Classification of Dance Music by Periodicity Patterns

Simon Dixon
Austrian Research Institute for AI
Freyung 6/6, Vienna 1010, Austria
simon@oefai.at

Elias Pampalk
Austrian Research Institute for AI
Freyung 6/6, Vienna 1010, Austria
elias@oefai.at

Gerhard Widmer
Austrian Research Institute for AI
Freyung 6/6, Vienna 1010, Austria
and
Department of Medical Cybernetics and AI
University of Vienna
gerhard@oefai.at

Abstract

This paper addresses the genre classification problem for a specific subset of music, standard and Latin ballroom dance music, using a classification method based only on timing information. We compare two methods of extracting periodicities from audio recordings in order to find the metrical hierarchy and timing patterns by which the style of the music can be recognised: the first method performs onset detection and clustering of inter-onset intervals; the second uses autocorrelation on the amplitude envelopes of band-limited versions of the signal as its method of periodicity detection. The relationships between periodicities are then used to find the metrical hierarchy and to estimate the tempo at the beat and measure levels of the hierarchy. The periodicities are then interpreted as musical note values, and the estimated tempo, meter and the distribution of periodicities are used to predict the style of music using a simple set of rules. The methods are evaluated with a test set of standard and Latin dance music, for which the style and tempo are given on the CD cover, providing a "ground truth" by which the automatic classification can be measured.

1 Introduction

Genre classification is an important problem in music information retrieval. Automatic classification at a coarse level, such as distinguishing classical from rock music, is not a difficult problem, but more fine-grained distinctions amongst pieces sharing similar characteristics are more difficult to establish (Tzanetakis and Cook, 2002). In this paper we consider the recognition of genre within the various styles of standard and Latin ballroom dance music. These styles have certain common characteristics (for example, a strong beat and a mostly constant tempo), but at the same time have clearly recognisable differences (consider tango, waltz and jive), which humans are usually able to distinguish with minimal training. Since the major feature of dance music is rhythm, this paper focusses entirely on classification based on temporal features of the music, although we recognise that other features (such as instrumentation and articulation) are also important in helping dancers to choose the appropriate dance style for a particular piece of music.

We compare two methods of generating a ranked list of periodicities from audio recordings in order to find the metrical hierarchy and timing patterns by which the style of the music can be recognised. The first method is based on an onset detection algorithm taken from a performance analysis and visualisation system (Dixon et al., 2002), which processes the audio signal by detecting onsets of musical notes, calculates the time intervals between pairs of onsets, and uses a clustering algorithm to find the significant periodicities in the music. The second method is based on a system for calculating the similarity of rhythmic patterns (Paulus and Klapuri, 2002), which splits the audio signal into a number of frequency bands, smooths each one to produce a set of amplitude envelopes, and finds periodicities in each frequency band as peaks in the autocorrelation function.

The periodicities are then processed to find the best-fitting metrical hierarchy, by assigning each periodicity to a musical note value, expressed as a simple integer fraction representing the number of beats. The distribution of note values and their weights, as well as the rounding errors, are used in determining the most likely metrical structure, which in turn determines the tempo and meter of the music. Finally, a simple rule-based system is used to classify the piece by dance style, based on the tempo, meter, patterns of periodicities and their strengths.

It is not clear that periodicity patterns provide sufficient information to correctly classify all dance music. No tests have been made with human subjects to compare their performance on such a task. As it stands, the greatest source of error is in the selection of the metrical hierarchy. Once this is correctly determined, classification accuracy compares favourably with other systems. The results from a test set of over 100 standard and Latin ballroom dance pieces indicate that when the tempo and meter are correctly estimated, the style recognition rules attain up to 80% accuracy.

The following 2 sections describe the two periodicity detection methods respectively, and then in section 4, the algorithm for determining tempo, meter and finally dance style is presented. Section 5 contains the results of testing on a set of dance CDs, and the paper concludes with a discussion of the results and an outline of planned future work.

2 Inter-Onset Interval Clustering

Most rhythmic information is conveyed by the timing of the beginnings (onsets) of notes. For this reason, many tempo induction and beat tracking systems which work with audio input start by estimating the onset times of musical notes and/or percussive sounds (e.g. Large and Kolen, 1994; Goto and Muraoka, 1995; Dixon, 2001). Other systems work with MIDI input (e.g. Rosenthal, 1992) and necessarily use the onset times in their processing. The subsequent processing is then performed symbolically, without further reference to the audio data, in which case onset detection can be seen as a preprocessing step for an algorithm that is not audio-based. Tempo information is then derived from the time durations between pairs of onsets, which correspond to various rhythmic units, such as quarter notes, half notes and dotted quarter notes. These durations are called inter-onset intervals (IOIs), referring to the time intervals between both consecutive and non-consecutive pairs of onsets. Assuming the tempo does not vary greatly during the analysis period, clustering of similar IOIs will reveal the main periodicities in the music and filter out most spuriously detected onsets. This section describes a periodicity detection algorithm based on Dixon et al. (2002).

2.1 Audio Processing

The audio input is read in linear PCM format (after being converted from a compressed format if necessary). If the input has more than one channel, a single channel signal is created by averaging all channels. The audio data is processed in blocks by a smoothing filter which calculates the RMS amplitude for 40ms blocks of data, using a 10ms hop size. Onsets are detected using a simple time-domain method, which finds local peaks in the slope of this smoothed amplitude envelope (see figure 1), where the slope is calculated using a 4-point linear regression. Thresholds in amplitude and slope are used to delete spurious peaks, and the remaining peaks are taken to be the note onset times. Although a relatively simple algorithm is used for event detection, it has been shown that it works sufficiently well for the successful extraction of tempo.

2.2 Clustering

The onset times are used by the clustering algorithm given in figure 2 to find significant periodicities in the data. The clustering algorithm begins by calculating all IOIs between pairs of onsets up to 5 seconds apart, weighting the intervals by the geometric mean of the amplitudes of the onsets, and summing across equally spaced onset pairs, to give a weighted IOI histogram, as shown in figure 3.

The IOIs are then clustered using an iterative best-first algorithm which sequentially finds the cluster with the greatest average amplitude, marks its IOIs as used, and continues searching for the next cluster. The width of a cluster is adjusted according to the IOI duration, so that clusters representing longer durations allow greater variation in the IOIs. Each cluster C_i is ranked by its weight S_i, which is calculated as the average weight $\overline{w_j}$ of its component IOIs, and the centroid T_i of each cluster is calculated as the weighted average of the IOIs t_j themselves. The next best clusters and their weights are calculated by marking the IOIs which have been used in a previous cluster and repeating the above calculations ignoring the marked IOIs.

It is usually the case in traditional Western music that time inter-

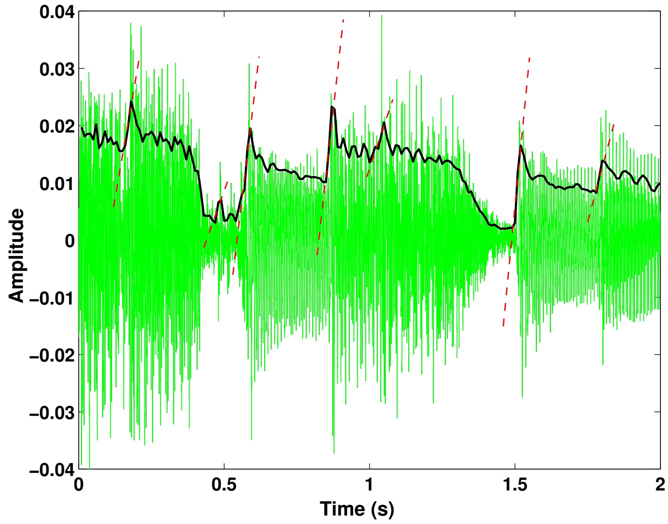

Figure 1: Onset detection method, showing the audio signal with the smoothed amplitude envelope overlaid in bold and the peaks in the slope marked by dashed lines.

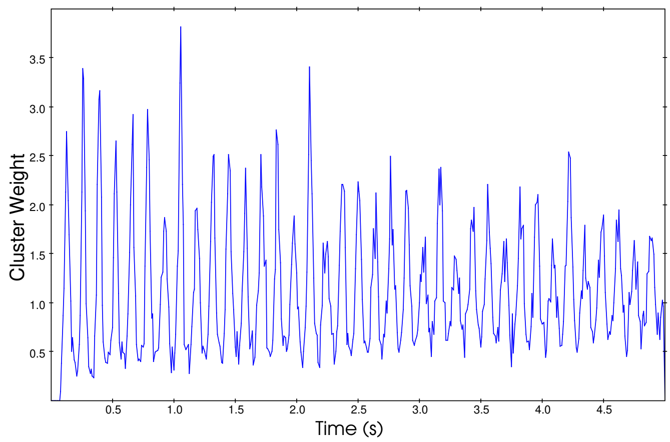

Figure 3: Example showing weighted IOI histogram for a samba piece at 56 measures per minute (224 BPM). The 3 highest peaks correspond to 4 beats (one measure), 8 beats and 1 beat, respectively.

```
For times t_i from 0.1s to 5.0s in 0.01s steps
    Find all pairs of onsets which are t_i apart
    Weight w_i = sum of the mean amplitude of the onset pairs
While there are unmarked IOIs
    For times t_i from 0.1s to 5.0s in 0.01s steps
        Cluster width s_i = 0.01⌊t_i/0.3 + 8⌋
        Find average amplitude of unmarked IOIs in window [t_i, t_i + s_i]
    Find t_M which gives maximum average amplitude
    Create a cluster containing the IOIs in the range [t_M, t_M + s_M]
    Mark the IOIs in the cluster as used
For each cluster
    Find related clusters (multiples or divisors)
    Adjust related clusters using weighted average
```

Figure 2: Algorithm for clustering of inter-onset intervals

vals are approximately related by small integer ratios; the cluster centroids also tend to reflect this property. In other words, the cluster centroids are not independent; they represent related musical units such as quarter notes and half notes. Since the clusters are part of a single metrical hierarchy, the centroids can be used to correct each other, since we expect them to exhibit simple integer fraction ratios. Thus an error in a single cluster can be corrected by reference to the other clusters. This is the final step of periodicity detection, where the cluster centroids and weights are adjusted based on the combined information given by all of their related clusters. The cluster centres and their weights define a ranked list of periodicities which are then used in determining tempo, meter and style.

3 Periodicity Detection with Autocorrelation

An alternative approach to periodicity detection uses autocorrelation. This method has been used for detecting the meter of musical scores by Brown (1993), and for pulse tracking by Scheirer (1997), using the Meddis and Hewitt pitch model with much larger time windows. We base this work on the more recent research of Paulus and Klapuri (2002), which was developed for measuring the similarity of rhythmic patterns.

3.1 Audio Processing

The second method of periodicity detection was implemented by converting the audio input data to the mid-level representation advocated by Paulus and Klapuri. The suggested preprocessing step which removes sinusoids from the signal was not performed, since we could not guarantee the existence of drums in all pieces in our data set. The aim of the audio processing step was to reduce the audio signal to a set of amplitude envelopes from which temporal information could be derived.

The audio data was taken from audio CDs, converted to a single channel by averaging, and then passed through an 8 channel filter bank, with the first band up to 100 Hz and the remaining bands equally spaced (logarithmically) at just over an octave wide to cover the full frequency range of the signal. Then for each of the 8 frequency bands, the signal was rectified, squared, decimated to a sampling rate of 980Hz, and smoothed with a 20Hz low-pass filter. Finally the dynamic range was compressed using a logarithmic function.

3.2 Periodicity Calculation

Periodicities are found in each frequency band by examining the peaks of the autocorrelation function for time lags between 0 and 5 seconds. After normalising the autocorrelation by the magnitude of the lag 0 value, this peak is discarded, and the three highest peaks are collected from each frequency band. Figure 4 shows an example of these peaks for a samba piece at 224 beats per minute. The periodicities corresponding to the beat (268ms) and the measure (1057 ms) are prominent in several frequency bands.

Rather than summing the autocorrelation results across the various frequency bands, we match peaks (periodicities) in different frequency bands which differ by less than 20ms. Matched sets of periodicities are combined by averaging the period. A weight S_i for each resulting periodicity T_i is calculated as the sum of the mean autocorrelation value and the number of matched periodicities in the set. The weights S_i are used to rank the elements in the list of periodicities; these values are used in the subsequent estimation of tempo, meter and style.

4 Determining Tempo, Meter and Style

Both of the methods discussed in the previous sections are generally successful in finding peaks at the periodicities corresponding to the beat and measure level of the metrical hierarchy. The difficulty is that peaks also occur at higher, lower and intervening metrical levels, as well as at commonly occurring note durations which are not directly part of the metrical hierarchy, for example the dotted quarter note in a samba rhythm.

We use an exhaustive approach to evaluate the suitability of each periodicity as the measure level of the metrical hierarchy. For each periodicity T_k, the ratio r_i of the other periodicities T_i to T_k is expressed as a simple integer fraction $\frac{p_i}{q_i}$, attempting to keep p_i and q_i small while minimising the error $|p_i - q_i r_i|$ for each i. Formally, the constraints are as follows for each i:

$$p_i < 16 \quad (\text{for } q_i > 1)$$

$$q_i \in \{1, 2, 3, 4, 6, 8, 12, 16\}$$

$$\mathrm{GCD}(p_i, q_i) = 1$$

$$\neg\exists\, p', q' \text{ such that } |p' - q' * r_i| \; < \; |p_i - q_i * r_i|$$

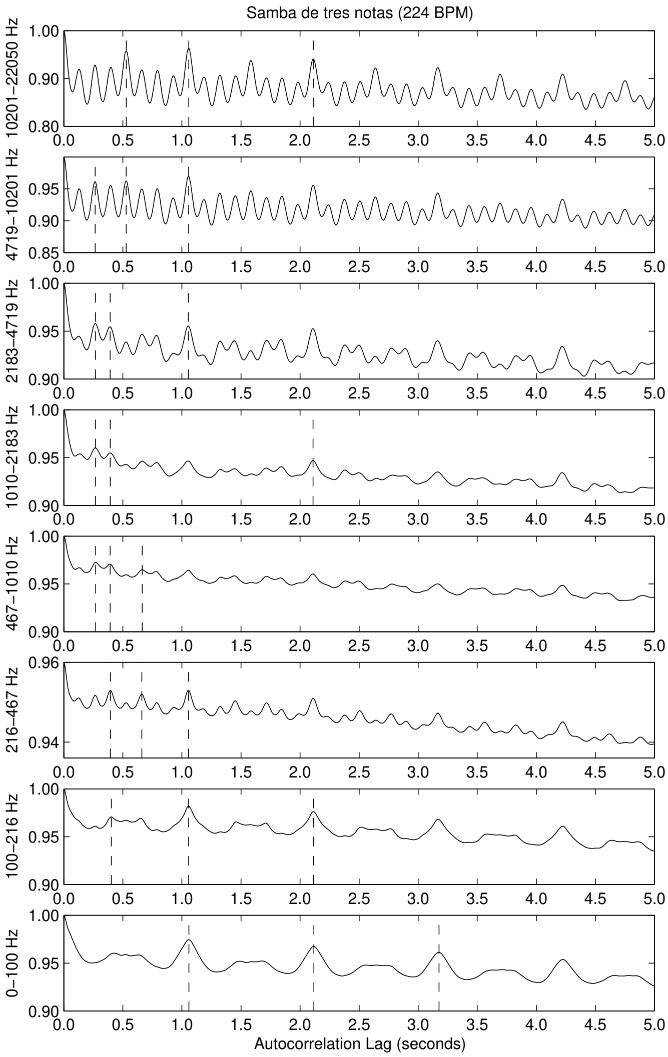

Figure 4: Autocorrelation of the amplitude envelopes for each frequency band, showing lags up to 5 seconds. The top three peaks (excluding the lag 0 peak) are marked by a dashed vertical line. The input is the same piece as used in figure 3.

Style	Tempo Range				Met.	#
	Actual		Suggested			
	Min	Max	Min	Max		
Blues	20	20	-	-	4	2
Rumba	26	29	26	26	4	11
Tango	28	33	30	32	4	10
Slow fox	29	30	28	30	4	5
Disco	30	30	-	-	4	2
Slow waltz	29	30	28	30	3	9
Cha cha	30	33	32	32	4	27
Jive	32	44	44	44	4	21
Rock and roll	42	42	-	-	4	1
Boogie	44	46	-	-	4	2
Foxtrot	44	50	-	-	4	3
Quickstep	50	52	50	52	4	7
Samba	50	63	50	50	4	9
Mambo	56	56	-	-	4	4
Viennese waltz	43[1]	62	58	60	3	9
Paso doble	60	65	-	-	2	6
Polka	76	76	-	-	4	1
Miscellaneous	-	-	-	-	4	32

Figure 5: The distribution of styles in the test set, including the range of tempos (where given), the suggested tempo ranges for each styles taken from Ballroomdancers.com (2002) (where available), and the meter and number of pieces in that style.

The next step involves calculating a weighted sum of the periodicity weights S_i described in previous sections. This is computed separately for even and odd meters, since the distribution patterns vary depending on the meter. For periodicity T_i with weight S_i, the weighted sums $S^*_{i,e}$ and $S^*_{i,o}$ are given by:

$$S^*_{i,e} = \sum_{i \neq k} w_e(p_i, q_i)\ S_i\ u(T_i)\ v(E_i)$$

$$S^*_{i,o} = \sum_{i \neq k} w_o(p_i, q_i)\ S_i\ u(T_i)\ v(E_i)$$

where w_e and w_o are empirically determined matrices of weights for even and odd meters respectively, u is a tempo weighting function which restricts the range of allowed tempos, v is an error weighting function which penalises periodicities which deviate from the supposed simple integer fraction relationship, and $E_i = |p_i - q_i r_i|$ represents the error in the rational approximation $\frac{p_i}{q_i}$. The maximum value of $S^*_{j,k}$ determines the periodicity of the measure level and the meter. We currently assume the meter is either $\frac{3}{4}$ or $\frac{4}{4}$, so the quarter note (beat) level of the metrical hierarchy is also determined by this step.

The final step is to choose the style, which we do using a simple rule-based approach, which combines information from the tempo, meter and periodicity distribution to make its decision. The allowed tempo range and the meter of the dances, given in figure 5, are used as constraints, and in cases where multiple dances are possible for a given tempo and meter, the periodicity distribution is used to distinguish between genres.

The simplest rules are for the pieces in triple meter, since there are only two genres, slow waltz and Viennese waltz, and these

[1]The Viennese waltz at 43 MPM is an outlier; all other instances of this style are at least 60 MPM.

are separated by a large tempo difference. The following 2 rules, expressed in Prolog-like notation, ensure correct classification of instances of these classes:

```
viennesewaltz(Meter, Tempo) :-
        Meter = 3,
        Tempo > 40.
slowwaltz(Meter, Tempo) :-
        Meter = 3,
        Tempo <= 40.
```

The remaining rules are used for duple and quadruple meters, conceptually by first dividing the pieces into tempo ranges, and then using the strengths of various periodicities to choose the most likely genre. For ease of reading, we express each of the rules independently; the implementation is otherwise.

```
polka(Meter, Tempo) :-
        Meter = 4,
        Tempo > 70.
pasodoble(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 70,
        Tempo > 58,
        weight(3/8) <= 3.
quickstep(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 54,
        Tempo > 48,
        weight(3/8) <= 3.
samba(Meter, Tempo) :-
        Meter = 4,
        (
            (   Tempo <= 70,
                Tempo > 48,
                weight(3/8) > 3
            );
            (   Tempo <= 58,
                Tempo > 54
            )
        ).
jive(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 48,
        Tempo > 35.
slowfox(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 35,
        Tempo > 29,
        maxWeightAt(1/2).
chacha(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 35,
        Tempo > 29,
        !maxWeightAt(1/2),
        weight(_/8) > 4.
tango(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 35,
        Tempo > 29,
        !maxWeightAt(1/2),
        weight(_/8) <= 4.
rumba(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 29,
        Tempo > 25.
blues(Meter, Tempo) :-
        Meter = 4,
        Tempo <= 25.
```

We briefly explain the two most complex rules, those for the samba and cha cha. The tempo of the samba has quite a wide range of possible values, so that it overlaps the paso doble at the higher end and the quickstep at the lower end. To distinguish these dances in the overlapping parts of the tempo range, we take advantage of the fact that the samba rhythm tends to have a strong periodicity at the dotted quarter note level (see figures 3 and 4) which is not present in the other dances which overlap its tempo range. The weight S_i of this periodicity is compared with a threshold, and if higher, the piece is classified as samba, otherwise, the tempo determines the classification.

The tempo range of the cha cha coincides with that of the slow fox and the tango. In this case, the weight of the half, quarter and eighth notes are used to classify the piece. If the periodicity with maximum weight is the half note level, then piece is classified as slow fox, otherwise the weight of the sum of the eighth note periodicities (i.e. $\frac{1}{8}$, $\frac{3}{8}$, $\frac{5}{8}$, ...) determines the genre, with a high value indicating cha cha and a low value indicating tango.

The current rule set was constructed in an ad hoc fashion, as a proof of concept, that a reasonable level of classification can be obtained based only on periodicity information. The danger with such a small set of data is that the rules overfit the data set and generalise poorly. For this reason, we omitted development of rules for some of the styles which have very few instances, and we did not try a more principled approach to rule generation using machine learning. If we are able to obtain more data, this is an interesting avenue for further work.

5 Tests and Results

One of the major difficulties in evaluating systems that deal with music similarity and style is that there is no ground truth, that is, no objective evaluation criteria or standard test sets. Instead, category definitions are subjective, they change over time, and most music consists of elements from a mixture of different categories. Since the current work focusses on temporal features, we chose a test set of music where rhythm is an important element, and for which somewhat objective evaluation criteria are available.

The test set consists of standard and Latin dance music, which is subdivided into various styles such as tango, slow waltz, Viennese waltz, foxtrot, quickstep, jive, cha cha, samba and rumba (see figure 5). Each of these styles are reasonably well-defined, in that dancers are generally able to identify and agree upon the style of such a piece within the first few bars of music, as evidenced by the general uniformity of the chosen dance styles at a ball. Furthermore, the test set consisted of CDs where the style of dance and/or tempo are printed on the CD cover, providing an independent means of evaluation. But despite the apparent clarity, there is known to be some degree of overlap between the various dance styles, such that some pieces of music fit more than one type of dance, so perfect results are never to be expected. The test set consists of 161 pieces for which the style

	IOI-Clustering	Correlation
Tempo	53/96	65/96
Meter	142/161	150/161
Style	36/52	52/65

Figure 7: Summary results for recognition of tempo, meter and style.

	IOI-Clustering	Correlation
Half tempo	4	10
Double tempo	24	16
Wrong meter	14	5
Other	1	0

Figure 8: Counts of each type of tempo error.

is given, and 96 for which the tempo is given. There are 17 different styles for which the tempo is given (but not for all instances), plus a further 9 styles for which the tempo is never given (the row marked "miscellaneous" in figure 5).

The first results deal with the periodicity detection algorithms. Both algorithms produce a ranked list of periodicities, and we compare these with the tempo and style printed on the CD cover. We call the periodicities corresponding to the measure and beat levels of the metrical hierarchy the measure period and the beat period respectively. Figure 6 shows, for each position from 1 to 10, the number of songs for which the measure period (respectively the beat period) was ranked at this position. From these results it appears that the correlation method ranks the important periodicities higher, but it is also the case that the correlation method produces shorter lists of periodicities, which may explain both the higher ranking and the higher number of unranked beat and measure periods.

The main results are shown in figure 7. The first row shows the number of songs for which the calculated measure period agrees with the given tempo on the CD (plus or minus 3 measures per minute). We use measures per minute for tempo instead of the more usual beats per minute, because this is the format of the details printed on the CD liner notes. This value will be wrong if either the beats per minute or the meter is wrong.

We examine the nature of the tempo errors in figure 8. The majority of errors are due to selecting the wrong metrical level, that is choosing a measure period which is half or double the correct (i.e. notated) value. In other words, the system chose a musically plausible solution which didn't correspond with the intention of the musicians. This is a common problem which is reported in tempo induction systems (Dixon, 2001) and a phenomenon that also occurs in tapping experiments with human subjects (Dixon and Goebl, 2002).

All other errors except one were due to selecting the wrong meter, so that even if the beat period were correct, the measure period would be wrong because it contains the wrong number of beats. The remaining error occurred on a piece that contains many triplets, and the system chose the triplets as the beat level, but (surprisingly) also chose a binary grouping of these triplets as the measure level. It is unlikely that people would make these types of errors.

The second row of results (figure 7) shows the meter recognition results, which appear to be very good. However, there are only 18 pieces in $\frac{3}{4}$ time, so the IOI-clustering results would be improved (marginally) by replacing this part of the system by one that always predicts $\frac{4}{4}$ time! More data is required to determine how well this part of the system really functions. In recent work, Gouyon and Herrera (2003) report over 90% accuracy in distinguishing duple and triple meters.

The style recognition results range from 69% (for the IOI-clustering data) to 80% (for the autocorrelation data), assuming the data set is restricted to pieces for which the tempo is correctly recognised. (Since none of the dance genres has a tempo range wide enough to accommodate a factor of two error, it is impossible for the system to predict style correctly once it has the wrong tempo. The wrong meter also guarantees failure in style recognition, but these were not deleted from the style results.) The confusion matrix (figure 9) shows the nature of the classification errors. Some errors, such as the confusion of cha cha with tango and slow fox, show a weakness in the classification rules, whereas others, such as classifying boogie and rock and roll as jive, are to be expected, as the genres are very closely related. In fact, since there are no rules for boogie or rock and roll, these pieces could not be correctly classified by the present system.

6 Conclusion

We presented a comparison of two methods of generating a ranked list of periodicities from an audio file, and found that an autocorrelation-based approach gave better results than one based on processing discretely detected onsets. The periodicity patterns were used to predict tempo, meter and genre of different types of dance music with some success. The major source of error was in choosing the periodicity which corresponds to the measure level of the music. When this was correctly chosen, the classification of autocorrelation-based periodicities reached 80% success. This is particularly surprising when one considers that no rhythmic patterns (i.e. sequences of durations) were used, nor timbral, nor melodic, nor harmonic features. Tzanetakis and Cook (2002) report a 61% success rate for classifying music into 10 (non-similar) genres, using features representing timbre, rhythm and pitch. The overall success rate in this work (including tempo detection errors) is perhaps lower, but it is impossible to make a meaningful comparison due to the different nature of the tasks, methods and data.

The current work is limited by the small test set and the accompanying danger of overfitting. In future work we hope to build a larger test set and investigate the use of automatic classification techniques.

Periodicities give information about the metrical structure of the music, but not the rhythmic structure, which arises from the relative timing of onsets within and between the various frequency bands. A fruitful area for further work would be to extract and encode commonly occurring rhythmic patterns, which is (intuitively at least) a better way of identifying genres of dance music. (Note that this is very different to examining periodicity distributions.) As a starting point, there is a large body of literature on beat tracking involving the analysis of sequences of temporal events in order to estimate tempo, meter and metrical boundaries (see Dixon, 2001, for an overview).

Rank:		1	2	3	4	5	6	7	8	9	10	none
Method 1:	Measure	13	16	16	9	11	13	13	3	2	0	0
IOI-Clustering	Beat	11	16	18	15	9	12	10	1	1	0	3
Method 2:	Measure	19	20	25	10	13	1	1	0	0	0	7
Correlation	Beat	30	25	20	7	4	0	0	0	0	0	10

Figure 6: Position of the periodicities corresponding to the bar and measure levels in the ranked lists of IOI clusters and autocorrelation peaks.

	PD	SA	TA	SF	QU	RR	RU	SW	CH	BO	WW	FO	JI	MA
PD	5	-	-	-	-	-	-	-	-	-	-	-	-	-
SA	-	3	-	-	-	-	-	-	-	-	-	-	-	1
TA	-	-	6	-	-	-	-	-	1	-	-	-	-	-
SF	-	-	-	4	-	-	-	-	1	-	-	-	1	-
QU	-	3	-	-	6	-	-	-	-	-	-	-	-	-
RR	-	-	-	-	-	-	-	-	-	-	-	-	-	-
RU	-	-	1	-	-	-	4	-	-	-	-	-	-	-
SW	-	-	-	-	-	-	-	6	-	-	-	-	-	-
CH	-	-	-	-	-	-	-	-	-	-	-	-	-	-
BO	-	-	-	-	-	-	-	-	-	-	-	-	-	-
WW	-	-	-	-	-	-	-	-	-	-	2	-	-	-
FO	-	-	-	-	-	-	-	-	-	-	-	-	-	-
JI	-	-	-	-	-	1	-	-	-	2	-	2	16	-
MA	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 9: Confusion matrix for correlation-based classification. The columns refer to the actual dance style, and the rows the predicted style. The abbreviations for the dance styles are: paso doble (PD), samba (SA), tango (TA), slow fox (SF), quickstep (QU), rock and roll (RR), rumba (RU), slow waltz (SW), cha cha (CH), boogie (BO), Viennese waltz (WW), foxtrot (FO), jive (JI) and mambo (MA).

Acknowledgements

This research is part of the project Y99-INF, sponsored by the Austrian Federal Ministry of Education, Science and Culture (BMBWK) in the form of a START Research Prize. The BMBWK also provides financial support to the Austrian Research Institute for Artificial Intelligence. Thanks also to the anonymous reviewers of this paper for their insightful comments.

References

Ballroomdancers.com (2002). Learn the dances. Retrieved 24 April, 2003, from:
`http://www.ballroomdancers.com/Dances/`.

Brown, J. (1993). Determination of the meter of musical scores by autocorrelation. *Journal of the Acoustical Society of America*, 94(4):1953–1957.

Dixon, S. (2001). Automatic extraction of tempo and beat from expressive performances. *Journal of New Music Research*, 30(1):39–58.

Dixon, S. and Goebl, W. (2002). Pinpointing the beat: Tapping to expressive performances. In *7th International Conference on Music Perception and Cognition (ICMPC7)*, pages 617–620, Sydney, Australia.

Dixon, S., Goebl, W., and Widmer, G. (2002). Real time tracking and visualisation of musical expression. In *Music and Artificial Intelligence: Second International Conference, ICMAI2002*, pages 58–68, Edinburgh, Scotland. Springer.

Goto, M. and Muraoka, Y. (1995). A real-time beat tracking system for audio signals. In *Proceedings of the International Computer Music Conference*, pages 171–174, San Francisco CA. International Computer Music Association.

Gouyon, F. and Herrera, P. (2003). Determination of the meter of musical audio signals: Seeking recurrences in beat segment descriptors. In *Presented at the 114th Convention of the Audio Engineering Society*, Amsterdam, Netherlands.

Large, E. and Kolen, J. (1994). Resonance and the perception of musical meter. *Connection Science*, 6:177–208.

Paulus, J. and Klapuri, A. (2002). Measuring the similarity of rhythmic patterns. In *Proceedings of the 3rd International Conference on Musical Information Retrieval*. IRCAM Centre Pompidou.

Rosenthal, D. (1992). Emulation of human rhythm perception. *Computer Music Journal*, 16(1):64–76.

Scheirer, E. (1997). Pulse tracking with a pitch tracker. In *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk, NY.

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