Audio Engineering Society

Convention Paper

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain



# Detrended Fluctuation Analysis of Music Signals: Danceability Estimation and further Semantic Characterization

Sebastian Streich, Perfecto Herrera

Universitat Pompeu Fabra, Barcelona, 08003, Spain
{sstreich,pherrera}@iua.upf.es

## ABSTRACT

Detrended fluctuation analysis (DFA) has been proposed by Peng et al. [1] to be used on biomedical data. It originates from fractal analysis and reveals correlations within data series across different time scales. Jennings et al. [2] used a DFA-derived feature, the detrended variance fluctuation exponent, for musical genre classification introducing the method to the music analysis field. In this paper we further exploit the relation of this low-level feature to semantic music descriptions. It was computed on 7750 tracks with manually annotated semantic labels like "Energetic" or "Melancholic". We found statistically strong associations between some of these labels and this feature supporting the hypothesis that it can be linked to a musical attribute which might be described as "danceability".

## 1. INTRODUCTION

In current research on Music Information Retrieval many attempts are made to automatically extract a semantic description of the content from the raw audio signal itself. Relating semantic labels with digital music archives becomes a richer and more convenient form of interaction with a music database than what is currently the case using low-level descriptors. This relationship can be achieved in several different ways: through pure signal processing, through machine learning techniques, or through manual annotation. The latter option however requires a huge effort of human intervention, which might not be feasible or desirable for large collections. Hence, a fully automatic approach bears a clear attractiveness.

In this paper we focus mainly on one very specific low-level feature of a musical audio signal: the detrended variance fluctuation exponent. The method of detrended fluctuation analysis (DFA) was first proposed by Peng et al. in 1994 [1]. Since then it has been used rather extensively in the analysis of time series from biomedical or financial data (e.g. heart rate time series or currency exchange rate time series). It originates from fractal analytical techniques and has the ability to indicate long-range correlations in non-stationary time series. In music analysis so far DFA-derived features are not common and, in contrast to many other low-level features like spectral centroid or zero crossing rate, their potential in this domain has not been exploited yet. Jennings et al. in 2003 [2] reported about their results in musical genre classification by using exclusively the detrended variance fluctuation exponent, a low-level feature derived from DFA of the audio signal. We will consecutively use the term "DFA exponent" in this paper when referring to this feature. Jennings et al. stated in their publication that the DFA exponent can be seen as an indicator for the "danceability" of a piece of music. Hence, we got inspired to examine this particular feature a bit further.

## 2. THE DFA EXPONENT FOR AUDIO DATA

First, since the method of DFA is relatively unknown in music analysis, we dedicate some space to explain the computation process and also focus on some of its characteristics in this particular context. In contrast to commonly used features in music analysis, the DFA exponent can not be computed on a frame-by-frame basis, but comprises a longer-scale statistical analysis of the data. Naturally, it therefore yields a very compact representation. For our experiments we collapsed the feature to only one single value per track. It is interesting to note that DFA is particularly suited for non-stationay data, whereas many methods originating from the signal processing domain rely on quasi stationary properties.

### 2.1. Computation procedure

Following Jennings et al. [2] the audio signal is segmented into non-overlapping blocks of 10ms length. For each block the standard deviation $s(n)$ of the amplitude is computed. The values $s(n)$ resemble a *bounded*, non-stationary time series, which can be associated with the averaged physical intensity of the audio signal in each block (see figure 2.1). In order to obtain the *unbounded* time series $y(m)$, $s(n)$ is integrated:

$$y(m) = \sum_{n=1}^{m} s(n) \qquad (1)$$

This integration step is crucial in the process of DFA computation, because for *bounded* time series the DFA exponent (our final feature) would always be 0 when time scales of greater size are considered. This effect is explained in more detail in [3].

The series $y(m)$ can be thought of as a random walk in one dimension. $y(m)$ is now also segmented into blocks of $\tau$ elements length. This time, we advance only by one sample from one block to the next in the manner of a sliding window. There are two reasons for this extreme overlap. First, we obtain more blocks from the signal, which is of interest, since we will obtain better statistics from a larger number of blocks. Secondly, we avoid possible synchronization with the rhythmical structure of the audio signal, which would lead to arbitrary results depending on the offset we happen to have. However, performing the computation in this manner the number of operations is enormously increased. We will further comment on this in section 2.3.

From each block we now remove the linear trend $\hat{y}_k$ and compute $D(k,\tau)$, the mean of the squared residual:

$$D(k,\tau) = \frac{1}{\tau}\sum_{m=0}^{\tau-1}\left(y(k+m) - \hat{y}_k(m)\right)^2 \qquad (2)$$

We then obtain the detrended fluctuation $F(\tau)$ of the time series by computing the square root of the mean of $D(k,\tau)$ for all $K$ blocks:

$$F(\tau) = \sqrt{\frac{1}{K}\sum_{k=1}^{K} D(k,\tau)} \qquad (3)$$

As indicated, the fluctuation $F$ is a function of $\tau$ (i.e. of the time scale in focus). The goal of DFA is to reveal correlation properties on different time scales. We therefore repeat the process above for different values of $\tau$ that are within the range of our interest. Jennings et al. [2] use a range from 310ms ($\tau = 31$) to 10s not specifying the step size in their paper. Relating these time scales to the musical signal they are reaching from the beat level through the bar level up to a level of simple rhythm patterns.

The DFA exponent $\alpha$ is defined as the slope on a double log graph of $F$ over $\tau$ (eq. 4). It therefore makes sense to increase $\tau$ by a constant multiplication factor rather than a fixed step size. Apart from giving equally spaced supporting points on the logarithmic axis it also reduces the computational operations without affecting the accuracy too much. We chose a factor of 1.1 giving us 36 different values for $\tau$ covering time scales from 310ms to 8.8s.

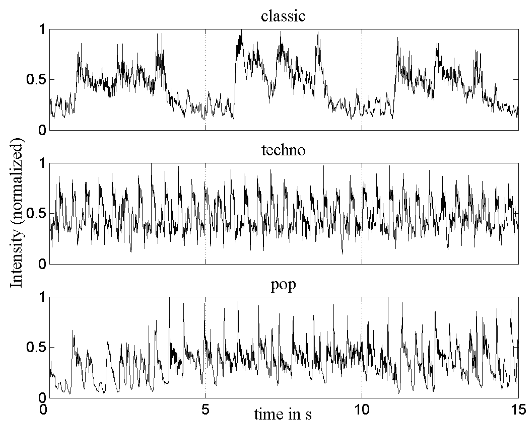


Figure 2.1: Excerpts from the time series $s(n)$ for three example pieces from different musical genres.

For small values of $\tau$ an adjustment is needed in the denominator when computing $\alpha$ (cp. [4]) giving us the following formula for the DFA exponent:

$$\alpha(i) = \frac{\log_{10}\left(F(\tau_{i+1})/F(\tau_i)\right)}{\log_{10}\left((\tau_{i+1}+3)/(\tau_i+3)\right)} \qquad (4)$$

As $\tau$ grows, the influence of the correction becomes negligible. In case that the time series has stable fractal scaling properties within the examined range, the double log graph of $F$ over $\tau$ is a straight line making $\alpha(i)$ a constant function. We find a constant value of 0.5 for a completely random series (white noise), a value of 1 for a series with 1/f-type noise, and 1.5 for a Brown noise series (integrated white noise) [3].

## 2.2. Interpreting the function $\alpha(i)$

For music signals normally we don't have stable scaling properties (see figure 2.2). Opposed to heart rate time series for example, there is much more variance in $\alpha(i)$ for music. Still, we can find that music with sudden jumps in intensity will generally yield a lower level of $\alpha$ than music with a smoother varying series of intensity values. That means music with pronounced percussion events and emphasized note onsets shows lower $\alpha$ values than music with a more floating, steady nature. Apart from that, we will see in the following section that interesting information about the music can be obtained also from the way the stable scaling properties are violated.

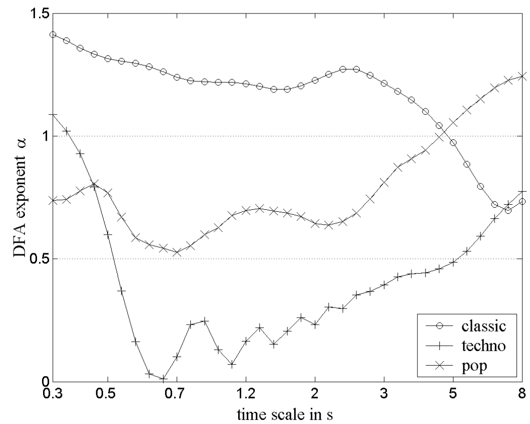


Figure 2.2: DFA exponent functions for the three example tracks.

### 2.2.1. An exemplary comparison

Figure 2.2 shows the evolution of $\alpha$ over the different time scales for three musical pieces. As can be seen, the DFA exponent varies significantly within each single piece. The most stable scaling behavior is found for the classical piece at short time scales, where in contrast the pop piece shows an intermediate and the techno piece a high instability. This is due to the presence of a strong and regular beat pattern in the two latter cases (cp. figure 2.1).

In the techno piece the periodic beat dominates the intensity fluctuation completely since intensity variations on larger time scales are negligible in comparison. This strong periodic trend deteriorates the scaling properties of the series and causes $\alpha$ to drop significantly. Towards larger time scales however, the influence of the periodic intensity variation fades off and $\alpha$ raises back towards its normal level.

In the pop music piece there is also a regular beat, but it is less dominant than in the techno piece. As can be seen in figure 2.1, there are also some noticeable changes in intensity on a larger time scale. Still, $\alpha$ is clearly decreased by the periodic trend. Towards larger time scales, we can observe the same effect as in the techno piece.

For the classical piece no dominant, regular beat pattern can be identified in the time series. Thus, the scaling properties are not affected in the corresponding range. But in contrast to the other two examples the series reveals a larger scale pattern in some parts, which can also be seen in figure 2.1. This causes $\alpha$ to drop in the upper range.

### 2.2.2. A measure of "danceability"

In the previous section we saw how the presence of a strong and regular beat influences the DFA exponent. We can make the simplified assumption, that the presence of a strong and regular beat is also the main characteristic of danceable music. It can further be argued that the easy identification of note onsets or percussion events facilitates a synchronized body movement. In that case it is straightforward to use the DFA exponent as an indicator for the danceability of a musical track (i.e. the easiness with which one can dance to it). The simplest way to do this is to compute the average $\alpha$ of the track, a low value indicating high danceability and vice versa. In figure 2.3 we plotted the average $\alpha$ for 120 tracks, half of them being electronic techno music (in general very danceable), the other half being film score tracks performed with classical orchestra instruments (in general not danceable).

Jennings et al. [2] already found in their experiments that the average values of $\alpha$ for tracks from typical dance music genres were significantly lower than those for tracks from "high art" genres, which are not very danceable in the majority of cases. Since the mean values of the average $\alpha$ for different genres were compared in their study, single exceptions (like e.g. classical polkas or waltzes) didn't influence the general result so much. As it was our first experiment with this descriptor, we followed the simple way of collapsing $\alpha(i)$ into a single value per track by taking the average. In section 5 we will briefly point to some ideas for alternative and advanced exploitation.

Additionally we would like to mention that we consider "danceability" as a special form of a rhythmic complexity description. In [5] we introduced a multifaceted concept of musical complexity and its possible applications in music information retrieval scenarios, like collection browsing, playlist generation, or music recommendation. The DFA exponent appears to form a useful descriptor in this context. It might be further enhanced and combined with other features to model rhythmic complexity in all its facets.

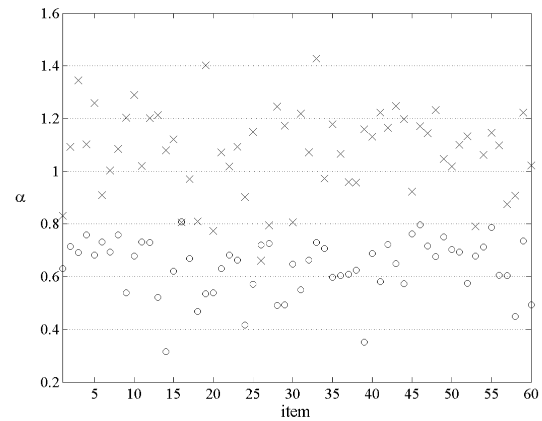


Figure 2.3: $\alpha$ -levels for 60 techno (o) and 60 film score tracks (x), unordered.

## 2.3. Practical issues for the computation

As the computation of the DFA exponent involves statistical methods it is of importance that the time series under examination is sufficiently long. Jennings et al. [2] used 4 minute long excerpts in their study. For our experiment we always used the complete track for the computation. Tracks with less than 30s length were disregarded.

In this context memory consumption can become an issue when very long tracks are being processed. We used a MATLAB implementation and experienced problems with insufficient memory. We therefore processed the tracks in portions of 60s length.

As mentioned in section 2.1. the massive overlap of the blocks for detrending enormously increases the number of operations. Particularly for the larger values of $\tau$ the detrending is computationally expensive. For reasons of efficiency it might therefore be useful to increase the hop size with increasing $\tau$ . In any case it should be kept small relative to $\tau$ in order to avoid unintentional synchronization effects.

Willson and Francis claim the equivalence of the DFA method with a special form of spectral analysis [6]. This alternative would mean a clear efficiency advantage, since the costly detrending operation is not needed there. In a short test we only obtained partly similar results to the DFA method and therefore kept using the latter.

## 3. EXPLOITING THE DFA EXPONENT IN A LARGE COLLECTION

### 3.1. Properties of the collection

We computed the average DFA exponent on a dataset of 7750 tracks from MTG-DB [7], where each track refers to a full piece of music. The dataset also contains manually annotated semantic labels for each item. In our experiments we used the artist names and also "tone" labels consisting in abstract attributes that are associated with the music, such as "Rousing", "Sentimental", or "Theatrical". The list of "tone" labels is composed of a total of 172 different entries. In the statistical analysis only a subset of 136 labels were considered, because the remaining ones appeared less than 100 times each. It must be noted that these labels are originally assigned to the artists and not to the individual tracks. Therefore a certain degree of fuzziness has to be accepted with these descriptions when working on the track level. The data set contains very different, mostly popular styles of music from a total of 289 different artists. A maximum of 34 labels are assigned to a single artist, while the average is 11 labels per artists. Figure 3.1 shows a bar plot of the eight labels that are assigned to the highest number of artists. The average number of artists sharing a label is 18.

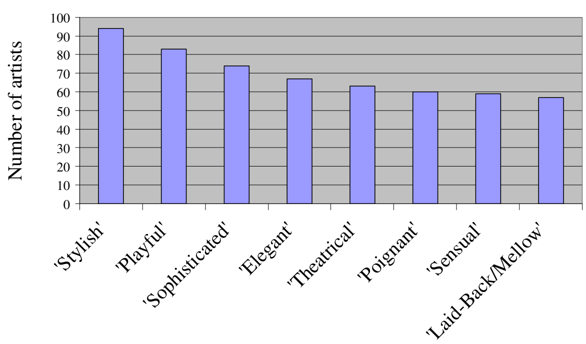


Figure 3.1: Top-eight labels with the highest number of assigned artists.

### 3.2. Experimental procedure

Once the descriptor was computed for the whole set, we started the evaluation by manually checking its consistency. This was done by randomly picking tracks at different levels of $\alpha$ and judging the danceability in direct comparison by listening. A formal user study was not carried out.

As a more objective evaluation we applied also general statistical methods and machine learning methods in order to explore relations between the semantic labels, or certain artists and the DFA exponent. The rationale behind this is to prove a systematic variation of the DFA exponent subject to certain semantic attributes assigned to the music.

#### 3.2.1. Statistical evaluation

For the statistical evaluation we followed two opposed approaches. In one analysis we looked at the distributions of the DFA exponent for all tracks with a certain label. We checked first whether each of the distributions differed significantly from a normal distribution at a 5% level with the Lilliefors test [8]. The Lilliefors test is similar to the Kolmogorov-Smirnoff test for normality, but doesn't require the mean and variance to be known. We applied this test as a pre-selection, because the following t-test is designed for data with normal distributions. So only for those labels that didn't show a significant deviation from a normal distribution of $\alpha$, we applied the generalized t-test (eq. 5) as a second step. The Null-hypothesis here was that the true mean values of the label's $\alpha$-distribution and the global $\alpha$-distribution of the entire collection were identical giving us the following inequation:

$$-2.58 < \frac{m_{global} - m_{label}}{\sqrt{\frac{\sigma^2_{global}}{n_{global}} + \frac{\sigma^2_{label}}{n_{label}}}} < 2.58 \qquad (5)$$

with $m$ referring to the sample mean, $\sigma^2$ referring to the variance, and $n$ referring to the number of tracks. A deviation is significant on the 1% level when the inequation doesn't hold. Significant deviations from the global mean indicate a correlation between music with a certain semantic label and the typical level of $\alpha$ for this music.

Secondly, we looked at the distribution of labels in different partitions, where each partition contained all tracks within a certain range of $\alpha$ values. The collection was partitioned into deciles with 775 tracks each. We then tested with the $\chi^2$-test whether the percentages of the individual labels in each decile differed significantly from an equal distribution, which was assumed as the Null-hypothesis. Here, we used a significance level of 1%. Deviations from the equal distribution indicate a correlation between levels of $\alpha$ and the frequency of appearance of a certain semantic label assigned to the music.

#### 3.2.2. Machine learning methods

Since the labels are originally assigned to artists rather than to individual tracks, we also estimated the recognition rate for simplified artist identification tasks with machine learning algorithms. A major conceptual difference to the statistical approach lies in the fact that an artist usually shares a combination of labels which are not necessarily synonymic. It is therefore not useful to try arbitrary artist recognition using $\alpha$ as the single feature. Instead, we selected artists where a tendency to a certain level of $\alpha$ could be expected. A recognition rate clearly above chance level indicates a correlation between a certain artist's typical music style (or at least some aspect of it) and the DFA exponent.

For our experiments the feature by itself consists of only one single value per track. Hence, the possibilities of applying machine learning methods are limited. We therefore tested only a rule-based learner and the "lazy learner" method k-nearest-neighbor with different values for k. A 10-fold cross validation was used in all experiments.

**Two-class decision**

In a two class decision experiment we used 238 Frank Sinatra tracks and 238 tracks from nine other artists who either had the labels "Brittle" or "Outrageous" assigned to them. For the artist "Sinatra" a total of 18 labels are listed in our data set, among them "Romantic", "Calm/Peaceful", and "Sentimental". From the results of the statistical analysis (cp. 3.3.1.) we would expect the Sinatra songs to be distributed around a greater value of the DFA exponent than the other ones. The classes should therefore be separable up to a certain degree. It must be noted, that among the nine selected artists we find also assigned labels like "Whistful" and "Reflective", which are linked to higher $\alpha$ values (cp. table 3.1).

**Three-class decision**

In another experiment we used three different classes: 108 tracks composed by Henry Mancini, 65 tracks composed by Bernard Herrmann, and 52 tracks of dance music with a strong beat. We purposely didn't use the labels for selecting the artists in this case. Mancini and Herrmann were both film music composers, but while Herrmann mainly uses "classical" orchestration, Mancini often arranges his music in a jazz-like style. We had to select dance music from different artists, because there was no single one with a sufficient number of tracks in the collection. In terms of the DFA exponent we would expect to find the highest values associated with Herrmann's music, because it is the least danceable in general terms. Intermediate values can be expected for Mancini's tracks, since there are many which at least possess a pronounced beat. The lowest values should be found for the dance music, which has strong and regular beat patterns.

### 3.3. Results

By the manual random checks we found the danceability estimations at the extreme ends being the most consistent ones. Comparing the tracks from these regions with each other and with the intermediate ones the underlying concept of the descriptor immediately became apparent (cp. also figure 2.3). The fine grain ranking within a local region however didn't appear comprehensible in many cases. This was especially noticeable in the very dense area of intermediate values.

#### 3.3.1. Statistical results

The results of the statistical tests sustain the findings from manual random evaluation. Strong coherence of high statistical significance was found for several labels that are semantically close to the concept "danceable" or "not danceable" respectively. For example the labels "Party/Celebratory" and "Energetic" in the context of music have a clear relation with danceability, whereas "Plaintive" and "Reflective" appear more appropriate descriptions for music that is not well suited for dancing. But, even exceeding the aspect of danceability, the results reveal a consistency on a high abstraction level. In some cases we find labels that are quasi antonyms for opposite deviations. We report in more detail about the findings in the following subsections.

| Label | $\overline{\alpha}$ | $n$ | Label | $\overline{\alpha}$ | $n$ |
|---|---|---|---|---|---|
| Party/ Celebratory | 0.796 | 706 | Romantic | 0.911 | 1399 |
| Clinical | 0.761 | 489 | Wistful | 0.909 | 1308 |
| Hypnotic | 0.804 | 951 | Plaintive | 0.922 | 659 |
| Energetic | 0.820 | 898 | Reflective | 0.901 | 1805 |
| Visceral | 0.808 | 422 | Calm/ Peaceful | 0.908 | 1102 |
| Trippy | 0.824 | 998 | Autumnal | 0.916 | 604 |
| Outrageous | 0.781 | 102 | Intimate | 0.897 | 1709 |
| Exuberant | 0.839 | 1383 | Stately | 0.908 | 730 |
| Irreverent | 0.830 | 657 | Gentle | 0.892 | 1327 |
| Sparkling | 0.790 | 116 | Elegant | 0.886 | 2506 |

Table 3.1: Most significantly differing labels in each direction from the global mean of 0.863 ($n$=number of tracks).

**The generalized t-test**

For 139 labels the hypothesis of a normal distribution could not be rejected on a 5% confidence level. Of these, 24 showed a significantly higher and 35 a significantly lower mean value than the global mean of 0.863. The ten most significant ones for each direction are listed in table 3.1.

When looking at the two lists of labels a certain affinity can be noted in many cases on either side. The group of labels for higher $\alpha$ wakes associations of *Softness*, *Warmness*, *Tranquility*, and *Melancholy*. For the others we might form two subgroups, one around terms like *Exuberance* and *Vehemence*, the other around *Tedium* and *Coldness*. Comparing labels from both lists with each other, we can identify several almost antonymous pairs, for example: Outrageous – Refined/Mannered, Boisterous – Calm/Peacefull, Carefree – Melancholic, Clinical – Intimate.

**The $\chi^2$-test**

As mentioned above we split the entire collection into deciles. The boundaries for $\alpha$ are the following: 0.2491, 0.6718, 0.7366, 0.7846, 0.8214, 0.8577, 0.8938, 0.9339, 0.9799, 1.0343, 1.4270. Evaluating the proportions of tracks within these ten partitions for each individual label we found a significant deviation from a hypothesized equal distribution beyond the 1% level in 68 cases. In 56 cases even the 0.1% level was exceeded.

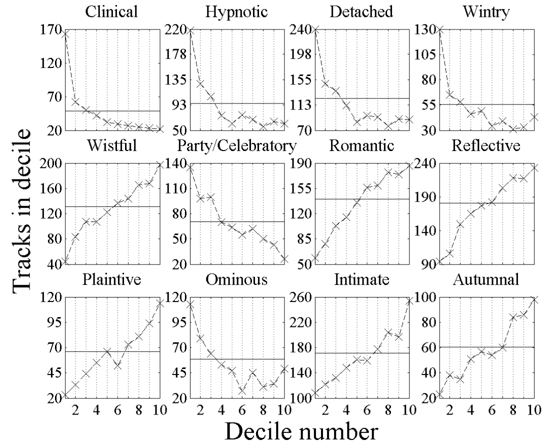


Figure 3.2: Distributions on deciles for the twelve labels with most significant deviation from equal distribution (solid lines).

Of the 68 labels only 39 coincide with the ones found significant in the generalized t-test, 23 for the lower $\alpha$ values and 16 for the higher ones.

Figure 3.2 shows the distributions on the deciles for the twelve most significant labels. We can see two different, basic types of distributions here. The first row, as well as the "Ominous" tracks, show a high concentration in the first one or two deciles rapidly decaying to a lower, relatively stable level for the following ones. For the other labels, the percentage of tracks is almost linearly growing (or decaying) over the full range of $\alpha$.

### 3.3.2. Machine learning results

Both machine learning experiments were revealing the expected range allocation when the rule based method was used. The results are far from an optimal class separation, but clearly indicate the relevance of the used feature. We discuss the outcome of the two experiments in the following sections in some more detail.

**Two-class experiment**

Since the two classes had exactly the same amount of tracks in this experiment we had a baseline of 50% for the classification. With the decision table classifier an optimal threshold for $\alpha$ of 0.914 was found. As expected, the Sinatra tracks were associated with the higher $\alpha$ values and the "Brittle"/"Outrageous" tracks with the lower ones. This way, a level of 72.27% of correctly classified items was achieved, which is clearly better than guessing.

| | Predicted class: | |
|---|---|---|
| **True class:** | **Sinatra** | **B/O** |
| **Sinatra** | 174 | 64 |
| **B/O** | 63 | 175 |

Table 3.2: Confusion matrix for knn classifier with k=20.

The knn-classifier needed to consider a relatively large number of neighbors in order to reach the same accuracy. For k=20 we achieved 73.32%, while the basic k=1 only yielded 64.29% of correct classifications. The confusion matrix for k=20 is shown in table 3.2. It is almost symmetric.

**Three-class experiment**

For the three class experiment the best guessing strategy would have been to always classify as “Mancini”, because it was the class with the most instances. The baseline is therefore 48.0% in this experiment.

Again, the decision table classifier found the anticipated order of the classes assigning a threshold of 0.7 to discriminate the dance music from Mancini’s music and another one of 1.0 to discriminate Mancini’s from Herrmann’s music. 61.78% of correctly classified instances where achieved this way. The figure is less impressive than the one from the previous experiment, but the task was also more demanding.

| | Predicted class: | | |
|---|---|---|---|
| **True class:** | **Mancini** | **Herrmann** | **dance** |
| **Mancini** | 84 | 18 | 6 |
| **Herrmann** | 23 | 41 | 1 |
| **dance** | 25 | 1 | 26 |

Table 3.3: Confusion matrix for knn classifier with k=28.

This time, the knn-classifier needed an even bigger neighborhood for optimal results. A basic k=1 neighborhood only yielded 55.56% of accuracy. But with a recognition rate of 67.11% for k=28 the classification by fixed thresholds was outperformed in this task. Table 3.3 shows the confusion matrix, where we see that Herrmann’s music was almost never confused with the dance music and vice versa. The “Mancini” instances in contrast overlap quite a lot with the other two classes.

## 4. CONCLUSION

From our experiments and observations the hypothesis put forward by Jennings et al. seems to be held, and the DFA exponent can be considered a good indicator for the danceability of music. However, this should be seen rather in the broad sense, classifying music into a small number of categories from “extremely easy” over “moderately easy” to “very difficult” to dance to. Currently, a fine grain ordering by the DFA exponent inside such a category is not beneficial. Due to subjectivity effects such an ordering might not prove useful anyway. By averaging the DFA exponent function we used an extremely compact and simple representation in the experiments. It should be possible to improve the results by a more sophisticated reduction of the function $\alpha(i)$ or by choosing a less compact representation (see also section 5).

It can further be concluded from our results, that the DFA exponent shows to be meaningful also in revealing even higher level attributes of music. It thus might form a valuable addition for different music classification tasks, like artist identification or musical genre recognition.

## 5. FUTURE WORK

As pointed out above, we see some more potential of the DFA in music retrieval application, not limited to a danceability estimation. Our experiments were done with a very basic setup leaving many ways for future enhancement of the processing. An obvious point to improve is the averaging of the function $\alpha(i)$. Instead of computing only the mean it might be beneficial to take also the variance into account. Even the computation of other features from this function, like the identification of peaks and valleys could prove useful. Furthermore, the selection of $\tau$ can be optimized considering both, performance and computational efficiency. Probably the upper bound can be reduced in favor of a finer resolution. Concerning the computational cost it is also interesting to explore some more the alternative approach proposed by Willson and Francis [6].

## 6. ACKNOWLEDGEMENTS

This research has been partially funded by the EU-FP6-IST-507142 project SIMAC (Semantic Interaction with Music Audio Contents). More information can be found at the project website http://www.semanticaudio.org.

The authors also want to thank their colleagues Pedro Cano, Markus Koppenberger, and Oscar Celma from Universitat Pompeu Fabra for their remarkable contributions to the experiments.


## 7. REFERENCES


[1] C.-K. Peng et al., "Mosaic organization of DNA nucleotides", *Phys. Rev. E* 49, 1685—1689 (1994).

[2] H. D. Jennings et al., "Variance fluctuations in nonstationary time series: a comparative study of music genres," *Condensed Matter* (2003). http://xxx.lanl.gov/abs/cond-mat/0312380

[3] C.-K. Peng et al., "Fractal Mechanisms in Neural Control: Human Heartbeat and Gait Dynamics in Health and Disease", Online Tutorial. http://www.physionet.org/tutorials/fmnc/

[4] S. V. Buldyrev et al., "Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis", *Phys. Rev. E* 51, 5084—5091 (1995).

[5] S. Streich, P. Herrera, "Towards describing perceived complexity of songs: computational methods and implementation", *Proc. of the 25th Int. AES Conf.* (2004).

[6] K. Willson, D. P. Francis, "A direct analytical demonstration of the essential equivalence of detrended fluctuation analysis and spectral analysis of RR interval variability", *Physiological Measurement* 24, N1—N7 (2003).

[7] P. Cano et al., "MTG-DB: A Repository for Music Audio Processing", *Proc. of the 4th Int. WEDEL Music Conf.* (2004).

[8] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown", *Journal of the American Statistical Association* 62, 339—402 (1967).