RHYTHMIC SIMILARITY IN TRADITIONAL TURKISH MUSIC

Andre Holzapfel
Institute of Computer Science, FORTH and University of Crete
hannover@csd.uoc.gr

Yannis Stylianou
Institute of Computer Science, FORTH and University of Crete
yannis@csd.uoc.gr

ABSTRACT

In this paper, the problem of automatically assigning a piece of traditional Turkish music into a class of rhythm referred to as *usul* is addressed. For this, an approach for rhythmic similarity measurement based on scale transforms has been evaluated on a set of MIDI data. Because this task is related to time signature estimation, the accuracy of the proposed method is evaluated and compared with a state of the art time signature estimation approach. The results indicate that the proposed method can be successfully applied to audio signals of Turkish music and that it captures relevant properties of the individual *usul*.

1. INTRODUCTION

Traditional music of Turkey has a big community of listeners, and the music is strongly related to the music of neighboring regions. For example, in Greece and Arabian countries music melodies of traditional music are often based on similar modal systems as in Turkey. Concerning rhythm, there is a correspondence in classes of rhythm found in Arabic music (*iqa'*) and in Turkey (*usul*), and dances encountered in Turkey have influenced rhythms played in Greek *Rembetiko* music. Thus, automatic retrieval of this information not only enables a better understanding of an important cultural heritage but may also be of major commercial interest. Methods for this type of retrieval can be assigned to the branch of computational ethnomusicology, as introduced in [20]. Only recently, first research results on the classification of Turkish music into melodic classes were presented [7]. The retrieval of rhythmic information from traditional Turkish music has not been addressed yet. In this paper, classification of samples of Turkish music into rhythmic classes is proposed. These classes are referred to as *usul* [16]. A data set containing samples of songs composed in six different *usul* has been compiled to conduct experiments. As it will be shown in the later Sections, in the context of this data set the classification into a specific rhythmic class is related to the recognition of the time signature in Western music.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.
© 2009 International Society for Music Information Retrieval.

In [18], an approach was presented to estimate the time signature of a piece of music based on symbolic descriptions (MIDI). This approach uses autocorrelation coefficients (ACF) derived from the annotated onsets. In [21], a time signature estimation system for audio signals was proposed and evaluated on a set of percussive music. The system estimates the tatum [2] of the signal using inter-onset intervals (IOI) and in parallel, ACF are computed from the amplitude envelope of the signal. Beat and bar length are chosen from the peaks of the ACF, taking into account the estimated tatum. In [8], the determination of musical meter was reduced to a classification into either binary or ternary meter. Beat indexes are extracted in a semi-automatic way and then ACF on a chosen set of features are used to decide on the meter type. Using audio signals, the general problem of rhythmic similarity was addressed previously in [10] [1] in the context of traditional music, in both cases by applying Dynamic Time Warping techniques. In [4], rhythmic patterns were computed from samples of Western ballroom dances.

In [14] a system was proposed for the automatic estimation of the musical meter, i.e., the estimation of the position of tatum, beat and bars in the signal. The estimation of bar positions in $\frac{3}{4}$ time signatures is mentioned to be error-prone. Compound time signatures such as $\frac{9}{8}$ are not mentioned and to the best of our knowledge no reliable method has been presented to estimate the meter in such signals.

On the other hand, compound or complex time signatures are commonly encountered in traditional music of Turkey. The time signatures can take various forms, as it will be detailed in Section 2. Furthermore, the goal of the approach presented in this paper is not only the correct estimation of a time signature, but a description of the rhythmic properties of a class, because *usul* cannot be only distinguished by time signature in all cases. In [11], audio samples of traditional dances were compared: ACF were computed from onset strength signals (OSS) and these ACF were transformed into the scale domain by using the scale transform [3]. This results in descriptors that do not vary due to tempo changes. Thus, the scale transform magnitudes (STM) can be used to compare the rhythmic content of audio using simple point to point distance measures without the need of meter estimation. The approach in [11] was shown to be superior to the DTW based approach presented in [10]. In this paper, it will be combined with the approach presented in [18] and applied to a set of MIDI data. MIDI data was chosen as a first step to approach the problem of automat-

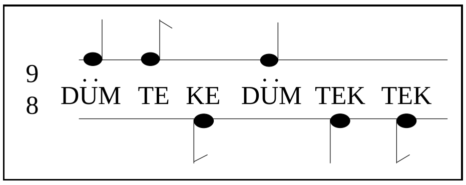

Figure 1. Symbolic description of the *usul Aksak*

ic rhythm description in Turkish music. This approach can easily be applied to audio signals by replacing the type of OSS, has no need of meter estimation and is robust to large tempo deviations.

Section 2 introduces the basic concepts of the analyzed type of music and describes the data set. Section 3 introduces the descriptors based on scale transform and proposes a method of comparison. Section 4 gives experimental results and Section 6 concludes the paper.

2. DATA SET

Compositions in Turkish traditional music follow certain schemes regarding their melodic and rhythmic content. Melodies are characterized by a modal system referred to as *makam*, and it defines a melodic texture consisting of specific tonal segments, progressions, directionality, temporal stops, tonal centers and cadences [13]. The rhythmic schemes encountered in traditional Turkish music are referred to as *usul*. An *usul* is a rhythmic pattern of certain length that defines a sequence of strong and weak intonations. An example is shown in Figure 1: the *usul Aksak* has a length of nine beats. The notes on the upper line labelled *düm* have the strongest intonation while the notes on the low line denote weak intonations. The note durations in the sequence shown in Figure 1 can be described as the string `xoxxxoxox`, where `x` symbolizes the start of a note and `o` metric unit without note [19]. Note that this representation is a further simplification of the one shown in Figure 1, because no differentiation of the intonation strength is contained. However these representations can be used for estimating the similarity between rhythms of same lengths by computing a chronotonic distance, as detailed in [19].

Unlike in [19], the length of the *usul* varies. According to H. Sadeddin Arel (1880-1955), the *usul* can be divided into minor and major *usul*. Minor *usul* have a length of up to 15 time units, while the major *usul* have up to 124 time units. As denoted in [16], minor usul are related to small musical forms, while larger musical forms employ the major usul in most cases. Musical forms that are usually composed in major usul are, e.g., *Presrev* and *Beŝte*. Two examples of small musical forms are *Sarkı* and *Türkü*. The latter are folk songs of unknown composers, while the former are short songs based usually on four lines of text with known composer. Both forms have in common that a song follows a certain minor *usul* and a certain *makam*, and both forms are vocal music. The most popular songs in Turkish music are composed in these forms. Because of that, along with a system for the recognition of the *makam* as presented in [7], an approach for the recognition of the *usul* represents an essential element in automatic retrieval of information from this music. Apart from that, the relation between melody and *usul* has not been investigated and an automatic approach like the one presented here can give valuable insight into the relation between melody and *usul*.

The data set used in this paper consists of Turkish songs in the forms of *Sarkı* and *Türkü*. They are following six different types of rhythmic schemes having lengths from 3 up to 10: *Aksak* ($\frac{9}{8}$), *Curcuna* ($\frac{10}{8}$), *Düyek* ($\frac{8}{8}$), *Semai* ($\frac{3}{4}$), *Sofyan* ($\frac{4}{4}$), and *Türk Aksaği* ($\frac{5}{8}$). The software *mus2okur* [13] has been used to obtain a data set consisting of 288 songs distributed along the six classes as shown in the second line of Table 1. Each sample consists of a MIDI description of the song melody, in most cases also a MIDI voice with a percussive accompaniment is contained. This percussive accompaniment has been left out, in order to be able to focus on the rhythmic properties of the melody. Due to the character of this music, there exists no chord accompaniment.

As all *usul* in the data set have different length, the recognition of the *usul* can be reduced to a recognition of its length. This is closely related to the task of time signature recognition and motivates the experimental setup described in the following Sections. The lower two lines in Table 1 depict the mean values of the tempi in *bpm* (beats per minute) and the standard deviation of the tempi, respectively. It is apparent that there are large overlaps between the tempo distributions of the *usul*. Thus, a system for *usul* length estimation for a given audio signal has to be robust to the tempo deviations and overlaps.

Table 1. Data set: number of songs, mean and standard deviation of tempi in *bpm*

CLASS	AKS	CUR	DUY	SEM	SOF	TUR
N_{Songs}	64	57	47	22	60	38
MEAN	98.5	98.3	70.7	131.9	81.3	73.1
STD	27.9	13.5	12.6	26.3	16.7	22.3

3. TIME SIGNATURE ESTIMATION

3.1 Rhythm Description

3.1.1 *Tempo-invariant ACF*

In order to describe and compare the content of the samples, an autocorrelation based method as presented in [18] has been combined with a method used for estimating rhythmic similarity presented in [11]. The onset times are read from the MIDI files and each onset is assigned a weight. In [18], different methods to set the weights were evaluated, and in this paper the three most successfull weighting schemes have been applied: the weight of an onset can either be related to the note duration as proposed in [15], to characteristics of the melody [17], or all onsets are assigned the same weight. The best weighting scheme will be

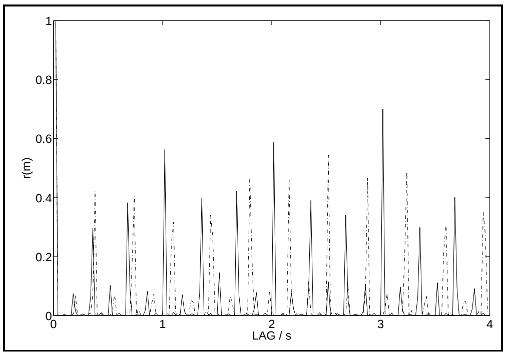

Figure 2. Autocorrelations r_u derived from two samples of *usul aksak*

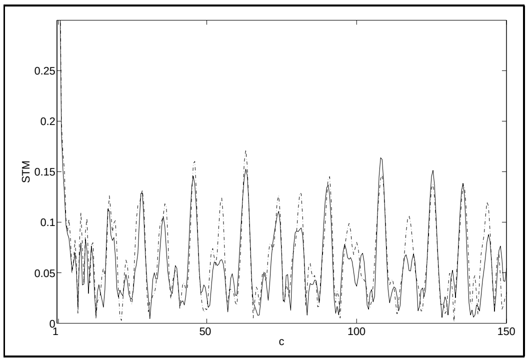

Figure 3. Two STM derived from the two *aksak* examples shown in Figure 2

determined in Section 4. In the method presented in [18], an onset strength signal (OSS) is generated at a sampling frequency related to the eighth note of the piece. This OSS has an impulse of height according to the assigned weight at the positions related to the onset time. From an OSS $o(n)$ an ACF $r(m)$ can be derived

$$r(m) = \frac{\sum_n o(n)o(n-m)}{\sum_n o(n)^2} \tag{1}$$

Note that the autocorrelations are not affected by tempo differences, when the OSS are computed at a sampling frequency that changes with the tempo (eighth note). Because of this, changing the tempo will result in constant ACF, which will be denoted as r_c.

3.1.2 Tempo-variant ACF

As mentioned in [18], beat tracking is a necessary step when applying the above described approach to audio. It is necessary to correctly estimate all metric levels in order to determine the eighth note pulse of the piece. When dealing with compound rhythms of different type as they are contained in the data set and commonly encountered in the music of Turkey and the whole eastern Mediterranean, no method has been presented yet to perform this task. For that reason, the MIDI data contained in the data set as described in Section 2 is used to compute OSS using a constant sampling frequency of $f_s = 50Hz$. From the OSS autocorrelations are derived. For two pieces having the same time signature but different tempi, their autocorrelations will differ by an unknown scaling factor, as can be seen in Figure 2. This is particularly critical for the type of music examined in this paper due to the large tempo deviations as detailed in Section 2. In order to overcome this scaling problem, typically the beat tracking would be necessary in order to estimate the tempo difference between the pieces. However, in this paper the usage of the method introduced in [11] is proposed to avoid the intractable problem of beat tracking in the presence of complex and compound time signatures. Due to the unknown scaling factor depicted in Figure 2, a simple point-to-point distance measure cannot be applied when comparing these autocorrelations, which due to the unknown scaling will be denoted as r_u. In order to solve this problem, a scale transform has been applied to the autocorrelation sequence $r_u(t)$:

$$R(c) = \frac{1}{2\pi} \int_0^\infty r_u(t) e^{(-jc-1/2)\ln t} dt \tag{2}$$

The scale transform has the property that for a signal $r_u(t)$ and its time scaled version $\sqrt{a} r_u(at)$, with $a > 0$ being the scaling factor, the two computed scale transform magnitudes will be the same. This can be seen in Figure 3, where the two scaled autocorrelations from Figure 2 have been transformed to scale space. Due to the scale invariance property they are aligned and can be directly compared.

Thus, in this paper OSS will be computed from the MIDI files using a constant sampling frequency of $f_s = 50Hz$. Then, scale transform magnitudes (STM) are computed from the autocorrelations r_u using the discrete scale transform algorithm proposed in [22]. This results in a STM vector that describes the rhythmic content of the signal, the scale resolution was found to be of minor importance and has been set to $\Delta c = 0.5$. The accuracy in the task of time signature recognition when using either scaling free autocorrelations r_c or the STM derived from r_u will be compared. The results will indicate if by using a scale transform, the unsolved problem of meter estimation in complex time signatures can be avoided and the *usul* length could be determined by using this method.

3.2 Rhythm Dissimilarity

In order to determine the time signature of a piece the following approach will be applied: All pairwise dissimilarities between songs are computed using either the scale-free ACF r_c or the STM vectors, by using a cosine distance as proposed in [6] [9]. This results in dissimilarity matrices, having values close to zero whenever two pieces are found to be similar regarding their rhythmic content. In order to determine the accuracy of the proposed rhythmic similarity measure, the accuracies of a modified k-Nearest Neighbor (kNN) classification will be determined. For this, each single song will be used as a query for that a classification into one of the available classes is desired. This classification is performed by applying the modified kNN to the dissimilarity matrix. As shown in [10], a locally weighted kNN was found to improve accuracies on similar data, and

DURATION	MELODY	FLAT
80.2%	68.1%	72.9%

Table 2. Time signature recognition accuracies when using scale free r_c representation

therefore it has been used in the experiments. It assigns a weight $w_i = 1 - (d_i/d_{k+1})$ to the i-th training sample, where d_{k+1} is the distance of the $k+1$-nearest neighbor to the test sample. Thus, training samples more far away from the test sample contribute less to its classification.

An *usul* can be expressed in a simplified way as a string, as for example the string xoxxxoxox for *Aksak*. In Section 4, for some *usul* their string representations will be used to estimate their similarity using a method proposed in [19]: From the string representations chronotonic chains can be computed, by breaking down the rhythm into its smallest time unit on the *x-axis* and assigning to each element a height on the *y-axis* according to the beat-to-beat interval. This results in the chronotonic chain $[211221]$ in case of *Aksak*. As proposed in [19], in order to compare two such chronotonic chains, then a discrete form of the Kolmogorov Variational Distance (DKVD) can be applied. Given two chronotonic chains g and f of same length L, this distance can be computed as

$$K = \sum_{i=1}^{L} |f[i] - g[i]| \tag{3}$$

and is equal to the $1 - norm$ distance between the chains. Thus, by depicting an *usul* pair as two strings of same length, their rhythmic similarity can be estimated. In this paper, this method will be applied to pairs of *usul* for that samples frequently were confused in the time signature recognition.

4. EXPERIMENTS

4.1 Scale-free ACF

Three different weighting schemes have been evaluated in the experiments: the duration accent as proposed in [15], the melodic accent [17], and the flat accent (i.e., using the same accent weight for all onsets). Using the r_c autocorrelations computed using these three accents in the classification approach as described in Section 3.2, resulted in the best accuracies for the duration accent, as documented in Table 2. This contradicts with the findings in [18], where the melodic and flat accents were found to be preferable. Furthermore, using a selected range of autocorrelation coefficients could not further improve results on this data set, while in [18] using the coefficients of longer lags and leaving out the coefficients of short lags was found superior. This must be assigned to the differences between the data sets.

In Table 3 the confusion matrix for the best classification in Table 2 is shown. The biggest confusion happens between the $\frac{8}{8}$ time signature *usul* and the $\frac{4}{4}$ *usul* (*Düyek* and *Sofyan*, respectively). The pieces in the $\frac{8}{8}$-*usul* could be equivalently annotated in a $\frac{8}{4}$ time signature by changing their degree, referred to as *mertebe*, to four. The second biggest confusion happens between *Curcuna* and *Türk Aksaği*. The time signatures are related by a factor of two as well ($\frac{10}{8}$ and $\frac{5}{8}$). These types of errors have been denoted as typical as well in [18]. Still, the confusion between between *Düyek* and *Sofyan* is larger. This can be attributed to the different degree of similarity of the *usul*, which can be estimated using the approach proposed in [19]: In Table 4, the symbolic descriptions for the two confused *usul*-pairs are depicted as vectors of same length. From these descriptions the chronotonic chains have been derived that are depicted in Table 4. Note that *Sofyan* would be typically denoted as $[211]$ as its smallest beat-to-beat interval is a fourth note. In order to get chains of equal length, the eighth note has been chosen as smallest unit. Computing the Kolmogorov Variational Distances between the chronotonic chains, and normalizing by the length of the vectors it can be seen that the *usul Düyek* and *Sofyan* are more similar than the other pair. This is reflected in the higher confusion in Table 3. Thus, it can be concluded that the applied autocorrelation method is not only suitable for determining time signatures, but can as well capture rhythmic similarities contained in the piece.

		Predicted					
		9/8	10/8	8/8	3/4	4/4	5/8
Notated	9/8	62	0	1	0	1	0
	10/8	0	50	0	0	1	6
	8/8	1	4	24	0	18	0
	3/4	0	0	0	20	2	0
	4/4	2	0	12	0	46	0
	5/8	0	9	0	0	0	29

Table 3. Confusion matrix for r_c using duration accent

Symbolic Description			
Düyek:	xxoxxoxo	*Curcuna*:	xoxxoxoxox
Sofyan:	xoooxoxo	*Türk Aksaği*:	xoooxoooxo
Chronotonic Chains			
Düyek:	12212222	*Curcuna*:	2212222221
Sofyan:	44442222	*Türk Aksaği*:	4444444422
Normalized DKVD betw. Chronotonic Chains			
10/8=1.25		18/10=1.8	

Table 4. Computing chronotonic distances between confused *usul*

4.2 Scale Transform Magnitudes

The results presented in Section 4.1 have been obtained using the known note values that have been read from the MIDI files. As discussed above, when audio signals have to be examined instead of MIDI, this knowledge can only be obtained by means of beat tracking, which is an unsolved

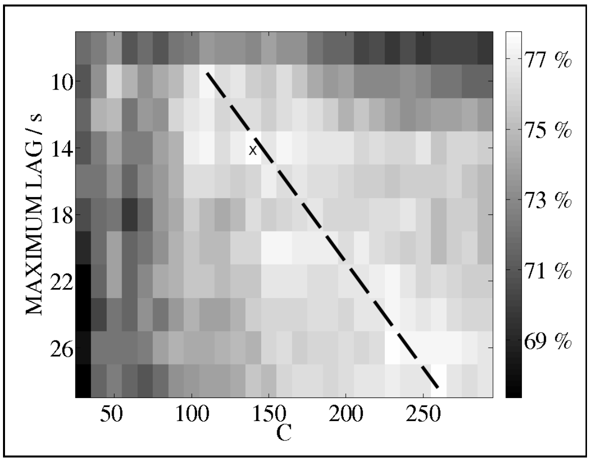

Figure 4. Result of the parameter grid search using the STM descriptors

		Predicted					
		9/8	10/8	8/8	3/4	4/4	5/8
Notated	9/8	51	3	3	1	3	3
	10/8	0	52	2	0	0	3
	8/8	1	1	30	2	11	2
	3/4	3	0	3	15	1	0
	4/4	0	2	8	1	48	1
	5/8	2	4	3	0	1	28

Table 5. Confusion matrix for STM at $C = 140$ and maximum lag of $14s$

task for the time signatures obtained in the data set. Thus, the STM represent a solution to avoid beat tracking, and in this Section the influence of its application on the resulting accuracies will be documented.

The parameters to set when using the STM are the maximum lag considered in the autocorrelation r_u and the number of scale coefficients C that is to be used when computing the cosine distance.

The influence of these parameters has been evaluated in a grid search. The resulting accuracies are depicted in Figure 4. It can be seen that by increasing the maximum lag size and the maximum scale coefficient the accuracies are improved until a level of about 77% is reached. The highest accuracy achieved at some points on the dotted line in Figure 4 is 77.8%, for example at $C = 140$ and at a maximum lag of $14s$ (marked in Figure 4). Choosing a point with small maximum lag leads to faster computation of the scale transform, and choosing a small value of C means a more compact STM description.

The related confusion matrix is shown in Table 5 and comparing it with the confusion matrix shown in Table 3 reveals very similar structure. The decrease in accuracy seems to be caused by some misclassification that cannot be justified by a similarity of the *usul*, as for example the $\frac{9}{8}$-time signature, which for the STM descriptor is randomly misclassified. Thus it appears that transforming autocorrelations to scale domain in the proposed way introduces some noise to the rhythm descriptors. However, the performance is only 2.4% lower than for using the scale-free autocorrelations (77.8% instead of 80.2%). Hence, by including scale transform the currently infeasible step of beat tracking in this kind of meters is avoided and time signature estimation is made feasible, when presented with arbitrary types of music signals having a compound or complex meter.

5. FUTURE WORK: TOWARDS AUDIO SIGNALS

As mentioned in [18], in order for the above described approach to work on audio instead of MIDI three algorithmic steps have to be added: onset detection, pitch estimation and beat tracking. The first step appears to be necessary, because the onset locations are not known as it is the case for MIDI. The pitch estimation is necessary only when the weights in the OSS are desired to be influenced by the pitch properties of the melody. On audio data, this can be approached using a fundamental frequency based OSS as proposed in [12], otherwise this step can be left out and an OSS as described in [5] can be used instead. The most error-prone step when dealing with audio is the beat tracking: it is necessary to correctly estimate all metric levels in order to determine the eighth note pulse of the piece, when the method as described in Section 3.1.1 is desired to be applied. Fortunately, the results using the STM as described in Section 3.1.2 avoids this step of beat tracking. Thus, time signatures and rhythmic properties can be captured by computing an OSS from an audio signal, and computing ACF and STM as described above. In order to evaluate the accuracy of the approach on audio data, a set of audio recordings similar to the MIDI data set will have to be compiled.

6. CONCLUSIONS

In this paper the application of scale transform for the recognition of time signatures is proposed. Using a data set of MIDI data with high class intern tempo deviations it is shown that this method achieves almost the same accuracy as a method that assumes that the metric levels of the piece are known. Thus, this method can be applied to the time signature recognition of audio signals by estimating an OSS suitable for the character of the signal and then computing the STM descriptors as proposed. This represents a significant achievement because the estimation of the metric levels in music signals having compound or complex meters is not a solved problem. The proposed approach is computationally simple because the scale transform can be performed using FFT algorithms. Furthermore, the proposed descriptors seem to capture a reasonable amount of information about the rhythmic properties of the *usul*, as could be seen in the relation between symbolic similarity and the confusion. As the rhythmic properties of Turkish music have never been studied using computational methods, this indicates an interesting direction for future studies. Next steps of these studies have to be the usage of audio signals and the examination of *usul* of same length.

7. REFERENCES

[1] Iasonas Antonopoulos, Angelos Pikrakis, Sergios Theodoridis, Olmo Cornelis, Dirk Moelants, and Marc Leman. Music retrieval by rhythmic similarity applied on greek and african traditional music. In *Proc. of ISMIR - International Conference on Music Information Retrieval*, Vienna, Austria, 2007.

[2] Jeffrey A. Bilmes. *Timing is of the Essence*. PhD thesis, Master Thesis, Massachusetts Institute Of Technology, 1993.

[3] L. Cohen. The scale representation. *IEEE Transactions on Signal Processing*, 41(12):3275–3292, 1993.

[4] Simon Dixon, Fabien Gouyon, and Gerhard Widmer. Towards characterisation of music via rhythmic patterns. In *Proc. of ISMIR - International Conference on Music Information Retrieval*, 2004.

[5] Daniel P. W. Ellis. Beat tracking by dynamic programming. *Journal of New Music Research*, 36(1):51–60, 2007.

[6] Jonathan Foote, Matthew D. Cooper, and Unjung Nam. Audio retrieval by rhythmic similarity. In *Proc. of ISMIR - International Conference on Music Information Retrieval*, pages 265–266, 2002.

[7] Ali C. Gedik and Baris Bozkurt. Automatic classification of turkish traditional art music recordings by ariel theory. In *Proc. of CIM08, 4th Conference on Interdisciplinary Musicology*, Thessaloniki, Greece, 2008.

[8] Fabian Gouyon and Perfecto Herrera. Determination of the meter of musical audio signals: Seeking recurrences in beat segment descriptors. In *114th Convention of the Audio Engineering Society*, 2003.

[9] Andre Holzapfel and Yannis Stylianou. Musical genre classification using non-negative matrix factorization based features. *IEEE Transactions on Audio, Speech and Language Procesing*, 16(2):424–434, 2008.

[10] Andre Holzapfel and Yannis Stylianou. Rhythmic similarity of music based on dynamic periodicity warping. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP*, pages 2217–2220, 2008.

[11] Andre Holzapfel and Yannis Stylianou. A scale transform based method for rhythmic similarity of music. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP*, pages 317–320, 2009.

[12] Andre Holzapfel, Yannis Stylianou, Ali C. Gedik, and Barı¸s Bozkurt. Three dimensions of pitched instrument onset detection. *Accepted for publication in IEEE Trans. on Audio, Speech and Language Procesing*, 2009.

[13] M. K. Karaosmanoğlu, Süleyman Metin Yılmaz, Ömer Tören, Secgi Ceran, Utku Uzmen, Gülhan Cihan, and Emre Başaran. *Mus2okur*. Data-Soft Ltd., http://www.musiki.org/, Turkey, 2008.

[14] A. P. Klapuri, A. J. Eronen, and J. T. Astola. Analysis of the meter of acoustic musical signals. *IEEE Transactions on Acoustics Speech and Signal Processing*, 14(1):342–355, 2006.

[15] R. Parncutt. A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11(4):409–464, 1994.

[16] Tolga Bektaş. Relationships between prosodic and musical meters in the beste form of classical turkish music. *Asian Music*, 36(1), Winter/Spring 2005.

[17] M. T. Thomassen. Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, 71:1596–1605, 1982.

[18] Petri Toiviainen and Tuomas Eerola. Autocorrelation in meter induction: The role of accent structure. *Journal of the Acoustical Society of America*, 119(2):1164–1170, 2006.

[19] Godfried T. Toussaint. A comparison of rhythmic similarity measures. In *Proc. of ISMIR - International Conference on Music Information Retrieval*, 2004.

[20] George Tzanetakis, Ajay Kapur, Andrew Schloss, and Matthew Wright. Computational ethnomusicology. *Journal of interdisciplinary music studies*, 1(2):1–24, 2007.

[21] Christian Uhle and Juergen Herre. Estimation of tempo, micro time and time signature from percussive music. In *Proc. of the Int. Conference on Digital Audio Effects (DAFx)*, 2003.

[22] W.J. Williams and E.J. Zalubas. Helicopter transmission fault detection via time-frequency, scale and spectral methods. *Mechanical systems and signal processing*, 14(4):545–559, July 2000.