TRACKING MELODIC PATTERNS IN FLAMENCO SINGING BY ANALYZING POLYPHONIC MUSIC RECORDINGS

A. Pikrakis
University of Piraeus, Greece
pikrakis@unipi.gr

F. Gómez, S. Oramas
Polytechnic University of Madrid, Spain
{fmartin, soramasm}@eui.upm.es

J. M. D. Báñez, J. Mora, F. Escobar
University of Sevilla, Spain
{dbanez, mora, fescobar}@us.es

E. Gómez, J. Salamon
Universitat Pompeu Fabra, Spain
{emilia.gomez, justin.salamon}@upf.edu

ABSTRACT

The purpose of this paper is to present an algorithmic pipeline for melodic pattern detection in audio files. Our method follows a two-stage approach: first, vocal pitch sequences are extracted from the audio recordings by means of a predominant fundamental frequency estimation technique; second, instances of the patterns are detected directly in the pitch sequences by means of a dynamic programming algorithm which is robust to pitch estimation errors. In order to test the proposed method, an analysis of characteristic melodic patterns in the context of the flamenco fandango style was performed. To this end, a number of such patterns were defined in symbolic format by flamenco experts and were later detected in music corpora, which were composed of un-segmented audio recordings taken from two fandango styles, namely Valverde fandangos and Huelva capital fandangos. These two styles are representative of the fandango tradition and also differ with respect to their musical characteristics. Finally, the strategy in the evaluation of the algorithm performance was discussed by flamenco experts and their conclusions are presented in this paper.

1. INTRODUCTION

1.1 Motivation and context

The study of characteristic melodic patterns is relevant to the musical style and this is especially true in the case of oral traditions that exhibit a strong melodic nature. Flamenco music is an oral tradition where voice is an essential element. Hence, melody is a predominant feature and many styles in flamenco music can be characterized in melodic terms. However, in flamenco music the problem of characterizing styles via melodic patterns has so far received very little attention. In this paper, we study characteristic melodic patterns, i.e., melodic patterns that make a given style recognizable.

In general, it is possible to adopt two main approaches to the study of characteristic melodic patterns. According to the first approach, music is analysed to discover characteristic melodic patterns [2] (distinctive patterns in the terminology of [2]); see, for example, [3] for a practical application of this approach to finding characteristic patterns in Brahms' string quartet in C minor. Typically, the detected patterns are assessed by musicologists to determine how meaningful they are. Therefore, this type of approach is essentially an inductive method. The second approach is in a certain sense complementary to the first one: specific melodic patterns, which are known or are hypothesized to be characteristic, are tracked in the music stream. The results of this type of method allow musicologists to study important aspects of the given musical style, e.g., to confirm existing musical hypotheses. The techniques to carry out such tracking operations vary greatly depending on the application context, the adopted music representation (symbolic or audio), the musical style and the available corpora. This type of approach can be termed as deductive.

In this paper, we adopted the second approach. Specifically, certain characteristic melodic patterns were carefully selected by a group of flamenco experts and were searched in a corpus of flamenco songs that belong to the style of fandango. Tracking patterns in flamenco music is a challenging task for a number of reasons. First of all, flamenco music is usually only available as raw audio recordings without any accompanying metadata. Secondly, flamenco music uses intervals smaller than a half-tone and is not strict with tuning. Furthermore, due to improvisation, a given abstract melodic pattern can be sung in many different ways, sometimes undergoing dramatic transformations, and still be considered the same pattern within the flamenco style. These facts obviously increase the complexity of the melody search operation and demand for increased robustness.

Preliminary work on detecting ornamentation in flamenco music was carried out in [6], where a number of predefined ornaments were adapted from classical music and were looked up in a flamenco corpus of *tonás* styles. In [9] a melodic study of flamenco a cappella singing styles was performed.

1.2 Goals

Two main goals were established for this work: the first one was of technical nature -transcription of music and location of melodic patterns-, and the second one of musicological nature -the study of certain characteristic patterns of the Valverde fandango style.

From an algorithmic perspective, two major problems had to be addressed. The first problem was related to the transcription of music, since flamenco is an oral music tradition and transcriptions are meagre. In addition, our corpus consisted of audio recordings that contained both guitar and voice and predominant melody (pitch) estimation was applied in order to extract the singing voice. The output of this processing stage was a set of pitch contours representing the vocal lines in the recordings. Note that even though we use a state-of-the-art algorithm, theses lines will still contain estimation errors, and our algorithm must be able to cope with them. The second problem was related to the fact that the patterns to be detected were specified by flamenco experts in an abstract (symbolic) way and we had to locate the characteristic patterns directly on the extracted pitch sequences. To this end, we developed a tracking algorithm that operates on a by-example basis and extends the context-dependent dynamic time warping scheme [10], which was originally proposed for pre-segmented data in the context of wind instruments.

Musicologically speaking, the goal was to examine certain melodic patterns as to being characteristic of the Valverde fandango style. Those patterns were specified in a symbolic, abstract way and were detected in the corpus. Both the pattern itself and its location were important from a musicological point of view. The tracking results were reviewed and assessed by a number of flamenco experts. The assessment was carried out with respect to a varying similarity threshold that served as means to filter the results returned by the algorithm. In general, the subjective evaluation of the results (experts' opinion) was consistent with the algorithmic output.

2. THE FANDANGO STYLE

Fandango is one of the most fundamental styles in flamenco music. In Andalusia, there are two main regions where fandango has marked musical characteristics: Malaga (verdiales fandangos) and Huelva (Huelva fandangos).

Verdiales fandangos are traditional folk *cantes* related to dance and a particular sort of gathering. The singing style is melismatic and flowing at the same time [1].

Huelva fandangos are usually sung in accompaniment with a guitar. The oldest references about Huelva fandangos date back to the second half of the XIX century. At present, Huelva fandangos are the most popular ones and display a great number of variants. They can be classified based on the following criteria: (1) Geographical origin: from the mountains (Encinasola), from Andévalo (Alosno), from the capital (Huelva capital fandango); (2) Tempo: fast (Calañas), medium (Santa Barbara), or slow (valientes from Alosno); (3) Origin of tradition: village (Valverde), or personal, i.e., fandangos that are attributed to important singers (Rebollo and other important singers, for example). More information on the different styles of fandango can be found in [7].

From a musicological perspective, all fandangos have a common formal and harmonic structure which is composed of an instrumental refrain in flamenco mode (major Phrygian) and a sung verse or *copla* in major mode. The interpretation of fandangos can be closer to the folkloric style, or to the flamenco style, with predominant melismas and greater freedom in terms of rhythm. The reader may refer to [5] for further information on their musical description.

The study of the fandangos of Huelva is of particular interest for the following reasons: (1) Identification of the musical processes that contribute to the evolution of folk styles to flamenco styles; (2) Definition of styles according to their melodic similarity; (3) Identification of the musical variables that define each style; this includes the discovery of melodic and harmonic patterns.

3. THE CHARACTERISTIC PATTERNS OF FANDANGO STYLES

Patterns heard in the exposition (the initial presentation of the thematic material) are fundamental to recognizing fandango styles. The main patterns identified in the Valverde fandango style are shown in Figure 1 (chords shown in Figure 1 are played by the guitar; pitches are notated as intervals from the root) . These patterns are named as follows: *exp-1*, *exp-2*, *exp-4*, and *exp-6*. The number in the name of the pattern refers to the phrase in which it occurs in the piece.

Pattern *exp-1* is composed of a turn-like figure around the tonic. Pattern *exp-2* basically goes up by a perfect fifth. First, the melody insists on the B flat, makes a minor-second mordent-like movement, and then rises with a leap of a perfect fourth. Pattern *exp-4* is a fall from the tonic to the fourth degree by conjunct degrees followed by an ascending leap of a fourth. Pattern *exp-6* is a movement from B flat to the tonic. Again, the B flat is repeated, then it goes down by a half-tone and raises to the tonic with an ascending minor third. The rhythmic grouping of the melodic cell is ternary (three eighth notes for B flat and three eighth notes for A).

Again, notice that this is a symbolic description of the actual patterns heard in the audio files. Any of these patterns may undergo substantial changes in terms of duration, sometimes even in pitch, not to mention timbre and other expressive features.

3.1 The Corpus of Fandango

The corpus of our study was provided by *Centro andaluz de flamenco de la Junta de Andalucía*, an official institution whose mission is the preservation of the cultural heritage

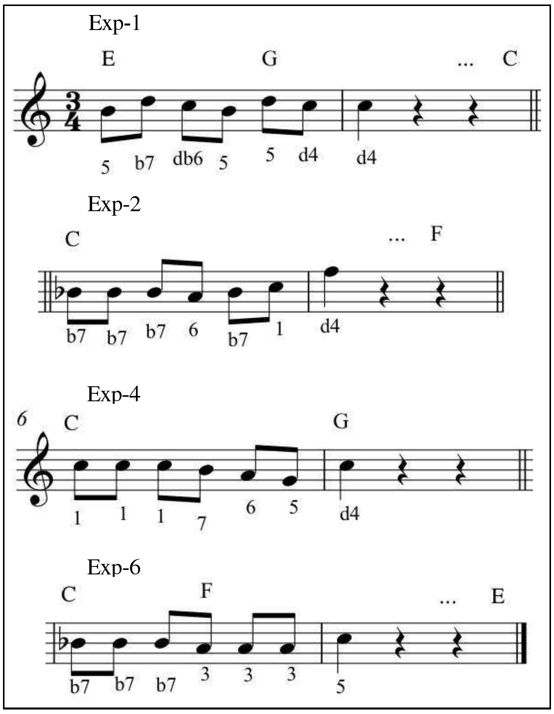

Figure 1. Characteristic patterns in the Valverde fandango style.

of flamenco music. This institution possesses around 1200 fandangos, from which 241 were selected. The selection was based on the following four criteria: (1) Audio files must contain guitar and voice; (2) Audio files are of acceptable recording quality to permit automatic processing; (3) Fandangos must be interpreted by singers from Huelva or acknowledged singing masters; (4) The time span of the recordings must be broad and in our case it covers six decades, from 1950 to 2009.

The corpus was gathered for the purposes of a larger project that aims at investigating fandango in depth. The sample under study is broadly representative of styles and tendencies over time. The current paper is an attempt to study 60 fandangos in total (30 Valverde fandangos and 30 Huelva capital fandangos). In this experimental setup we excluded Valientes of Huelva fandangos, Valientes de Alosno fandangos, Calañas fandangos, and Almonaster fandangos. All recordings were available in PCM (wav) single-channel format, with a 16 bit-depth per sample and 44 kHz sampling rate.

4. COMPUTATIONAL METHOD

4.1 Audio Feature Extraction

As mentioned earlier, written scores in flamenco music are scattered and scant. This can be explained to some extent by the fact that flamenco music is based on oral transmission. Issues related to the most appropriate transcription method have been quite controversial in the context of the flamenco community. Some authors, like Hurtado and Hurtado [8], are in favour of Western notation, whereas others propose different methods, e.g., Donnier [4], who advocates the use of plainchant neumes. In view of this controversy, we adopted a more technical approach that is based on audio feature extraction.

We now describe how the audio feature extraction algorithm operates. Our goal was to extract the vocal line in an appropriate, musically meaningful format that would also serve as input to the pattern detection algorithm. The audio feature extraction stage was mainly based on predominant melody (fundamental frequency, from now on $F0$) estimation from polyphonic signals. For this, we used the state-of-the art algorithm proposed by Salamon and Gómez [11]. Their algorithm is composed of four blocks. First, they extract spectral peaks from the signal by taking the local maxima of the short-time Fourier transform. Next, those peaks are used to compute a salience function representing pitch salience over time. Then, peaks of the salience function are grouped over time to form pitch contours. Finally, the characteristics of the pitch contours are used to filter out non-melodic contours, and the melody $F0$ sequence is selected from the remaining contours by taking the frequency of the most salient contour in each frame. Further details can be found in [11].

4.2 Pattern Recognition Method

The pattern detection method used in this paper builds upon the "Context-Dependent Dynamic Time Warping" algorithm (CDDTW) [10]. While standard dynamic time warping schemes assume that each feature in the feature sequence is uncorrelated with its neighboring ones (i.e. its context), CDDTW allows for grouping neighboring features (i.e. forming feature segments) in order to exploit possible underlying mutual dependence. This can be useful in the case of noisy pitch sequences, because it permits canceling out several types of pitch estimation errors, including pitch halving or doubling errors and intervals that are broken into a sequence of subintervals. Furthermore, in the case of melismatic music, the CDDTW algorithm is capable of smoothing variations due to the improvisational style of singers or instrument players. For a more detailed study of the CDDTW algorithm, the reader is referred to [10].

A drawback of CDDTW is that does not take into account the duration of music notes and focuses exclusively on pitch intervals. Furthermore, CDDTW was originally proposed for isolated musical patterns (pre-segmented data). The term isolated refers to the fact that the pattern that is matched against a prototype has been previously extracted from its context by means of an appropriate segmentation procedure, which can be a limitation in some real-world scenarios, like the one we are studying in this paper. Therefore, we propose here an extension to the CDDTW algorithm, that:

- Removes the need to segment the data prior to the application of the matching algorithm. This means that the prototype (in our case the time-pitch representation of the MIDI pattern) is detected directly in the pitch sequence of the audio stream without prior segmentation, i.e. the pitch sequence that was extracted from the fandango.

- Takes into account the note durations in the formulation of the local similarity measure.
- Permits to search for a pattern iteratively, which means that multiple instances of the pattern can be detected, one per iteration.

A detailed description of the extension of the algorithm is beyond the scope of this paper. Instead, we present the basic steps:

Step 1: The MIDI pattern to be detected is first converted to a time-pitch representation

$$P = \{[f_1, t_1]^T, [f_2, t_2]^T, \ldots, [f_J, t_J]^T\},$$

where f_i is the frequency of the i-th MIDI note, measured in cents (assuming that the reference frequency is 55 Hz) and t_i is the respective note duration (in seconds), for a MIDI pattern of J notes.

Step 2: Similarly, the pitch sequence of the audio recording is converted to the above time-pitch representation,

$$R = \{[r_1, tr_1]^T, [r_2, tr_2]^T, \ldots, [r_I, tr_I]^T\},$$

where r_i is a pitch value (in cents) and tr_i is always equal to the short-term step of the feature extraction stage ($10ms$ in our case), for an audio recording of I notes. In other words, even if two successive pitch values are equal, they are still treated as two successive events, each of which has a duration equal to the short-term step of the feature extraction stage. This approach was adopted to increase the flexibility of the dynamic time warping technique at the expense of increased computational complexity. For the sake of uniformity of representation, each time interval that corresponds to a pause or to a non-vocal part is inserted as a zero-frequency note and is assigned a respective time duration.

Step 3: Sequences R and P are placed on the vertical and horizontal axis of a similarity grid, respectively. The CDDTW algorithm is then applied on this grid, but, this time, the cost to reach node (i, j) from an allowable predecessor, say $(i-k, j-1)$, depends both on the pitch intervals and the respective note durations. More specifically, the interpretation of the transition $(i-k, j-1) \to (i, j)$ is that the pitch intervals in the MIDI pattern and audio recording are equal to $f_j - f_{j-1}$ and $r_i - r_{k-1}$, respectively. Note that on the y-axis, the pitch interval only depends on the end nodes of the transition and not on any intermediate pitch values, hence the ability to cancel out any intermediate pitch tracking phenomena. In the same spirit, the time duration that has elapsed on the $x-$axis and $y-$axis is equal to t_j and $\sum_{i-k+1}^{i} tr_k$, respectively. It is worth noticing that we do not permit omitting notes from the MIDI pattern, and therefore any allowable predecessor of (i, j) must reside in column $j-1$. The pitch intervals and respective durations are fed to the similarity function of Eq. (1), that yields a score, $S_{(i-k,j-1)\to(i,j)}$, for the transition $(i-k, j-1) \to (i, j)$, i.e.,

$$\begin{aligned} S_{(i-k,j-1)\to(i,j)} = 1 - f(\frac{\sum_{i-k+1}^{i} tr_k}{t_j}) \\ -g(r_i - r_{k-1}, f_j - f_{j-1}) \end{aligned} \tag{1}$$

where

$$f(x) = \begin{cases} (1-x)^{1.1}, & 1 \le x \le 4 \\ 1.5^{1.1}(1-x)^{1.1}, & \frac{1}{3} \le x < 1 \\ 3 - 6x, & 0 < x < \frac{1}{3} \\ \infty, & \text{otherwise} \end{cases}$$

and

$$g(x_1, x_2) = \begin{cases} (1 - \frac{x_1}{x_2})^{0.7}, & \text{if } 0.98 \le \frac{x_1}{x_2} \le 1.02 \\ \infty, & \text{otherwise} \end{cases}$$

The interpretation of this function is that it penalizes excessive time warping and does not tolerate much deviation in terms of pitch intervals. More specifically, $f(x)$ is a piecewise function that operates on the basis that duration ratios are not penalized uniformly and that any ratio outside the interval $[\frac{1}{3}, 1)$ should receive a stronger penalty. Similarly, function $g(x)$ implies that, taking the music interval of the MIDI pattern as reference, the respective sum of intervals of the audio recording exhibits at most a 2% deviation. The scalars involved in the formulae of $f(x)$ and $g(x)$ are the result of fine-tuning with respect to the corpus under study. The computation of the transition cost is repeated for every allowable predecessor of (i, j). In the end, one of the predecessors is selected to be the winner by examining the sum of the similarity that has been generated by the transition with the accumulated similarity at the predecessor.

Step 4: After the accumulated cost has been computed for all nodes in the grid, the maximum accumulated cost is selected and normalized and, if it exceeds a predefined threshold, a standard backtracking procedure reveals which part of the audio recording has been matched with the prototype; otherwise, the algorithm terminates.

Step 5: All nodes in the best path are marked as stop-nodes, i.e. forbidden nodes and Steps 1-4 are repeated in order to detect a second occurrence of the prototype and so on, depending on how many patterns (at most) the user has requested to be detected.

5. EVALUATION

5.1 Methodology

Four different exposition patterns were defined by the experts, which are distinctive of the Valverde style. The Valverde fandango has 6 exposition phrases in each *copla* (sung verse), where 1, 3 and 5 are usually the same pattern, and 2, 4 and 6 have different patterns each. Therefore, 4 exposition patterns (1, 2, 4, and 6) were chosen to be put to the test. Again, we insist that these patterns are abstract representations of the actual patterns heard in the audio recordings. Our algorithm was then run to locate those four patterns in the corpus of Valverde fandangos and Huelva capital fandangos. Therefore, our ground-truth in this study consists of all the melodic patterns plus their specific locations. For example, exposition pattern 1 has to be located 90 times, as it occurs three times in each of the 30 pieces that make up the corpus of the Valverde fandangos. If this pattern is found elsewhere (not in the exposition phrase), then it will be considered as a true negative. Once the results of the experiments were obtained, they were manually checked by the flamenco experts, both in terms of pattern occurrence and respective position.

5.2 Results

Results are summarized in Tables 1 and 2 with respect to the similarity threshold, which is a user-controlled variable. Once the threshold is set to a specific value, the algorithm filters out any patterns whose similarity score does not exceed the threshold. In our study, we experimented with values of the similarity threshold ranging from 30% to 80%. In Table 1, T_e stands for the total number of expected occurrences of each pattern in the corpus of Valverde fandangos (based on the ground truth that is provided by the musicological knowledge), T_f is the total number of detected instances (both true and false), T_p is the number of true positives, F_p is the number of false positives, and *Prec.*, *Rec.* and F are the values of precision, recall and the F-measure, respectively. In Table 2, we focus on the corpus of Huelva fandangos.

Figure 2 shows the average F-measure (over all patterns) as a function of similarity threshold. The maximum value is obtained at threshold 50%.

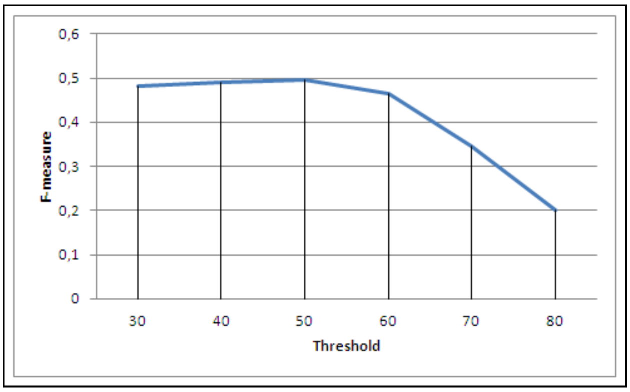

Figure 2. Average F-measure (over all patterns) with respect to the similarity threshold.

Next, we attempt to detect the Valverde patterns in the Huelva collection. Hence, one would expect that it would be otiose to reproduce computations like those in Table 1, as the total expected number of occurrences would be zero. However, Table 2 summarizes the detection results in the corpus of the Huelva capital fandangos for the four exposition patterns under study and we make an attempt to provide an interpretation of the detected occurrences.

Valverde fandangos								
	Sim.	T_e	T_f	T_p	F_p	Prec.	Rec.	F
Exp-1	30%	90	38	32	6	84%	36%	0.5
	40%	90	36	32	4	89%	36%	0.5
	50%	90	31	30	1	97%	33%	0.49
	60%	90	25	24	1	96%	27%	0.41
	70%	90	15	15	0	100%	17%	0.28
	80%	90	6	6	0	100%	7%	0.12
Exp-2	30%	30	13	13	0	100%	43%	0.6
	40%	30	13	13	0	100%	43%	0.6
	50%	30	13	13	0	100%	43%	0.6
	60%	30	13	13	0	100%	43%	0.6
	70%	30	7	7	0	100%	23%	0.37
	80%	30	1	1	0	100%	3%	0.06
Exp-4	30%	30	27	11	16	41%	37%	0.38
	40%	30	26	11	15	42%	37%	0.39
	50%	30	21	11	10	52%	37%	0.43
	60%	30	16	9	7	56%	30%	0.39
	70%	30	11	6	5	54.5%	20%	0.29
	80%	30	3	3	0	100%	10%	0.18
Exp-6	30%	30	34	14	20	41%	47%	0.43
	40%	30	31	14	17	45%	47%	0.46
	50%	30	27	13	14	48%	43%	0.45
	60%	30	15	10	5	66.6%	33%	0.44
	70%	30	8	8	0	100%	27%	0.42
	80%	30	3	3	0	100%	10%	0.18

Table 1. Experimental results for Valverde fandangos.

Huelva capital fandangos						
	30%	40%	50%	60%	70%	80%
Exp-1	7	4	1	1	0	0
Exp-2	1	0	0	0	0	0
Exp-4	29	27	23	17	8	4
Exp-6	27	22	19	17	9	5

Table 2. Experimental results for the Huelva capital fandangos.

Overall, from a quantitative point of view, the algorithm has exhibited a reasonably good performance in finding the patterns in the melody, despite the problems posed by the polyphonic source, the highly melismatic content, and the note-duration variation. Regarding performance measures, on the one hand, precision is quite high, but, on the other hand, recall is low. Most of the values of F-measure are around $0.3 - 0.45$, with a few isolated exceptions. In other words, the algorithm is capable of detecting well localized occurrences of the patterns, but fails to locate a significant number of occurrences. The best performance of the F-measure occurs with a threshold of 50%.

From a qualitative point of view, we make the following remarks.

Exp-1: This pattern is the exposition of the first phrase of the fandango. Interestingly enough, not only does the algorithm detect the pattern correctly in the first phrase of the Valverde fandango, but also in other phrases, as expected. Indeed, it identifies the pattern as a leit-motiv throughout the piece. This pattern was detected only a few times by the algorithm in the Huelva capital fandangos.

Exp-2: This is the pattern of the second exposition phrase in Valverde fandangos. This is the musical passage with the amplest tessitura. The algorithm detects it with high precision in the Valverde corpus (even for a similarity threshold equal to 30%), and very few matches are encountered in the Huelva capital fandangos.

Exp-4: In the Valverde corpus, for a threshold equal to 80%, the algorithm only detects the pattern in *cantes* sung by women who have received music training in flamenco clubs in Huelva. These clubs are called *peñas flamencas* and organize singing lessons. Women from *peñas* are trained to follow very standard models of singing and therefore do not contribute to music innovation like other fandango performers e.g., Toronjo or Rengel). For a 70% similarity threshold (and below), the pattern is also detected in the voices of well-known fandango singers.

In the Huelva capital fandango corpus this pattern is frequently detected by the algorithm in the transition between phrases. Note that we can state that the pattern is there, more or less blurred or stretched, but it is present, so these are not considered to be false positives.

Exp-6: This pattern is used to prepare the final cadence of the last phrase. In the Valverde corpus, irrespective of the similarity level, the algorithm returns correct results, although as stated above, many occurrences fail to be detected. In the Huelva capital corpus and when the threshold is low, the algorithm detects the pattern in the first, the middle and the final section. When the threshold is raised to 80%, it is only located in the final cadence.

6. CONCLUSIONS

In this paper we presented an algorithmic pipeline to perform melodic pattern detection in audio files. The overall performance of our method depends both on the quality of the extracted melody and the precision of the tracking algorithm. In general, the system's performance, in terms of precision and recall of detected patterns, was measured to be satisfactory, despite the great amount of melismas and the high tempo deviation. From a musicological perspective, we carried out a study of fandango styles by means of analyzing archetypal melodic patterns. As already mentioned, written scores are not in general available for flamenco music. Therefore, our approach was to design a system that operated directly on raw audio recordings and circumvented the need for a transcription stage. In the future, our study could be extended to other Huelva fandango styles. A more ambitious goal would be to carry out the analysis for the whole corpus of fandango music. Also, other musical features could be taken into account and thus perform a more general analysis, i.e., embrace more than what melodic descriptors can offer.

7. ACKNOWLEDGEMENTS

The authors would like to thank the *Centro Andaluz del Flamenco, Junta de Andalucía* [1] for providing the music collection. This work has been partially funded by AGAUR (mobility grant), the COFLA project [2] (P09 - TIC - 4840 Proyecto de Excelencia, Junta de Andalucía) and the Programa de Formación del Profesorado Universitario of the Ministerio de Educación de España.

8. REFERENCES

[1] M. A. Berlanga. *Bailes de candil andaluces y fiesta de verdiales. Otra visión de los fandangos*. Colección monografías. Diputación de Málaga, 2000.

[2] D. Conklin. Discovery of distinctive patterns in music. *Intelligent Data Analysis*, 14(5):547–554, 2010.

[3] D. Conklin. Distinctive patterns in the first movement of Brahms' string quartet in C minor. *Journal of Mathematics and Music*, 4(2):85–92, 2010.

[4] P. Donnier. Flamenco: elementos para la transcripción del cante y la guitarra. In *Proceedings of the IIIrd Congress of the Spanish Ethnomusicology Society*, 1997.

[5] Lola Fernández. *Flamenco Music Theory*. Acordes Concert, Madrid, Spain, 2004.

[6] F. Gómez, A. Pikrakis, J. Mora, J.M. Díaz-Báñez, E. Gómez, and F. Escobar. Automatic detection of ornamentation in flamenco. In *Fourth International Workshop on Machine Learning and Music MML*, NIPS Conference, December 2011.

[7] M. Gómez(director). Rito y geografía del cante flamenco II . Videodisco. Madrid: Círculo Digital, D.L., 2005.

[8] A. Hurtado Torres and D. Hurtado Torres. *La voz de la tierra, estudio y transcripción de los cantes campesinos en las provincias de Jaén y Córdoba*. Centro Andaluz de Flamenco, Jerez, Spain, 2002.

[9] J. Mora, F. Gómez, E. Gómez, F. Escobar Borrego, and J. M Díaz Báñez. Characterization and melodic similarity of a cappella flamenco cantes. In *Proceedings of ISMIR*, pages 9–13, Utrecht School of Music, August 2010.

[10] A. Pikrakis, S. Theodoridis, and D. Kamarotos. Recognition of isolated musical patterns using context dependent dynamic time warping. *IEEE Transactions on Speech and Audio Processing*, 11(3):175–183, 2003.

[11] J. Salamon and E. Gómez. Melody Extraction from Polyphonic Music Signals using Pitch Contour Characteristics. *IEEE Transactions on Audio, Speech and Language Processing*, 20(6):1759-1770, Aug. 2012.

[1] http://www.juntadeandalucia.es/cultura/centroandaluzflamenco/
[2] http://mtg.upf.edu/research/projects/cofla