure or the gender of the vocalist, others are more perceptual such as the perceived tempo or the recording quality.

In this paper, we bypass the error-prone step of automatic feature extraction and use the observable annotations by musicologists as input for the classification algorithm introduced in the following sections. Our motivation is to perform a selection of initial classification experiments and keep the focus on the classification method itself.

**Properties** Properties are rules that characterize songs of a given genre. A property translates one aspects of the musical description of a class into an explicit boundary condition on the value of a feature. Each property consists of an *observable* $O$, a specific *relation* $R$ such as "is greater than" or "is equal to" and a definite *threshold* $V$.

We discern *mandatory properties* $P_M$, which strictly need to be fulfilled, and *frequent properties* $P_F$, which are not compulsory for modeling a class. According

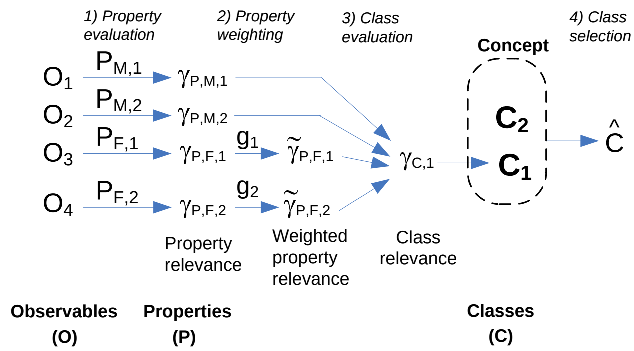



**Fig. 1.** Rule-based classification framework.

to their importance, each frequent property can be associated with a *weighting factor* $0 \leq g \leq 1$. Each property can either be *omnipresent* ($P_O$) and thus constantly valid or *conditional* ($P_C$). For instance, if a property depends on the presence of a certain instrument, it only needs to be considered, if this instrument truly appears in a song. The presence of the instrument thus can be formulated as a condition for this property.

### 4.2 Relevance values

Each property is evaluated by computing a *property relevance* $\gamma_P = r(O)$ $(0 \leq \gamma_P \leq 1)$. The function $r(O)$ is a mathematical expression of the property relation and translates the feature value $F$ calculated for a given song into a measure to what extend a property is fulfilled or not. This results in a high or low value of $\gamma_P$. The *class relevance* $(0 \leq \gamma_C \leq 1)$ provides a aggregated measure for the importance of each class for the musical labeling of a song. Complying with the different types of properties that are defined for a class, the class relevance is derived from all property relevance values by means of two algorithmic steps as illustrated in Fig. 1.

At first, the *weighted property relevance* $\tilde{\gamma}_{P,F,i}$ is derived by scaling the actual property relevance $\gamma_{P,F,i}$ of the frequent properties $P_F$ according to its weighting factor $g$ as

$$\tilde{\gamma}_{P,F,i} = 0.5 - 0.5g_i + g_i\,\gamma_{P,F,i}\ . \tag{1}$$

This ensures that properties of low importance can only contribute to a low degree to the overall importance of a class. The result of this scaling is illustrated for a minor important and a very important property on the right side of Fig. 2. If all mandatory properties $P_M$ are fulfilled, the class relevance $\gamma_C$ for class $k$ is derived by a weighted mean over all weighted property relevance values,

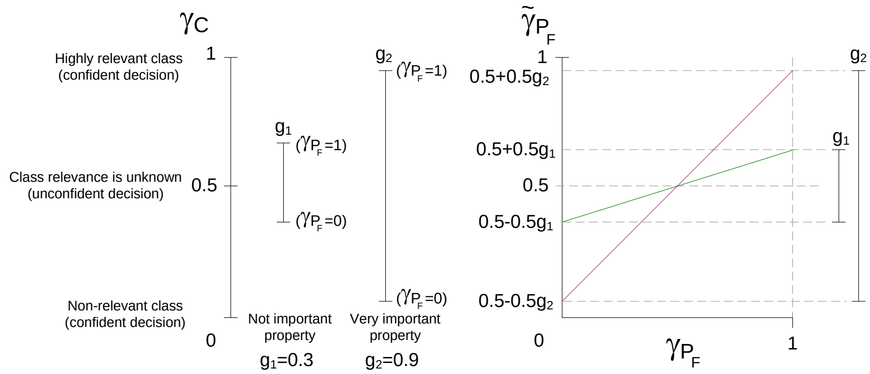



**Fig. 2.** The relevance value of each property is scaled according to its weighting factor. This affects its influence onto the class relevance. Two examples are illustrated.

otherwise it is set to zero:

$$\gamma_{C,k} = \frac{1}{N_{P_F}} \sum_{i=1}^{N_{P_F}} \tilde{\gamma}_{P,F,i} \text{ if } \gamma_{P,M,j} = 1 \; \forall \; P_M \tag{2}$$

If we consider the different scaling in Eq. 4.2, the value range $[0, 1]$ should be interpreted as follows. Confident decisions for a high or low importance of a class can be derived from values of $\gamma_C$ around 1 or 0. A value of $\gamma_C \approx 0.5$ implies that no confident decision based on the importance of all class properties can be taken. This is illustrated on the left side of Fig. 2. Finally, the class

$$\widehat{C} = k \Leftrightarrow \arg\max_k \gamma_{C,k} \tag{3}$$

is selected as the genre of the song.

# 5 Evaluation

## 5.1 Genre selection

Seemingly, the ideal test set for a classification system would consist of typical and clear examples of the given music genres. The situation on today's music markets, however, is different (see Sec. 2 for details). In today's fusion music the original styles have not entirely vanished and classification frameworks can help to trace them. This paper tests the possibilities of a one-genre-only classification in that regard as well as its limitations. The genres chosen are *Reggae*, *Hip Hop*, *Gypsy Brass*, *Klezmer*, the Mozambican *Marrabenta* and the Brazilian *Forró*. Our choice of 24 music pieces of the wide ranged catalogue of *Piranha Musik & IT* reflects a variety of challenges.

While Northern Brazilian Forró is a very unique style (as well within our selection as in Brazil), Klezmer and Gypsy Brass (Roman brass music from the Balkans) are very similar to each other and partly derive from the same musical origins[2]. Marrabenta is in some terms unique within our selection, but shares certain features with regard to the instrumentation and the playing style with Reggae, while Hip Hop and modern forms of Reggae share the use of electronics and vocal styles. All three of the latter styles also share a strong influence of western drums, scales and music structure. Within the genres, the song selection features songs from different artists with different degrees of fusion and modern adaptation.

### 5.2 Annotation process

The *Annotation Tool* is a software application, used by the musicologists or music experts in order to attribute values to song segments. It offers a three window view that displays a list of songs at the bottom, a colored segment-stream and lists of observable categories such as “Melody / Harmony”,“Rhythm”, and “Instrumentation” in the middle, and a visualization of existing observable annotations in a time-line at the top. Observable values can be allocated to a particular segment of a song (like the highlighted gray one in Fig. 5 in the appendix).

To speed up the annotation process we apply an automatic song segmentation algorithm. The segments of each song are entities with a varied length of time, that are predetermined by an algorithm in a following way. Our approach is following the ideas successfully applied for Speaker Segmentation [9]. Firstly, based on low-level and mid-level features, we detect possible segment change points while computing Generalized Likelihood Ratio (GLR) for pairs of adjacent windows of the same size and subsequently shifting these windows by a fixed step. The local minima of the GLR measure point out possible change points candidates. Secondly, we utilize Bayesian Information Criterion (BIC) to group segments, belonging to the same segment state. Here with a segment state we understand a set of time segments sharing similar distribution of the feature vectors. We use BIC in an agglomerative clustering scenario, where initially derived segment candidates are firstly considered as individual segment states. At each iteration two closest segment states with highest and positive BIC are merged. The procedure is stopped when all BIC measures between segment states become negative.

Duration, number and combination of the automatically detected segments can be changed manually for each observable individually. Often it turned out to be meaningful to create a new specific segmentation for a chosen category. For instance, the structure of a song regarding its rhythm often differs from its instrumentation or vocal structure. Although the results ultimately provide only a classification of complete songs, the high degree of detail in our meta-data (i.e. annotations) made it possible to discern songs and genres by their temporal structure.

[2] The Jewish as well as the Roman Diaspora led to musical blendings, both inspired by the traditions of the Semite as well as the Oriental region.

### 5.3 Sets of rules

To test the rule based approach on global music catalogues, the musicologists have developed a set of rules for each of the 6 music genres introduced in Sec. 5.1. These sets of rules are depicted in the Appendix 8. For each genre all defined properties are listed in forms of their property type ((m)andatory, (f)requent, (c)onditional, (o)mnipresent, compare Sec. 4.1), the observable $O$, relation $R$, threshold $V$ and property weighting $g$. The property thresholds as well as the property weightings have been initially determined by the musicologists by a separated training dataset containing further songs of the applied genres that have not been applied for the experiments in this paper.

The experiment is limited to 6 genres and 4 songs per genre due to the goal of detailed analysis. The focus of this research is not as much a modeled multitude of meta-data, as a careful testing of the sets of rules - i.e. the choice of observables and their weighing (see Sec. 5.6. for details). For the same reason - to overview the effects of every single rule in detail - the sets of rules confined to few observables.

The rules were chosen in accordance with musicological research. For example it is a common musicological finding, that rhythm and instrumentation are central to discriminating global music styles. This was backed by the Evaluation of musicologists and students of musicology at the Humboldt University of Berlin in early 2009. In several pretests using popular Blues and Swing pieces as well as the catalogue of Piranha Musik & IT, this could be affirmed and therefore was a main criterion for the choice of the set of rules. Still, not only the most important observables were chosen, but a broad variety including rather marginal parameters as for example the presence of brass instruments in Reggae. The aim of the experiment was not to create 6 perfectly working sets of rules, but to display the different possibilities and problems of a rule-based approach.

Although the observable values have been annotated manually, the selection of properties for this rule-based approach reflects that the long term aim is to almost exclusively use automatic MIR algorithms for the extraction of the properties. Therefore, we used the properties, which can be automatically retrieved already nowadays or which are estimated to be retrievable within a small number of years. Only with this limitations, our experiments based on the rule-based classification approach become valuable for future designs paradigms of MIR and automated classification.

### 5.4 Reference classification system

We compare the presented rule-based approach with a state-of-the-art genre classification system (reference classification system) including feature extraction as well as different feature selection and feature space transformation methods.

**Feature extraction** For the experiment 5 described in Sec. 5.6, we utilize a broad palette of acoustic features and several mid-level representations [6]. These desciptors are computed on 5.12 seconds excerpts of the observed evolution of low-level features. With the help of mid-level representations, timbre texture [36] can be captured by descriptive statistics as well as by including additional musical knowledge.

*Timbre.* Although the concept of timbre is still not clearly defined with respect to music signals, it proved to be very useful for automatic music signal classification. To capture timbral information, we use Mel-Frequency Cepstral Coefficients (MFCC), Octave Spectral Contrast (OSC) [24], Audio Spectrum Centroid (ASC), Spectral Crest Factor (SCF), Spectral Flatness Measurement (SFM) and others. In addition, modulation spectral features [3] are extracted from the aforementioned features to capture their short term dynamics. We applied a cepstral low-pass filtering to the modulation coefficients to reduce their dimensionality and decorrelate them as described in [11].

*Rhythm.* All rhythmic features used in the current setup are derived from the energy slope in excerpts of the different frequency-bands of the Audio Spectrum Envelope (ASE) feature. These comprise the Percussiveness [38] and the Envelope Cross-Correlation (ECC). Further mid-level features [11] are derived from the Auto-Correlation Function (ACF). In the ACF, rhythmic periodicities are emphasized and phase differences annulled. Thus, we compute also the ACF Cross-Correlation (ACFCC). The difference to ECC again captures useful information about the phase differences between the different rhythmic pulses. In addition, the log-lag ACF and its descriptive statistics are extracted according to [21].

*Tonality.* Tonality descriptors are computed from a Chromagram based on Enhanced Pitch Class Profiles (EPCP) [26], [35]. The EPCP undergoes a statistical tuning estimation and correction to account for tunings deviating from the equal tempered scale. Most important, the so-called symmetry model, a pitch-space representations as described in [17] are derived from the Chromagram as mid-level features. Their usefulness for genre classification has been shown in [20]. The model provides an analytic description of aspects of musical consonance and dissonance, as well as functional relationships between probable notes.

**Feature selection (FS) & feature-space transformation (FST)** The following feature selection and feature space transformation techniques have been utilized to reduce the dimensionality of the feature space.

*Inertia Ratio Maximization using Feature Space Projection (IRMFSP).* IRMFSP was proposed in [32]. This FS algorithm is motivated by the ideas similar to Fisher's discriminant analysis. During each iteration of the algorithm, we look for the feature maximizing the ratio of between-class inertia to the total-class inertia. To avoid the next chosen feature to provide the same information on the next iteration, all features are orthogonalized to the selected one. In this evaluation we use the ISMFSP algorithms with the modifications proposed in [14].

*Linear Discriminant Analysis (LDA).* LDA is one of the most often used supervised FST methods [16]. It is successfully applied as a pre-processing for audio signal classification. Original feature vectors are linearly mapped into new feature space guaranteeing a maximum linear separability by maximization of the ratio of between-class variance to the within-class variance. This mapping is conducted by multiplying the original $K \times N$ dimension feature matrix $\mathbf{X}$ with the transformation matrix $\mathbf{T}$. Reducing the dimension of the transformed feature vector from $N$ to $D \leq N$ is achieved by considering only the first $D$ column vectors of $\mathbf{T}$ for multiplication.

**Classification** We applied the following four well-known methods for the purpose of classification.

*Support Vector Machines.* A Support Vector Machine (SVM) is a discriminative classifier, attempting to generate an optimal decision plane between feature vectors of the training classes [39]. Commonly for real-world applications, classification with linear separation planes is not possible in the original feature space. The transformation to the higher dimensional space is done using the so called kernel trick (we applied the Radial Basis Function kernel in this paper). The key idea of the kernel trick is to replace the dot product in a high-dimensional space with a kernel function in the original feature space. Transformed into a high-dimensional space, non-linear classification problems can become linearly solvable.

*Gaussian Mixture Models.* Gaussian Mixture Models (GMM) are commonly used generative classifiers. Single data samples of one class are interpreted as being generated from various sources and each source is modeled by a single multivariate Gaussian. The probability density function (PDF) is estimated as a weighted sum of the multivariate normal distributions. The parameters of a GMM can be estimated using the Expectation-Maximization algorithm [10].

*Naive Bayes Classifier.* Naive Bayes classifier (NB) is a simple probabilistic classifier. NB uses a strong assumption of feature dimensions being statistically independent and thus takes into account only means and variances over the feature dimensions for all training data of the class. Recently, applicability and efficiency of NB classifiers were discussed in detail in [40].

*$k$-Nearest Neighbor.* With $k$-Nearest Neighbor (kNN), the classification is based on the class assignment of the closest training examples in the feature space [13]. We used the Euclidean distance here. This type of discriminative classifier is also referred as instance based learning. The level of generalization of kNN can be tuned by adjusting the number of nearest neighbors $k$ taken into account.

### 5.5 Experimental setup

For the purpose of evaluation, a data-set consisting of 24 songs (see Appendix 8) was extensively annotated using approx. 65 observables that cover different aspects of instrumentation, rhythm, melody, and harmony. Therefore, for each of the 6 genres Reggae, Forró, Hip Hop, Klezmer, Marrabenta, and Gypsy Brass, 4 songs have been manually annotated by the musicologists. Each observable annotation consist of the *segment boundaries* and a specific *observable value*. We are aware of the fact that this small dataset does not allow strong generalizations. However, we aim to carry out a couple of initial experiments within this paper to motivate further research within this field.

For the purpose of automatic genre classification we assembled a database of 200 songs of the above mentioned six genres. A single genre label was manually assigned to each song. Song segmentation and multi-labeling have not been used. Due to the restricted size of the database, the classification results have been derived in a leave-one-out scenario.

### 5.6 Experiments & Results

**Experiment 1 - Is it meaningful to use mandatory properties?** Following the different types of properties introduced in Sec. 4.1, we aim to investigate whether the use of mandatory properties is a reasonable approach to separate different classes from each other. All of the 6 generated set of rules contain a certain number of mandatory properties (see Appendix 8). In this experiment, we compare two scenarios to investigate the utility of mandatory properties. First, we apply the properties in such way as it is originally stated. Second, we convert all mandatory properties to frequent properties each with a (maximum) weighting factor of $g = 1$ to comply with their high importance for the description of a class. Classification accuracies of 58.33% and 83.33% for scenario 1 and 2 show a significant increase if mandatory properties are converted to frequent properties with a maximum property weighting of $g = 1$. Even with future improvements in mind, strict mandatory properties seem to work contraproductively. Hence, we only use frequent properties in the following experiments.

**Experiment 2 - What about the optimal number of properties to define a genre?** In the second experiment, we modified the maximum number of properties per genre from 1 to 24. Again, we used the modified set of rules with all mandatory properties being converted to frequent properties ($g = 1$). As depicted in Fig. 3a (Exp. 2), a saturation effect in terms of the accuracy can be seen for $N \geq 11$. The set of rules for Reggae contains 11 properties, all other sets contain a larger number. We interpret the results in such way that a larger number of properties generally represents more aspects of a music genre and thus increases the classification accuracy. On the other hand, different numbers of properties for different genres does not seem to affect the classifier.

**Experiment 3 - What's the influence of erroneous annotations?** In practice, the annotations of different observables will not be provided by educated musicologists but by algorithms that are specifically trained for the automatic annotation of single observables. These algorithms are error-prone and thus can result in incorrect observable values. We simulate this effect within two experiments. First, we randomly select a certain ratio $P_{err}$ of all manual annotations for each song and modify the annotated observable value towards a false value. Second we modify both the observable value as well as the segment boundaries of each annotated segment. As it can be seen in Fig. 3b (Exp. 3a) and Fig. 3c (Exp. 3b), observable annotation with false values but correct segment boundaries still contains semantic information. Even if all segments are modified this way, the accuracy is still about 37%. If the segment boundaries are furthermore changed randomly, the accuracy drops towards the random decision baseline of approx. 16.6%. The same behavior can be recognized for the first scenario of experiment 1 as depicted in Fig. 3d (Exp. 3c).

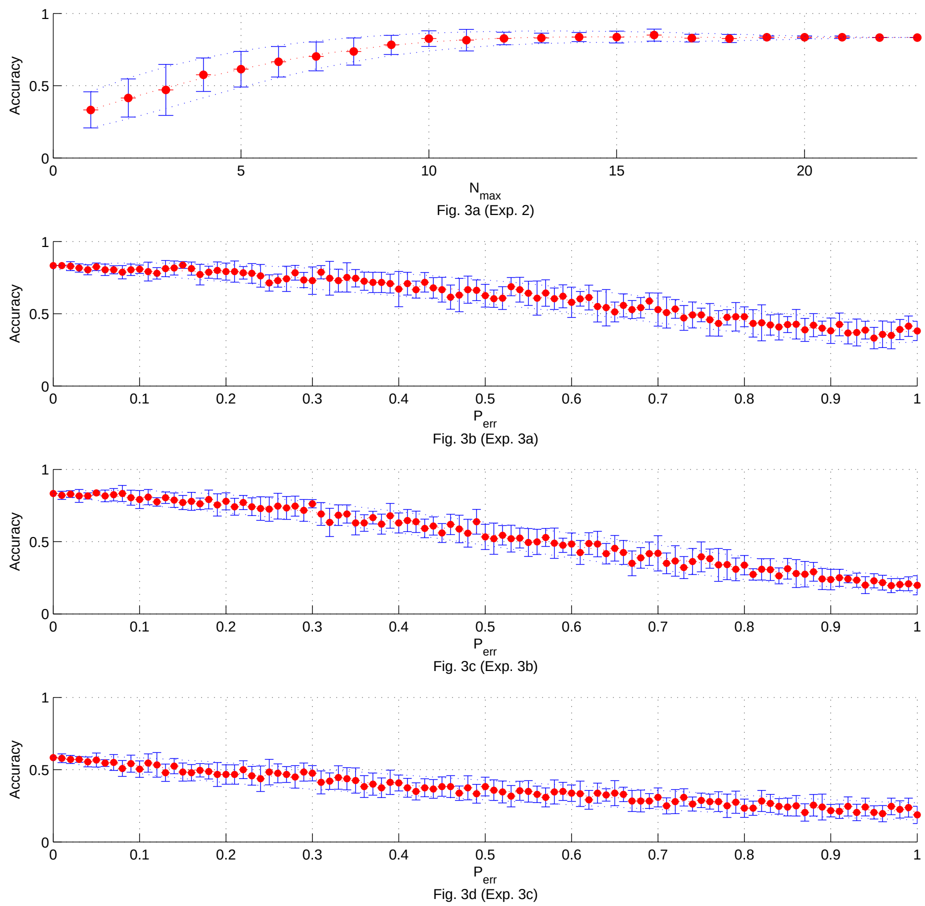
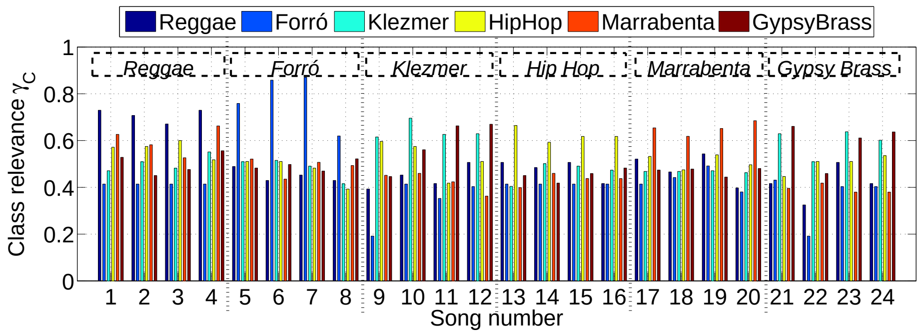
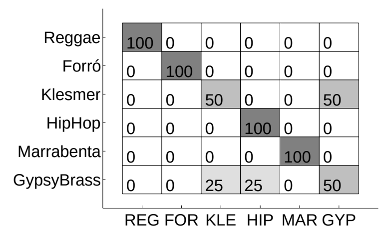



**Fig. 3.** Results of Experiment 2, 3a, 3b, and 3c.




(a) Calculated class relevance values for all songs (see Appendix 8, bar colors encode the related genre). All songs are grouped according to their truth genre label.




(b) Confusion matrix

**Fig. 4.** Detailed results of the rule-based classification for scenario 2 of experiment 1.

**Experiment 4 - What about the real case - how good is the machine nowadays?** We perform a reference experiment to evaluate the classification results we obtained from the presented rule-based approach. Therefore, we apply a state-of-the-art classification system that is based on different mid-level audio features. We compare the results using different feature selection, feature space transformation, and classification as previously explained in Sec. 5.4. As it have been mentioned above in Sec. 5.5, the classification has been performed in a leave-one-out scenario. Given a database of 200 songs, training and classification have been repeated 200 times. On each run the classifiers have been trained with 199 songs and tested with a remaining one. The results of the automatic classification are presented in Table 1. The best accuracy of 79.0% has been achieved while selecting 32 feature dimensions with IRMFSP, applying LDA for feature space transformation and GMM with 5 mixtures for classification.

## 6 Conclusions

The mean classification accuracy value of 83.33% for the rule-based approach is a good result for this new approach. For all genres but Forró, best class relevance value of between approx. 0.5 and 0.75 imply that the confidence of the

| Classifier / FS&FST | SVM | GMM2 | GMM3 | GMM5 | GMM10 | NB | kNN1 | kNN5 | kNN10 |
|---|---|---|---|---|---|---|---|---|---|
| LDA | 74.5 | 77.0 | 77.5 | 77.0 | 75.5 | 77.5 | 76.0 | 76.5 | 76.5 |
| IRMFSP(16) | 67.5 | 61.0 | 62.5 | 62.0 | 60.5 | *57.5* | 62.0 | 62.5 | 63.5 |
| IRMFSP(32) | 68.5 | 63.0 | 66.5 | 65.0 | 62.0 | 63.5 | 65.5 | 66.0 | 67.0 |
| IRMFSP(64) | 62.0 | 68.5 | 65.0 | 61.5 | 58.5 | 67.0 | 61.0 | 64.5 | 64.5 |
| IRMFSP(16)+LDA | 67.0 | 69.0 | 68.5 | 67.5 | 64.5 | 68.5 | 66.5 | 66.0 | 64.5 |
| IRMFSP(32)+LDA | 78.0 | 76.0 | 77.5 | **79.0** | 75.0 | 75.5 | 74.0 | 75.0 | 75.0 |
| IRMFSP(64)+LDA | 73.5 | 73.5 | 75.5 | 76.0 | 73.5 | 74.5 | 75.0 | 75.0 | 75.5 |

**Table 1.** Results of automatic genre classification: mean accuracy values, in %

decision of the classifier was not very high in comparison to the non-classified genres. As it can be seen in Fig. 4(a), for most of the songs the correct class relevance stands out by more than 0.1, often by 0.2 or 0.3, while all incorrect genres' relevances are approximately the same. It is worth to note, that in most cases of misclassification as well in all cases where a correct classification was obtained by less than 0.1 difference, the confusions are explainable by an actual musical similarity or even hybridity. Song no. 9 of Klezmer influenced Hip Hop artists *Solomon & Socalled* for example achieve a high relevance on Klezmer as well as Hip Hop. The music itself and the according minimal difference of the two classes' relevances around 0.6 call for a multi-label solution. The same applies to *Frank London's Klezmer Brass Allstars* (Song nos. 11 & 12) who in fact do mix the styles of Klezmer and Hip Hop. The only clear malfunction of our sets of rules is the high class relevance of Hip Hop for Gypsy Brass pieces, culminating in a misclassification of song no. 22, clearly not a Hip Hop track. For this case, a hybrid solution combining a rule-based as well as a common feature-based approach would be preferable. Detailed analysis of the performance of the set of rules shows that almost every rule and every observable chosen for this experiment gives insights to yet another aspect of the complex field of genre classification.

## 7 Outlook

Some of the observables such as song-origin can not be extracted using content-based methods but with a meta-data analysis of the song. Most observables are based on high-level features whose extraction generally consists of multiple processing steps and thus is prone to potential error propagation. For each automatic annotation of an observable, a confidence measure that can capture the reliability of the computation should be included in the framework to take this uncertainty into account. Furthermore, different approaches for the automatic optimization of the weighting factors $g$ based on a given set of training songs for each class will be investigated. As both approaches - rule based as well as statistically modeled - reach a classification accuracy of around 80%, both are promising. The authors suggest a hybrid framework combining these two approaches. This framework would be able to deal with both unknown classes (genres for which no set of rules exist yet) as well as with classes without a sufficient amount of ground truth data for an appropriate modeling of the statistical classifier.

## 8 Acknowledgments

This work has been partly supported by the German research project *Global-Music2One* funded by the Federal Ministry of Education and Research (BMBF-FKZ: 01/S08039B) (http://www.globalmusic2one.net). The authors would like to thank Peter Woitek for his work on the *Annotation Tool* and Julien Clavel and Mitko Eckert for their contribution in the song annotation process.

## References

1. J. Abeßer, H. Lukashevich, C. Dittmar, and G. Schuller. Genre classification using bass-related high-level features and playing styles. In *Proc. of the Int. Symp. of Music Information Retrieval (ISMIR), Kobe, Japan*, 2009.
2. A. Anglade, R. Ramirez, and S. Dixon. Genre classification using harmony rules induced from automatic chord transcriptions. In *Proceedings of the 10th Conference on Music Information Retrieval (ISMIR)*, pages 669–674, 2009.
3. L. Atlas and S. S. Shamma. Joint acoustic and modulation frequency. *EURASIP Journal on Applied Signal Processing*, 2003:668–675, 2003.
4. J.-J. Aucouturier and F. Pachet. Representing musical genre: A state of the art. *Journal of New Music Research*, 32(1):83–93, 2003.
5. J. Barthelemy and A. Bonardi. Similarity in computational music: a musicologist's approach. In *Proc. of the First International Conference on Web Delivering of Music*, pages 107–113, 2001.
6. J. P. Bello and J. Pickens. A robust mid-level representation for harmonic content in music signals. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, London, UK, 2005.
7. H. Blockeel. Top-down induction of first order logical decision trees. *AI Commun*, 12(1-2):119–120, 1999.
8. G. Buzzanca. A rule-based expert system for music style recognition. In *Proceedings of the 1st International Conference Understanding and Creating Music (UCM), Caserta, Italy*, 2001.
9. P. Delacourt and C. Wellekens. DISTBIC: A speaker-based segmentation for audio data indexing. *Speech Communication*, 32(1-2):111–126, 2000.
10. A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.
11. C. Dittmar, C. Bastuck, and M. Gruhne. Novel mid-level audio features for music similarity. In *Proceedings of the International Conference on Music Communication Science (ICOMCS)*, Sydney, Australia, 2007.
12. S. Doraisamy, S. Golzari, N. M. Norowi, N. Sulaiman, and N. I. Udzir. A study on feature selection and classification techniques for automatic genre classification of traditional malay music. In *Proc. of the International Symposium of Music Information Retrieval (ISMIR), Philadelphia, US*, 2008.

13. R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2nd edition, November 2000.
14. S. Essid. *Classification automatique des signaux audio-fréquences : reconnaissance des instruments de musique*. PhD thesis, Université Pierre et Marie Curie, Paris, France, December 2005.
15. R. Fiebrink. An exploration of feature selection as a tool for optimizing musical genre classification. Master's thesis, McGill University, Montreal, Canada, 2006.
16. K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 2nd edition, September 1990.
17. G. Gatzsche, M. Mehnert, D. Gatzsche, and K. Brandenburg. A symmetry based approach for musical tonality analysis. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR)*, Vienna, Austria, 2007.
18. Gebesmair. *Songs of the Minotaur - Hybridity and Popular Music in the Era of Globalization*, chapter Hybrids in the global economy of music: How the major labels define the Latin mass market, pages 1–20. LIT Verlag Münster, Hamburg, London, 2002.
19. R. O. Gjerdingen and D. Perrott. Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*, 37:93–100, 2008.
20. M. Gruhne and C. Dittmar. Comparison of harmonic mid-level representations for genre recognition. In S. Baumann, J. J. Burred, A. Nürnberger, and S. Stober, editors, *Proceedings of the 3rd Workshop on Learning the Semantics of Audio Signals (LSAS)*, pages 91–102, Graz, Austria, Dec 2009.
21. M. Gruhne, C. Dittmar, and D. Gaertner. Improving rhythmic similarity computation by beat histogram transformations. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)*, Kobe, Japan, 2009.
22. M. Hendler. Clave del son: die rhythmusformeln in der musik der karibik. *Musikwissenschaft*, 12, 2007.
23. F. J. and D. J.S. Interdisciplinary research issues in music information retrieval: Ismir 20002002. *Journal of New Music Research*, 32:121–131, 2003.
24. D.-N. Jiang, L. Lu, H.-J. Zhang, J. Tao, and L.-H. Cai. Music type classification by spectral contrast feature. In *Proc. IEEE Int. Conf. Multimedia and Expo (ICME)*, page 113, Lausanne, Switzerland, 2002.
25. O. E. Laske. *Musical Grammars and Computer Analysis*, chapter KEITH: A Rule-System for Making Music-Analytical Discoveries, pages 165–199. Florence: Olschki, 1984.
26. K. Lee. Automatic chord recognition from audio using enhanced pitch class profile. In *Proceedings of the International Computer Music Conference (ICMC)*, New Orleans, USA, 2006.
27. H. Lukashevich, J. Abeßer, C. Dittmar, and H. Großmann. From multi-labeling to multi-domain-labeling: A novel two-dimensional approach to music genre classification. In *Proc. of the International Symposium of Music Information Retrieval (ISMIR), Kobe, Japan*, 2009.
28. C. McKay and I. Fujinaga. Automatic genre classification using large high-level musical feature sets. In *Proc. of the International Symposium of Music Information Retrieval (ISMIR)*, 2004.
29. C. McKay and I. Fujinaga. Automatic music classification and the importance of instrument identification. In *Conference on Interdisciplinary Musicology*, 2005.
30. C. McKay and I. Fujinaga. jSymbolic: A feature extractor for MIDI files. In *International Computer Music Conference (ICMC)*, pages 302–305, 2006.
31. F. Pachet and D. Cazaly. A taxonomy of musical genres. In *Proc. of the Content-Based Multimedia Information Access Conference (RIAO)*, Paris, France, 2000.

32. G. Peeters and X. Rodet. Hierarchical gaussian tree with inertia ratio maximization for the classification of large musical instruments databases. In *Proceedings of the 6th Int. Conf. on Digital Audio Effects (DAFx)*, London, UK, 2003.
33. N. Scaringella, G. Zoia, and D. Mlynek. Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine*, 23:133–141, 2006.
34. J. Shepherd and P. Wicke. *Music and Cultural Theory.* Polity Press, Cambridge, UK, 1997.
35. M. Stein, B. M. Schubert, M. Gruhne, G. Gatzsche, and M. Mehnert. Evaluation and comparison of audio chroma feature extraction methods. In *Proceedings of the 126th AES Convention*, Munich, Germany, 2009.
36. G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.
37. G. Tzanetakis, A. Kapur, W. A. Schloss, and M. Wright. Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, 1(2):1–24, 2007.
38. C. Uhle, C. Dittmar, and T. Sporer. Extraction of drum tracks from polyphonic music using independent subspace analysis. In *Proceedings of the 4th International Symposium on Independent Component Analysis*, Nara, Japan, 2003.
39. V. N. Vapnik. *Statistical learning theory.* Wiley New York, 1998.
40. H. Zhang. The optimality of naive bayes. In V. Barr and Z. Markov, editors, *Proceedings of the FLAIRS Conf.* AAAI Press, 2004.

## Appendix: Annotation Tool

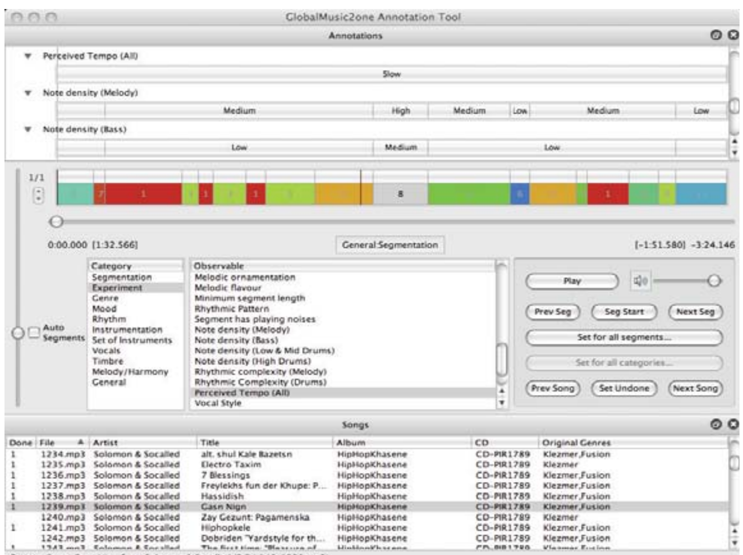



**Fig. 5.** Annotation tool.

## Appendix: Set of rules

All 6 set of rules that were used in this paper are depicted in Tab. 3. Several symbols are used to represent different *deduced observables*. Because the classification is performed song-based and all rules relate to complete songs, the segment-based observable annotations have to be further processed. "TIME_RATIO" implies the overall time ratio of a song, where one ("TIME_RATIO_SINGLE(Obs., Val.)") or multiple ("TIME_RATIO_MULT(Obs.1, Val.1, Obs.2, Val.2,...)") observable(s) have a specific value, with "AND", "OR", and "NOT" representing logical operations that are uses to join multiple observable-value pairs.

"PRES_OBS_VAL(Obs., Val.)" and "PRES_INSTR(Instr.)" are fulfilled if at least one annotation of the respective observable-value pair or instrument is given for a song.

| | **Type** | *O* | *R* | *V* | *g* |
|---|---|---|---|---|---|
| | | **Reggae** | | | |
| 1 | MO | PRES_OBS_VAL(Segment Type, Refrain) | $\equiv$ | 1 | - |
| 2 | MO | TIME_RATIO_SINGLE(Bar Measure, 4/4) | $\geq$ | 0.5 | - |
| 3 | MO | TIME_RATIO_SINGLE(HarmonyInstrument Accent, Off-beat) | $\geq$ | 0.6 | - |
| 4 | MO | TIME_RATIO_MULT_AND(HarmonyInstrument Accent, Off-beat, HarmonyInstrument Plucking Style, Skanking) | $\geq$ | 0.9 | - |
| 5 | MO | TIME_RATIO_SINGLE(Polyphony (Melody), No Polyphony) | $\geq$ | 0.9 | - |
| 6 | FO | PRES_OBS_VAL(Rhythmic Influence, Afro-caribean) | $\equiv$ | 1 | 0.8 |
| 7 | FO | PRES_INSTR(Electric Bass Guitar) | $\equiv$ | 1 | 0.4 |
| 8 | FO | PRES_OBS_VAL(Metallic guitar sound, Yes) | $\equiv$ | 1 | 0.4 |
| 9 | FO | TIME_RATIO_SINGLE_OR(Tempo of Off-beat, 60-80, 80-100) | $\geq$ | 0.5 | 0.6 |
| 10 | FO | TIME_RATIO_SINGLE(Presence of rhythm instrument, Yes) | $\geq$ | 0.5 | 0.35 |
| 11 | FO | TIME_RATIO_SINGLE(Presence of brass instrument, Yes) | $\geq$ | 0.5 | 0.25 |
| | | **Forró** | | | |
| 1 | M,O | PRES_INSTR(Accordion (Sanfona)) | $\equiv$ | 1 | - |
| 2 | M,O | PRES_INSTR(Bass drum (Zabumba)) | $\equiv$ | 1 | - |
| 3 | M,O | PRES_INSTR(Triangle) | $\equiv$ | 1 | - |
| 4 | M,O | TIME_RATIO_SINGLE(Danceable Rhythm, Yes) | $\geq$ | 0.5 | - |
| 5 | M,O | TIME_RATIO_SINGLE(Polyphony (Melody), No Polyphony) | $\geq$ | 0.9 | - |
| 6 | F,O | TIME_RATIO_SINGLE_OR(Note density (Low & Mid Drums), Low, Medium) | $\geq$ | 0.5 | 0.45 |
| 7 | F,O | TIME_RATIO_SINGLE(Note density (High Drums), High) | $\geq$ | 0.5 | 0.7 |
| 8 | F,O | PRES_OBS_VAL(Song origin, Brazil) | $\equiv$ | 1 | 0.4 |
| 9 | F,O | TIME_RATIO_SINGLE_NOT(Number of Vocalists, None) | $\geq$ | 0.5 | 0.1 |
| | | **Klezmer** | | | |
| 1 | M,O | TIME_RATIO_MULT_OR(Presence of brass instrument, Yes, Presence of string instrument, Yes, Presence of keyboard instrument, Yes) | $\geq$ | 0.1 | - |
| 2 | F,O | TIME_RATIO_SINGLE_OR(Number of Instruments, 2, >2) | $\geq$ | 0.5 | 0.45 |
| 3 | F,O | PRES_INSTR(Clarinet) | $\equiv$ | 1 | 0.9 |
| 4 | F,O | TIME_RATIO_SINGLE(Presence of brass instrument, Yes) | $\geq$ | 0.5 | 0.4 |
| 5 | F,O | TIME_RATIO_SINGLE(Presence of rhythm instrument, Yes) | $\geq$ | 0.5 | 0.45 |
| 6 | F,O | PRES_OBS_VAL(Melody has a howling intonation, Yes) | $\equiv$ | 1 | 0.7 |
| 7 | F,O | TIME_RATIO_SINGLE(Tonal complexity (Melody), High) | $\geq$ | 0.5 | 0.25 |
| 8 | F,O | TIME_RATIO_SINGLE(Melodic ornamentation, Vibrato, Glissando, Triller) | $\geq$ | 0.5 | 0.45 |
| 9 | F,O | TIME_RATIO_SINGLE(Time signature type, Binary) | $\geq$ | 0.5 | 0.4 |
| 10 | F,O | TIME_RATIO_SINGLE_OR(Note density (Melody), Medium, High) | $\geq$ | 0.5 | 0.4 |
| 11 | F,O | TIME_RATIO_SINGLE_OR(Tonal complexity (Melody), Medium, High) | $\geq$ | 0.5 | 0.35 |
| 12 | F,O | TIME_RATIO_SINGLE(Rhythmic Complexity (Drums), Low) | $\geq$ | 0.5 | 0.35 |
| 13 | F,O | TIME_RATIO_SINGLE(Polyphony (Melody), No Polyphony) | $\geq$ | 0.5 | 0.45 |
| 14 | F,O | PRES_INSTR(Accordion) | $\equiv$ | 1 | 0.35 |
| 15 | F,O | PRES_INSTR(Violin) | $\equiv$ | 1 | 0.3 |
| 16 | F,O | TIME_RATIO_SINGLE_OR(Perceived Tempo (All), Slow, Fast) | $\geq$ | 0.5 | 0.3 |
| 17 | F,O | TIME_RATIO_SINGLE(Perceived Tempo (All), Medium) | $\geq$ | 0.5 | 0.15 |
| 18 | F,O | TIME_RATIO_SINGLE(Tempo - Pres. of Accelerandi and Ritardandi, Yes) | $\geq$ | 0.5 | 0.3 |

| | | | | |
|---|---|---|---|---|
| | | **Hip Hop** | | |
| 1 | M,O | TIME_RATIO_SINGLE(Bar Measure, 4/4) | $\geq$ 0.8 | - |
| 2 | F,O | TIME_RATIO_SINGLE_NOT(Number of Vocalists, None) | $\geq$ 0.5 | 0.45 |
| 3 | F,O | TIME_RATIO_SINGLE(Presence of rhythm instrument, Yes) | $\geq$ 0.5 | 0.45 |
| 4 | F,O | TIME_RATIO_SINGLE_OR(Note density (Low & Mid Drums), Low, Medium) | $\geq$ 0.5 | 0.45 |
| 5 | F,O | TIME_RATIO_SINGLE_OR(Note density (High Drums), Low, Medium) | $\geq$ 0.5 | 0.45 |
| 6 | F,O | PRES_INSTR(Drum Machine) | $\equiv$ 1 | 0.25 |
| 7 | F,O | PRES_INSTR(Samples) | $\equiv$ 1 | 0.3 |
| 8 | F,O | TIME_RATIO_SINGLE_OR(Vocal Style, Rap, Rap&Skit, Singing+Rap) | $\equiv$ 1 | 0.7 |
| 9 | F,O | TIME_RATIO_SINGLE(Tonal complexity (Melody),Low) | $\geq$ 0.5 | 0.35 |
| 10 | F,O | PRES_INSTR(Scratches) | $\equiv$ 1 | 0.35 |
| 11 | F,O | TIME_RATIO_SINGLE(Rhythmic Complexity (Drums), Low) | $\geq$ 0.5 | 0.6 |
| 12 | F,O | TIME_RATIO_SINGLE_OR(Perceived Tempo (All), Slow, Medium) | $\geq$ 0.5 | 0.45 |
| 13 | F,O | TIME_RATIO_SINGLE(Variety of rhythm,Low) | $\geq$ 0.5 | 0.4 |
| 14 | F,O | TIME_RATIO_SINGLE(Rhythmic Complexity (Drums), Medium) | $\geq$ 0.5 | 0.2 |
| | | **Marrabenta** | | |
| 1 | M,O | TIME_RATIO_SINGLE_NOT(Number of Instruments, 1) | $\geq$ 0.5 | - |
| 2 | M,O | TIME_RATIO_SINGLE_NOT(Number of Vocalists, None) | $\geq$ 0.5 | - |
| 3 | F,O | PRES_INSTR_OR(Accoustic Guitar, Electric Guitar) | $\equiv$ 1 | 0.6 |
| 4 | F,O | TIME_RATIO_SINGLE(Guitar plays melodic patterns, Yes) | $\geq$ 0.5 | 0.6 |
| 5 | F,O | TIME_RATIO_SINGLE(Male vocals and multiple singers alternate, Yes) | $\geq$ 0.5 | 0.6 |
| 6 | F,O | PRES_INSTR(Drums) | $\equiv$ 1 | 0.4 |
| 7 | F,O | PRES_INSTR(Electric Bass Guitar) | $\equiv$ 1 | 0.4 |
| 8 | F,O | TIME_RATIO_SINGLE(Presence of brass instrument, Yes) | $\geq$ 0.5 | 0.35 |
| 9 | F,O | TIME_RATIO_SINGLE(Syncopation (Drums), Yes) | $\geq$ 0.5 | 0.3 |
| 10 | F,O | TIME_RATIO_SINGLE(Segment has African Percussion Instruments, Yes) | $\geq$ 0.5 | 0.4 |
| 11 | F,O | TIME_RATIO_SINGLE_NOT(Rhythmic Complexity (Drums), Low) | $\geq$ 0.5 | 0.4 |
| 12 | F,O | TIME_RATIO_SINGLE_OR(Variety of rhythm, Medium, High) | $\geq$ 0.5 | 0.3 |
| 13 | F,O | PRES_OBS_VAL(Metallic guitar sound, Yes) | $\equiv$ 1 | 0.6 |
| 14 | F,O | TIME_RATIO_SINGLE_OR(Note density (Low & Mid Drums),High,Medium) | $\geq$ 0.5 | 0.35 |
| 15 | F,O | TIME_RATIO_SINGLE(Note density (High Drums), High) | $\geq$ 0.5 | 0.25 |
| 16 | F,O | TIME_RATIO_SINGLE_OR(Note density (Bass), Medium, High) | $\geq$ 0.5 | 0.35 |
| 17 | F,O | TIME_RATIO_MULT_AND(segment_position, Beginning, Perceived Tempo (All),Slow) | $\geq$ 0.1 | 0.15 |
| 18 | F,O | TIME_RATIO_SINGLE(If male vocalists are singing, they have no low voices,Yes) | $\geq$ 0.5 | 0.4 |
| | | **GypsyBrass** | | |
| 1 | M,O | TIME_RATIO_SINGLE(Presence of brass instrument, Yes) | $\geq$ 0.5 | - |
| 2 | M,O | TIME_RATIO_SINGLE_NOT(Number of Vocalists, None) | $\leq$ 0.4 | - |
| 3 | M,O | TIME_RATIO_SINGLE_OR(Number of Instruments, 2, ¿2) | $\geq$ 0.5 | - |
| 4 | F,O | TIME_RATIO_SINGLE(Danceable Rhythm, Yes) | $\geq$ 0.5 | 0.6 |
| 5 | F,O | TIME_RATIO_SINGLE(Time signature type, Binary) | $\geq$ 0.5 | 0.45 |
| 6 | F,O | TIME_RATIO_SINGLE(Polyphony (Melody), No Polyphony) | $\geq$ 0.5 | 0.45 |
| 7 | F,O | TIME_RATIO_SINGLE(Unprecise tuning of brass instruments, Yes) | $\geq$ 0.5 | 0.6 |
| 8 | F,O | PRES_INSTR(Trumpet) | $\equiv$ 1 | 0.6 |
| 9 | F,O | TIME_RATIO_SINGLE(Presence of rhythm instrument, Yes) | $\geq$ 0.5 | 0.45 |
| 10 | F,O | TIME_RATIO_SINGLE(Syncopation (Drums), Yes) | $\geq$ 0.5 | 0.45 |
| 11 | F,O | TIME_RATIO_SINGLE_NOT(Perceived Tempo (All), Medium) | $\geq$ 0.5 | 0.4 |
| 12 | F,O | TIME_RATIO_SINGLE_OR(Perceived Tempo (All), Fast, Very Fast) | $\geq$ 0.5 | 0.4 |
| 13 | F,O | TIME_RATIO_SINGLE(Melodic flavour,Oriental ornamentation) | $\geq$ 0.5 | 0.4 |
| 14 | F,O | TIME_RATIO_SINGLE(Melodic scale, Based on minor-related scale) | $\geq$ 0.3 | 0.35 |
| 15 | F,O | NUM_SEG_SINGLE(Segment Type,Refrain) | $\geq$ 1 | 0.35 |
| 16 | F,O | TIME_RATIO_SINGLE(Melodic ornamentation, Vibrato, Glissando, Triller) | $\geq$ 0.5 | 0.35 |
| 17 | F,O | TIME_RATIO_SINGLE(Rhythmic Complexity (Drums),Low) | $\geq$ 0.5 | 0.35 |
| 18 | F,O | PRES_OBS_VAL(num_segments,11,$\leq$) | $\equiv$ 1 | 0.2 |
| 19 | F,O | TIME_RATIO_SINGLE(len_segments, 20000, $\geq$) | $\geq$ 0.5 | 0.2 |
| 20 | F,O | TIME_RATIO_SINGLE(Tempo - Pres. of Accelerandi and Ritardandi, Yes) | $\geq$ 0.5 | 0.15 |
| 21 | F,O | TIME_RATIO_SINGLE(Segment has playing noises, Yes) | $\geq$ 0.5 | 0.1 |
| 22 | F,O | TIME_RATIO_SINGLE(Number of Instruments, >2) | $\geq$ 0.5 | 0.45 |
| 23 | F,C | **Condition**: PRES_INSTR(Tuba) $\equiv$ 1 | | |
| | | **Property**: TIME_RATIO_SINGLE(Brass instr. plays rhythmically, Yes) | $\geq$ 0.5 | 0.4 |

## Appendix: List of songs

| Nr. | Genre | Artist | Title |
|---|---|---|---|
| 1 | **Reggae** | Ethio Stars | Tiz Balegn Gize |
| 2 | | Watcha Clan | Travelin' Shoes |
| 3 | | Fermin Muguruza | Mendebaldakareta |
| 4 | | Ethio Stars | Yekereme Fikir |
| 5 | **Forró** | Anastácia Azevedo | Xaxado (For Lampião) |
| 6 | | Oswaldinho Do Acordeon | Guadá E Live No Forró |
| 7 | | Cascabulho | Xodó de Sanfoneiro |
| 8 | | Cabruera | Espinhos |
| 9 | **Klezmer** | Solomon & Socalled | Gasn Nign |
| 10 | | The Klezmatics | Heyser Tartar-Tants |
| 11 | | Frank London's Klezmer Brass Allstars | Trink Nokh A Glezele Vayn, Moishele? |
| 12 | | Frank London's Klezmer Brass Allstars | Lieberman Husidl |
| 13 | **Hip Hop** | Sister Fa | Hip Hop Yaw La Fal |
| 14 | | African Achlou Bi | Andando |
| 15 | | Sister Fa | Amy Jotna |
| 16 | | Keur Gui | Liye Raam |
| 17 | **Marrabenta** | Orchestra Marrabenta Star De Moçambique & Wazimbo | Parabens |
| 18 | | Orchestra Marrabenta Star De Moçambique & Wazimbo | Matilde |
| 19 | | Ghorwane | Xizambiza |
| 20 | | Orchestra Marrabenta Star De Moçambique & Wazimbo | DJomela |
| 21 | **Gypsy Brass** | Boban Marković Orkestar | Mundo Čoček |
| 22 | | Fanfare Ciocărlia | Lume, Lume |
| 23 | | Fanrare Tirana | Zot, O Zot, Të Qofshin Fal |
| 24 | | Fanfare Ciocărlia | Hora Cu Strigaturii |

**Table 3.** List of songs