Journal of New Music Research
2011, Vol. 40, No. 2, pp. 119–125

Routledge
Taylor & Francis Group

Comparative Pattern Analysis of Cretan Folk Songs

Darrell Conklin[1,2] and Christina Anagnostopoulou[3]

[1]Universidad del País Vasco, Spain; [2]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain; [3]University of Athens, Greece

Abstract

This paper reports on data mining of Cretan folk songs for distinctive patterns. A pattern is distinctive if it occurs with higher probability in a corpus as compared to an anticorpus. A small set of transcribed Cretan folk songs was encoded, organized using a knowledge base of classes, and mined using distinctive pattern discovery methods. In this exploratory study several highly distinctive melodic patterns emerge, indicating the ability of distinctive patterns to describe subgroups of folk songs.

1. Introduction

In recent years there has been a renewed interest in folk song analysis partly driven by increasing interests in cultural heritage, and also by advances in music informatics methods. The ability to make predictions, from music content, of song properties such as region, dance type, tune family, instrumentation, modality, and social function is an important part of the management of large corpora. Music data mining methods play a key role in building predictive models for folk song classification.

In addition to predictive methods for folk song classification (Hillewaere, Manderick, & Conklin, 2009), descriptive methods are very important for obtaining an overall view of a corpus and for indicating comprehensible subgroups, distributions, and clusters in the data (Crane & Fiehler, 1970; Toiviainen & Eerola, 2001; Juhász, 2006; Taminau et al., 2009). The study here applies pattern discovery to a collection of Cretan folk songs which have been grouped into song types and regions. The comparative orientation distinguishes this approach from most work on pattern discovery in music, which is usually applied within a single piece of music to detect repetition rather than recurrence across pieces (Lartillot, 2004; Cambouropoulos, 2006). In terms of a broad machine learning approach, it can be viewed as an instance of supervised descriptive rule discovery (Novak, Lavrač, & Webb, 2009) which attempts to describe a subset of each class rather than develop predictive models of entire classes.

The traditional music of a region can often be divided into *dance* and *non-dance* music (Amargiannakis, 1994). In Crete, dance music can contain many different types; more than 100 dance variations exist in the whole island. Four dance types have been considered in this study (see Figure 1): Pentozalis, Maleviziotis, Syrtos, and Sousta. The non-dance songs of Crete can be divided according to the musical style and their function. The most well-known non-dance songs in Crete are Rizitika (solemn slow songs, possibly of Byzantine origin) and Tambahaniotika (city songs). The rest of the non-dance songs can be divided according to the function they are used in: lullabies, lament songs, and wedding songs. Finally, we include the popular Kalanda (new year song), and Erotokritos (epic song).

Although most of these types of songs are found throughout the island, each area has its own characteristic musical style. Different musical styles can be identified in the East and West of Crete (Hnaraki, 2007), although there are also many common songs known to the whole island. Therefore, in stylistic analysis of this music, it is interesting to look not only for patterns that distinguish between song types, but also those that distinguish between geographical regions (see Figure 1).

This paper reports on results obtained with a small collection of folk tunes from several areas across Crete,

Correspondence: Darrell Conklin, Department of Computer Science and Artificial Intelligence, Universidad del País Vasco, San Sebastián, Spain. E-mail: conklin@ikerbasque.org

DOI: 10.1080/09298215.2011.573562

taken from various archives of transcriptions, exploring the hypothesis that there may exist patterns of correlation between geographical origin and type of song on the one hand, and melodic patterns on the other. This study focuses on melodic interval patterns, as this is an important dimension which has not to date been studied systematically in Cretan folk music, with the hypothesis that there exist significant patterns in this parameter.

The pattern analysis method chosen is the MGDP method (Conklin, 2010a, 2010b) which discovers patterns that are distinctive and maximally general in a corpus. This paper discusses the encoding of a collection of pieces into a knowledge base for analysis, describes the application of the MGDP method to the corpus, reports on some interesting patterns found, and finally presents ideas for future work.

2. Method

This section describes the corpus, data preparation, and data mining method employed. The data mining method discovers patterns that are over-represented in a positive set (called the *corpus*) with respect to a background set of pieces (called the *anticorpus*).

2.1 Corpus of Cretan folk songs

In this study, four sources of transcribed Cretan folk songs were used. The first one was the collection of transcriptions done by Samuel Baud-Bovy in 1953–1954 (Baud-Bovy, 2006), published by the Laographic Music Archives of Melpo Merlie in Athens. It contains old songs which were recorded and transcribed in various villages around the island. The second source was by Peristeris (Spuridakis & Peristeris, 1968), as found in the Amargiannakis archives at the Academy of Athens and the University of Athens. This contains songs from all over Greece, including some Cretan songs. The third one was a collection of well-known songs, transcribed for lyra playing (Andreoulakis & Petrakis, 1994). Finally, a few pieces were added from the transcriptions done by Theodossopoulou (2003). The corpus contains well-known Cretan songs, embedded in the local life and culture. Not all of them are purely folk tunes; in certain cases they are composed by a specific composer or performer who imitates the old folk style.

All songs in the present corpus contain a melody which is sung and may be accompanied with a melodic instrument. There are two main melodic instruments in Crete which follow the singer's voice (either in unison, or in a question/answer type of texture): the lyra and the violin. Both these instruments have a similar range and way of playing. In the case of lament songs, and often with lullabies, there is no accompanying instrument.

Songs from printed scores were encoded using the Sibelius and Finale software, and exported to MIDI format for computational analysis. During the encoding of the pieces, a knowledge base was developed to describe the classes and other metadata features of pieces. Four broad categories of song classes were chosen (Figure 1): two basic categories of **type** and **area**, and two broader categories of song **supertype** and **superarea**. In addition to being meaningful groupings of song areas and types, these categories also compensate for the problem of small representation in the basic classes.

The area of the folk song was provided with the original scores, and in the case there is an identified composer, the home of the composer was also considered to determine the area. The area of **syria** was included as there are songs in Syria which are also sung in Crete. Song type classification has been mostly kept as specified by Baud-Bovy (2006). Superarea classification was made by dividing the island into East and West (as locals would also naturally divide it), and supertype classification was according to the classification of Greek folk music specified by Amargiannakis (1994) and Baud-Bovy (2006). In some cases it is hard to detect a song type simply by its score, especially in the case of non-dance songs, where the tempo and rhythmic structures can be more free. In this case, the lyrics might be more indicative of the song type, for example, one can distinguish between wedding songs and lullabies based on their lyrics.

While preparing the corpus and classifying the songs, several interesting points came up for discussion, and decisions had to be taken:

- as mentioned above, there are song types found throughout the island, such as the new year song (**kalanda**). These songs were therefore placed in all classes of **area** (and **superarea**);
- there are also songs which might have been recorded at a specific village (e.g. a small village of Chania), but these songs are sung all through the Western part of the island: these songs were thus added to both **rethymno** and **chania** areas;
- for a few songs, although the song type was clear, there was no indication in the sources regarding its area. Therefore, in our classification, these pieces are not considered part of any class for either **area** or **superarea** mining;
- upon initial inspection of some results, some pieces that were close variations of one another were found. For such variation clusters, only one of the set was retained;
- finally, some songs were noted as simply being **non-dance** songs, without any further song type classification, and from the music and lyrics it was not possible to deduce a more specific description. They were treated as not known for the **type** category.

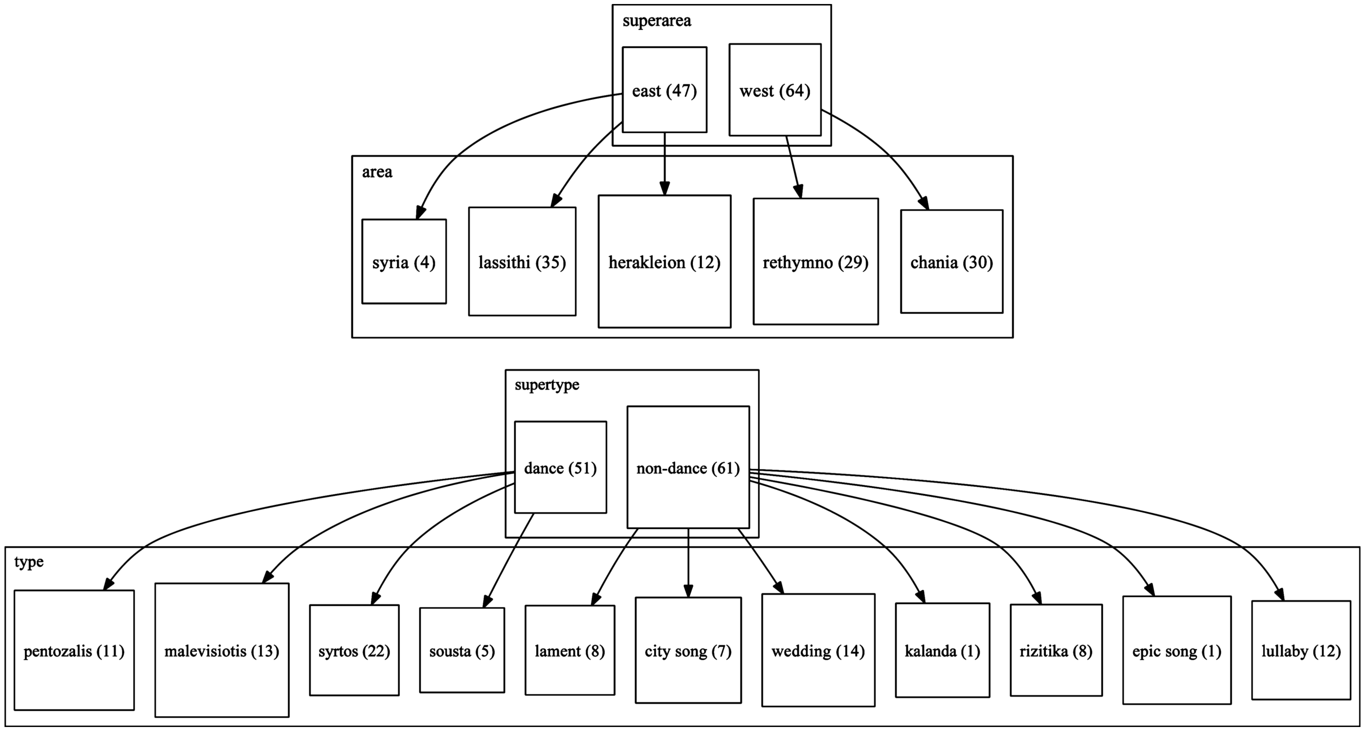

Fig. 1. The categories and classes used in this study, along with the number of pieces in each class.

The final corpus for this study contained 106 songs in our final collection. Figure 1 depicts the categories along with their classes and the number of pieces in each class. Due to the modifications described above, the number of pieces in the subclasses do not always sum to the number of pieces in the superclasses.

2.2 Pattern discovery

This section describes the method used for comparative pattern analysis of the Cretan folk tunes, including concepts of pattern, subsumption, corpus, anticorpus, and pattern distinctiveness. It finally outlines the concept of a maximally general distinctive pattern, and the pattern discovery algorithm.

A *pattern* is a sequence of event features. A rich set of features, including rhythmic ones, can be ascribed to events (Conklin, 2010a), but for this study only melodic intervals (measured in semitone steps) are used due to the small corpus size and the desire to fully explore melodic features prior to addressing rhythmic ones. A piece x *instantiates* a pattern P, written $P(x)$, if the pattern occurs one or more times in the piece: if the components of the pattern are instantiated by successive events in the piece.

A pattern P *subsumes* (is more general than) a pattern Q if all instances of Q are also instances of P: if $\forall x\ Q(x) \Rightarrow P(x)$ is valid (true in all possible corpora). For example, the pattern [+3] subsumes the pattern [+3, +1], which in turn subsumes, for example, [+2, +3, +1] and [+3, +1, −4].

To rank discovered patterns it is possible to partition a collection of pieces into an *anticorpus* (denoted $\ominus$), contrasting it with the analysis corpus (denoted $\oplus$). A *distinctive pattern* is one that is sufficiently over-represented in the corpus as compared to the anticorpus. An intuitive way to measure over-representation is according to the relative empirical probability of a pattern in the corpus and anticorpus:

$$\Delta(P) \stackrel{\text{def}}{=} \frac{p(P \mid \oplus)}{p(P \mid \ominus)}, \tag{1}$$

where

$$p(P \mid \oplus) \stackrel{\text{def}}{=} c^{\oplus}(P)/n^{\oplus},$$

$$p(P \mid \ominus) \stackrel{\text{def}}{=} c^{\ominus}(P)/n^{\ominus},$$

where $c^{\oplus}(P)$ (respectively, $c^{\ominus}(P)$) is the number of pieces in the corpus (anticorpus) containing the pattern P, and $n^{\oplus}$ ($n^{\ominus}$) is the number of pieces in the corpus (anticorpus). If $c^{\ominus}(P) = 0$, $\Delta(P)$ is defined to be ∞.

It is desirable, both for descriptive and for computational reasons, to focus attention primarily on short and frequent distinctive patterns found across several pieces within a corpus. The space of sequential patterns is too large to find or report all distinctive patterns. However, the set of the most general distinctive patterns is much smaller and can be found more efficiently. By definition, these patterns are also the most frequent, and therefore at one extreme of the distinctive set, useful for an analyst to explore prior to searching deeper for more specific patterns. More precisely, a pattern P is called a *maximally general distinctive pattern* (MGDP) if it is distinctive (with at least a specified minimum $\Delta(P)$), and if there does not exist a more general (subsuming) pattern that is also distinctive. Thus the set of all MGDP is the top border of the virtual pattern subsumption taxonomy such that all patterns above the border are not distinctive.

Figure 2 facilitates the explanation of the concepts outlined above. Nodes represent melodic interval patterns, and arrows between nodes represent pattern subsumption (which is a transitive relation, therefore transitive arrows are omitted in the figure). Each pattern is annotated with a hypothetical distinctiveness value $\Delta(P)$. The pattern [+2, +1], for example, is not distinctive, with a $\Delta(P)$ of 1.0 (occurring with equal probability in the corpus and anticorpus). Different MGDP sets result from different minimum values of $\Delta(P)$. For example, with $\Delta(P) \geq 1.4$ the patterns [+2, +1, −3] and [+4] would form the MGDP set. The pattern [+2, +1, +4], though distinctive, would not be included as it is subsumed by the distinctive pattern [+4]. However, with $\Delta(P) \geq 3$, it would indeed be in the MGDP set which then becomes [+2, +1, −3] and [+2, +1, +4]. Note that although the subsumption taxonomy of patterns remains unchanged, the MGDP set varies according to the choice of the minimum value for distinctiveness.

Highly distinctive patterns in a corpus are likely to be long, specific, and will occur in fewer pieces than less

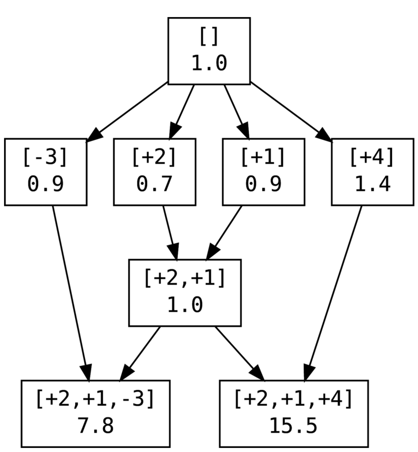

Fig. 2. A small network with eight patterns illustrating pattern subsumption and hypothetical distinctiveness values.

distinctive patterns. The choice of the minimum value of $\Delta(P)$ should therefore be guided by visual inspection of the results and considerations of the data mining purpose. A high value may be used to detect longer patterns in small groups of pieces, whereas low values (close to 1) will tend to detect shorter and more frequent patterns because the MGDP will be high in the subsumption taxonomy.

The MGDP discovery algorithm is a sequential pattern mining method similar to the method of Ayres, Gehrke, Yiu, and Flannick (2002). To compute the MGDP set for a corpus, a depth-first general-to-specific search of the subsumption taxonomy rooted at the empty pattern is performed. At each refinement step, two instance lists (one for the corpus, one for the anticorpus) are locally updated efficiently using hash map representations. The search space is structured so that no pattern need be evaluated more than once. A highly desirable computational property of the MGDP property is that it is *antimonotonic*: if P is an MGDP, and P subsumes Q, then Q cannot possibly be an MGDP. This means that a branch in the search tree may be immediately terminated with success if the minimal specified pattern distinctiveness is achieved by the pattern at a search node.

3. Results

For each category **supertype**, **superarea**, **type**, and **area**, the MGDP set is found, in turn, by considering each class as the corpus $\oplus$ and all remaining classes as the anticorpus $\ominus$. For example, for the **area** category, considering the class **chania**, the pieces in the classes **syria**, **lassithi**, **herakleion**, and **rethymno** (see Figure 1) are taken together as the anticorpus. For all experiments, the MGDP set using melodic intervals was found. A minimum distinctiveness value of $\Delta(P) \geq 3$ was found to be appropriate for this study (it produces neither an excessive number of trivial patterns nor patterns specific to just a few pieces). The same value was appropriate in other musicological studies using the MGDP algorithm (Conklin, 2010a, 2010b).

Table 1 shows a small subset of patterns found, presenting three patterns for each category that are highly distinctive. To restrict results for this presentation only those with piece count $c^{\oplus}(P) \geq 5$ are reported. In addition to a class $\oplus$ and a pattern P, each entry of the table also contains the piece count $c^{\oplus}(P)$, the size of the class $n^{\oplus}$, the pattern distinctiveness $\Delta(P)$, and finally a p-value for the pattern. This is computed using Fisher's one-tailed exact test and represents the probability of finding $c^{\oplus}(P)$ or more corpus pieces in $n^{\oplus}$ draws (without replacement) of pieces from the entire collection. This is a standard statistical test in functional genomics (Falcon & Gentleman, 2008) for detecting significant over-representation of gene ontology terms (here, patterns) in a set of genes of interest (here, the corpus). The remainder of this section discusses one pattern from each category.

Table 1. A selection of patterns found in the Cretan folk song corpus, organized by category.

class $\oplus$	pattern P	$c^{\oplus}(P)$	$n^{\oplus}$	$\Delta(P)$	p-value
	supertype				
dance	[+4, −4]	16	51	19.1	8.4e–06
dance	[−4, +2, +2, −4]	11	51	13.2	0.00069
dance	[+4, +1, +2]	19	51	11.4	3.3e–06
	superarea				
west	[−4, +2, −3]	14	64	∞	0.00023
east	[+1, +2, −2, +2, −2]	13	47	5.9	0.00081
east	[−2, +2, −2, −1, +1]	11	47	5.0	0.004
	type				
syrtos	[−5, −2]	5	22	∞	0.00032
malevisiotis	[+1, +2, +3]	5	13	34.2	8.6e–05
syrtos	[+3, +2, −3]	6	22	21.8	0.00033
	area				
lassithi	[−7, +4]	7	35	14.2	0.0017
lassithi	[−2, −1, +1, −1, +1, −1]	7	35	14.2	0.0017
rethymno	[−4, +2, −3]	10	29	6.6	0.00028

In the **dance** supertype the pattern [+4, −4] is found in all types of dances, throughout the island. The instances of this pattern are found in fast isochronous passages. Four instances, taken from different types of dances are shown in Figure 3 (first row). In most instances the second note occurs on a relatively strong beat. The pattern occurs in all types of dance songs, and in one non-dance song (Stin porta tsi paradeisos).

In the **west** superarea, an interesting and frequent pattern [−4, +2, −3] occurs in nearly a quarter of **west** pieces, in all types of songs (**rizitika**, **city song**, **syrtos**, **pentozalis**, and other dance songs) and remarkably does not occur in any **east** pieces. This same pattern is also distinctive, but to a lesser degree, in the **rethymno** area. In diatonic terms, this represents a drop of a M3, an ascending M2, followed by a descending m3, for instance, E-C-D-B or A-F-G-E. Figure 3 (second row) shows six instances of the pattern: the second and third are found in the same song, but with a different surrounding context. Thus the phrase is varied while this core pattern remains intact. Another **superarea** pattern [+1, +2, −2, +2, −2] occurs mainly in the east, in lullabies, wedding songs, dance songs. The pattern [−2, +2, −2, −1, +1] is found in all types of songs, especially fast dance songs.

For the **syrtos** song type, the pattern [−5, −2] occurs in nearly a quarter of all **syrtos** pieces and remarkably in

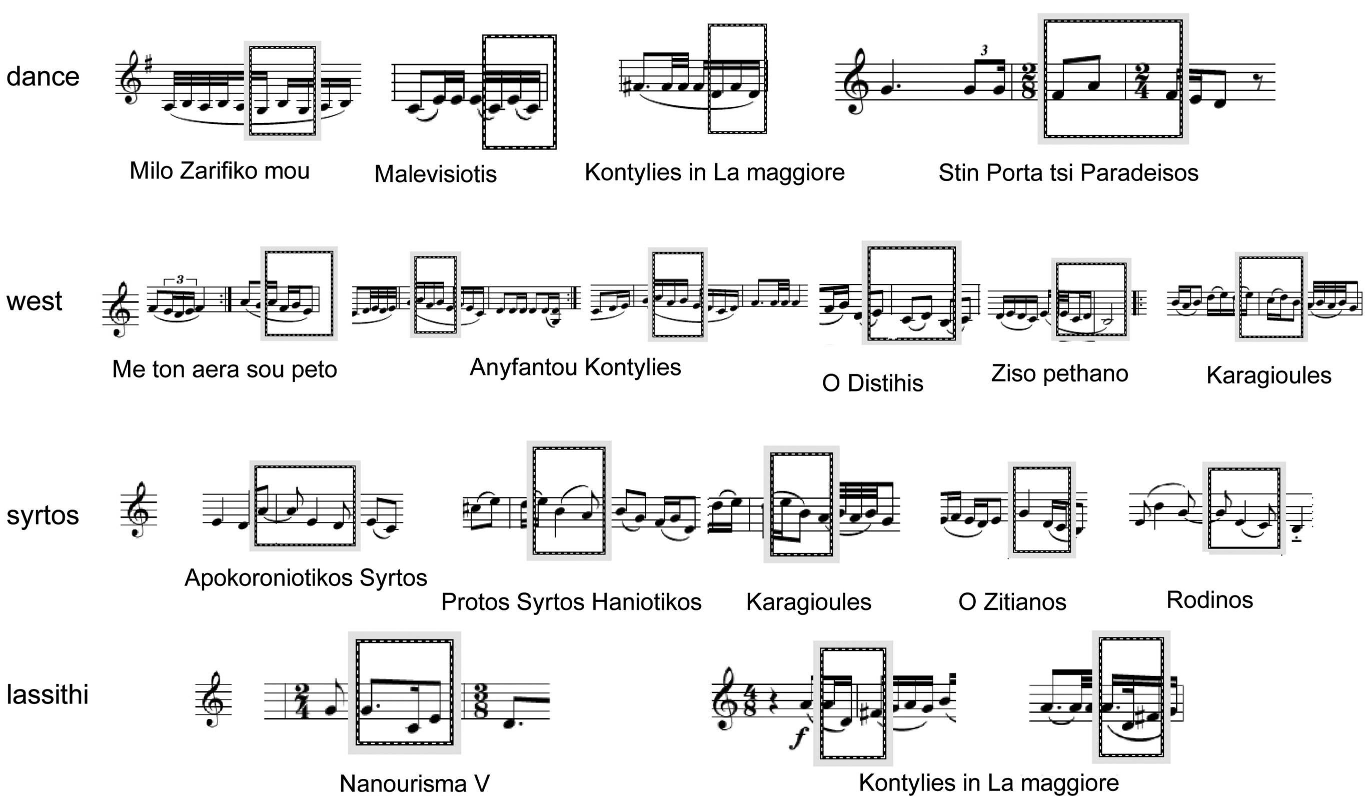

Fig. 3. Instances of distinctive patterns for **dance** [+4, −4], **west** [−4, +2, −3], **syrtos** [−5, −2] and **lassithi** [−7, +4].

no other song types. In diatonic terms, this pattern represents a descending P4 followed by a descending M2, for instance, A-E-D or E-B-A. Figure 3 (third row) shows five instances of the pattern. The rhythmic and metrical aspects of the instances vary considerably, indicating the abstraction provided by the melodic interval pattern.

In the **lassithi** area, the pattern $[-7, +4]$ was found. This pattern indicates a broken major triad. It is found mainly in dance songs and lullabies, and also in some dance songs from Rethymno. Figure 3 (fourth row) shows three instances of this pattern, the last two from the same piece. Another **area** pattern, the long pattern $[-2, -1, +1, -1, +1, -1]$, indicates a stepwise semitone movement changing direction continuously. It is found mainly in Lassithian dance songs, but also in one Rethymno song.

4. Conclusions

There exists a long tradition of computational studies of folk tunes, and indeed some of the earliest applications of computing were to folk music (Bronson, 1959). Many studies involve the clustering or grouping of pieces by similarity, driven by the need to represent and predict geographic localization and diachronic evolution of folk tune styles. Some studies, for example the indexing methods of Bartók on Hungarian folk tunes (Ling, 1997), even pre-date the electronic computer. Many early computational studies, such as those of Lomax (1968) and Suchoff (1970), and recent descriptive studies (Taminau et al., 2009) represent a tune entirely by global features such as length, overall shape, melodic range, ratio of words to music, position of final note, and so on. The melodic patterns discovered by the approach presented here can be viewed as new types of global features specific to a subgroup of pieces.

This paper reported on the application of comparative pattern analysis to Cretan folk songs. The exploratory nature of this study must be emphasized, and the results contain patterns distinctive within classes, but clearly a much larger corpus is needed before the patterns discovered can be considered musically and statistically significant. In Cretan folk music, while expanding the corpus particular attention should be given to more thoroughly populate the small classes such as **epic song**, **city song**, **sousta**, and **kalanda** (see Figure 1). Besides adding more support to distinctive patterns, the statistical significance of equally distinctive patterns will generally be higher for patterns with higher piece counts.

The pattern discovery method allows any number of features to describe music events (Conklin, 2010a). The approach functions on the surface of the music, considering only features of consecutive events. In this study attention was given to patterns of melodic intervals, and further studies with the Cretan folk song corpus could also study more abstract melodic features (for example, melodic contour) or rhythmic features. For the study of joint rhythmic and melodic features, the small corpus size would become problematic, as the data would permit only sparse population of the pattern space.

In this experiment the MGDP algorithm was applied by taking each class in turn as the corpus; alternatively it is entirely possible to form the logical disjunction of classes (e.g. **wedding** or **lament**) to form larger classes. It is also possible to form the logical conjunction of classes from different categories (e.g. **wedding** and **chania**) though the corpus would need to be much larger so that the new classes are not sparse. Concurrently, the data mining algorithm could be extended to use a background song hierarchy so, for example, the pattern $[-4, +2, -3]$ (Table 1) would not have been reported twice (in both **area** and **superarea**) but only the most distinctive occurrence.

There are some general considerations to bear in mind with the analysis of Cretan folk music (and folk music in general): first, transcription of the songs deviates from the live performance, and in the case of Cretan music, there are microtunings and rhythmic novelties that are difficult to transcribe into notation (Dragoumis, 2009). Furthermore, the exportation into MIDI files entails a loss of information, though not critical for the melodic interval analysis of this study. Second, there are songs that go beyond the boundaries of their origin, and they can be found in other areas with possible variations. Third, there may be a fluidity of melodies across styles, especially when different lyrics are sung to the same melody. Fourth, there may be discrepancies in opinion between informants and researchers (Rombou-Levidi, 1999), with the song classifications created mainly by researchers. Finally, in Cretan music especially, a lot of these songs are performed by various artists, and each lyra artist has a unique style of playing, which in this case means adding small characteristic turns and motives. Information about performer might be taken into account but would require more data collection for distinctive pattern analysis.

By extending pattern analysis approaches that work on a single piece, this comparative pattern analysis method produces patterns that recur across several pieces in predefined classes, occurring more frequently than expected given a background or anticorpus set. The identification of these patterns can stimulate discussion on the origin and similarities among pieces disparate at the music surface. The type of computational work reported in this paper can provide ideas to ethnomusicologists for further study, since it can propose distinctive patterns that might be difficult to recognize. It can be proposed that the concept of maximally general pattern discovery is a powerful descriptive approach for folk song analysis.

Acknowledgements

This work was partially funded by the John S. Latsis Foundation, through the project grant entitled 'Comparative computational analysis of Cretan music', in the call 'Meletes 2009'. Thanks also to N. Poulakis and R. Pylarinos for encoding of the corpus, and to Professor Markos Dragoumis from the Melpo Merlie Music Laographic Archives in Athens for discussions on the corpus and the results. Figures 1 and 2 were prepared using the Graphviz software.

References

Amargiannakis, G. (1994). *Gia mia morphologia tou Ellinikou Dimotikou Tragoudiou Towards a morphology of the Greek traditional song* (Technical report). Department of Music Studies, University of Athens, Athens.

Andreoulakis, I., & Petrakis, S. (1994). *Music and Mantinades in Crete*. Athens: Philippos Nakas Music Publications.

Ayres, J., Gehrke, J., Yiu, T., & Flannick, J. (2002). Sequential pattern mining using a bitmap representation. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining*, Edmonton, Canada, pp. 429–435.

Baud-Bovy, S. (2006). *Mousiki Katagrafi stin Kriti 1953–1954 (Music Recordings in Crete 1953–1954)*. Athens: Folk Music Archives Melpo Merlie.

Bronson, B.H. (1959). Toward the comparative analysis of British-American folk tunes. *The Journal of American Folklore*, *72*(284), 165–191.

Cambouropoulos, E. (2006). Musical parallelism and melodic segmentation: A computational approach. *Music Perception*, *23*(3), 249–269.

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

Conklin, D. (2010b). Distinctive patterns in the first movement of Brahms' String Quartet in C Minor. *Journal of Mathematics and Music*, *4*(2), 85–92.

Crane, F., & Fiehler, J. (1970). Numerical methods of comparing musical styles. In H.B. Lincoln (Ed.), *The Computer and Music* (Chap. XV, pp. 209–222). Ithaca, NY: Cornell University Press.

Dragoumis, M. (2009). *I Paradosiaki mas Mousiki (Our Traditional Music) Vol. 2*. Athens: Folk Music Archives Melpo Merlie.

Falcon, S., & Gentleman, R. (2008). Hypergeometric testing used for gene set enrichment analysis. In F. Hahne, W. Huber, R. Gentleman, & S. Falcon (Eds.), *Bioconductor Case Studies* (pp. 207–220). Berlin: Springer.

Hillewaere, R., Manderick, B., & Conklin, D. (2009). Global feature versus event models for folk song classification. In *ISMIR 2009: 10th International Society for Music Information Retrieval Conference*, Kobe, Japan, pp. 729–733.

Hnaraki, M. (2007). *Cretan Music – Unraveling Ariadne's Thread*. Athens: Kerkyra Publications.

Juhász, Z. (2006). A systematic comparison of different European folk music traditions using self-organizing maps. *Journal of New Music Research*, *35*(2), 95–112.

Lartillot, O. (2004). A musical pattern discovery system founded on a modeling of listening strategies. *Computer Music Journal*, *28*, 53–67.

Ling, J. (1997). *A History of European Folk Music*. Rochester, NY: University of Rochester Press.

Lomax, A. (1968). *Folk Song Style and Culture*. Washington, DC: American Association for the Advancement of Science.

Novak, P.K., Lavrač, N., & Webb, G.I. (2009). Supervised descriptive rule discovery: A unifying survey of contrast set, emerging pattern and subgroup mining. *Journal of Machine Learning Research*, *10*, 377–403.

Rombou-Levidi, M. (1999). Psifides horou ston evro. In L. Droulia & L. Liavas (Eds.), *Mousikes tis Thrakis: Mia diepistimoniki Proseggisi* (pp. 157–208). Evros: Syllogos oi Filoi tis Mousikis.

Spuridakis, G., & Peristeris, S. (1968). *Greek Folk Songs*. Athens: Academy of Athens Publications.

Suchoff, B. (1970). Computer-oriented comparative musicology. In H.B. Lincoln (Ed.), *The Computer and Music* (pp. 193–206). Ithaca, NY: Cornell University Press.

Taminau, J., Hillewaere, R., Meganck, S., Conklin, D., Nowe, A., & Manderick, B. (2009). Descriptive subgroup mining of folk music. In *MML 2009: International Workshop on Machine Learning and Music*, Bled, Slovenia, pp. 1–6.

Theodossopoulou, E. (2003). *Oi skopoi stin Elliniki Paradosiaki Moysiki (Tunes in Greek traditional music)* (PhD thesis). Department of Music Studies, University of Athens, Athens, Greece.

Toiviainen, P., & Eerola, T. (2001). Method for comparative analysis of folk music based on musical feature extraction and neural networks. In *Proc. VII International Symposium on Systematic and Comparative Musicology*, University of Jyväskylä, Finland, pp. 41–45.