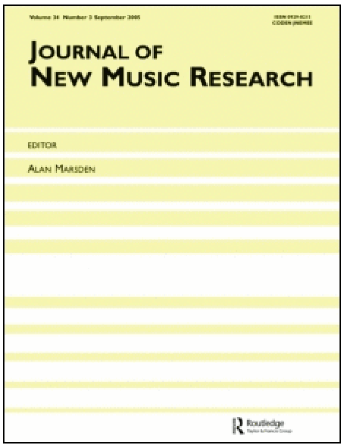


Journal of New Music Research
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713817838

Automatic Segmentation and Comparative Study of Motives in Eleven Folk Song Collections using Self-Organizing Maps and Multidimensional Mapping

Zoltán Juhász[a]
[a] Research Institute for Technical Physics and Materials Science, Hungary

To cite this Article Juhász, Zoltán(2009) 'Automatic Segmentation and Comparative Study of Motives in Eleven Folk Song Collections using Self-Organizing Maps and Multidimensional Mapping', Journal of New Music Research, 38: 1, 71 — 85
To link to this Article: DOI: 10.1080/09298210903029830
URL: http://dx.doi.org/10.1080/09298210903029830

Journal of New Music Research
2009, Vol. 38, No. 1, pp. 71–85

Routledge
Taylor & Francis Group

Automatic Segmentation and Comparative Study of Motives in Eleven Folk Song Collections using Self-Organizing Maps and Multidimensional Mapping

Zoltán Juhász

Research Institute for Technical Physics and Materials Science, Hungary

Abstract

A data-based system for automatic segmentation of large folk song corpora is described in this article. The algorithm is based on a self-organizing map that learns the most typical motive contours. Using this system, the typical motive collections of 11 cultures in Eurasia were determined. The analysis of the overlaps between the cultures allowed us to draw a graph of connections, which shows two main distinct groups, according to the geographical distribution. These groups are connected by the cultures of the Carpathian Basin, which in itself assures the unbroken structure of the system of connections. The mapping of the motive contours into points of an appropriate three-dimensional space opened the possibility to analyse the musical structures of the typical motives in different cultures. Based on the segmentation algorithm, we also defined a melody similarity measure, determining local similarities between the closest motive contours.

1. Introduction

Listening to the song of Ali Bekir in the year 1936 in Anatolia, Béla Bartók was amazed by the similarity to a Hungarian melody published by Zoltán Kodály before. He found that the contours of certain Anatolian and Hungarian melodies are so similar from the beginning to the end, that this relation cannot be attributed to an accident (Bartók, 1949). Similarly, the comparison of the whole contours of Mari, Chuvash and Hungarian melodies led to the conclusion that a genetic relation exists between certain characteristic melody structures of the Volga Region and Hungary (Kodály, 1971).

These exciting results raise the question, whether it is possible to describe a whole and clear system of musical contacts in Eurasia by a systematic comparison of a sufficient number of national or regional cultures? Our current possibilities allowed us to set up 11 folk song corpora, each of them consisting of 1100–2400 melodies, representing different nations as well as regions. We worked out an algorithm classifying our 11 data bases independently, using self-organizing maps. To make an unbiased and general analysis, we developed a method to measure the similarity of the cultures, and analysed this similarity of each culture to the others one by one.

A further question, raised by the above-mentioned classical results, refers just to the method of the analysis. The aim of the above-mentioned classical works was to find parallelism of entire melody structures. The similarity of whole melody contours seems to be really a sufficient condition to find genetic musical relations (Bartók, 1949; Kodály, 1971; Huron, 1996). However, the question rises: do less rigorous requirements also exist? Instead of comparing the complete melody structures, our aim was to find and analyse the smallest independent melodic units. It is well known that folk songs can usually be divided into certain phrases on the basis of musical and textual regularities. In a previous

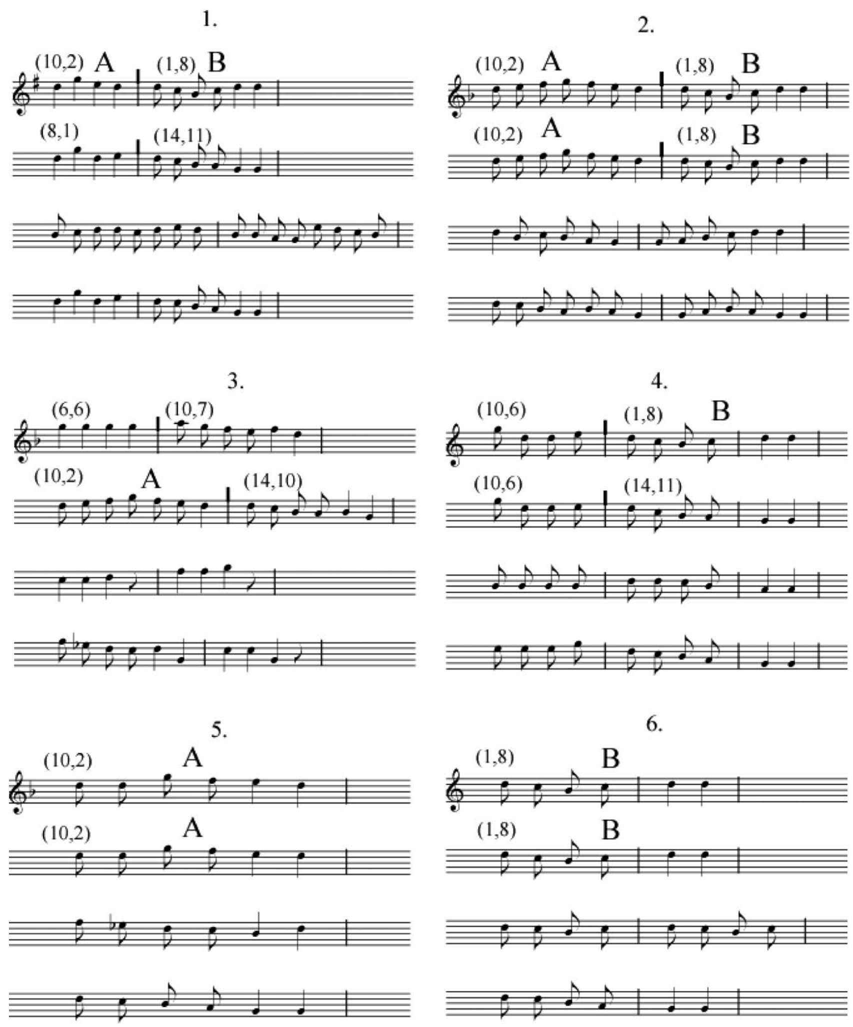
Correspondence: Zoltán Juhász, Research Institute for Technical Physics and Materials Science, POB 49, Budapest H-1525, Hungary. E-mail: juhasz@mfa.kfki.hu

DOI: 10.1080/09298210903029830

work, we have shown some results comparing individual phrases, as well as whole melodies of six European cultures (Juhász, 2006). However, the phrases are not necessarily the smallest intelligible musical units. For example, it seems straightforward to divide the first phrase of melody 1, in Example 1, into two motives with a boundary determined by the bar line. Close variants of the resulting motives, denoted by A and B, can also be identified in the first and second phrases in melody 2. The presentiment that the first phrases of melodies 1 and 2 are constructed by two parts could be justified by some examples with independent appearance of motives A and B. Such examples are shown in the second phrase of melody 3 for motive A, as well as in the first phrase of melody 4 for motive B. Since all of the melodies are Hungarian folk songs in Example 1, the examples justify that motives A and B are really considered as self-contained 'words' of the Hungarian 'musical language'. It is a special fortune that our motives can also be found as individual phrases of melodies 5 and 6. In these latter examples, the second phrases are simple repetitions of the first ones, so they make doubtless that motives A and B are independent and intelligible musical units in the context of Hungarian folk music.

The idea of a segmenting algorithm can be derived from the above example. We want to find the most frequently appearing motive types in a well-defined melody corpus, with the assumption that each motive type may have several variants, since motive A had three different forms, and motive B also had a major and minor representation in the previous examples.

The most frequently applied melody segmentation techniques can be divided into two main groups. In the first group, segmenting is based on pre-defined and data-independent rules (Cambourpoulos, 1998; Singer, 2004). Using such rules, the so-called Local Boundary Detection Model (LBDM) determines a boundary strength value between each couple of notes, and determines the segment boundaries at the maximal strength values (Cambouropoulos, 1996, 1998). Just



Example 1. Different variants of two typical motives (A and B) in Hungarian folk songs. The motive boundaries are marked by short thick lines on the scores.

due to their pre-defined rules, such rule-based methods are not available for the sake of a learning system. The second group of segmenting techniques is based on a learning process to determine the regularities of a given melody corpus. Such regularities can be characterized by the frequencies or conditional probabilities of the motives (Seneff, 1992; Charniak, 1996, 2000). The so-called Markov technique operating with conditional probabilities has already been applied for folk songs (Bod, 2001a,b). In a previous work, we described a further data-based self-learning method for segmenting a large corpus of folk songs, which determines the conditional entropy of the motives and defines an average entropy increment value for a given segmentation (Juhász, 2004).

The segmentation algorithm described in this article belongs to the latter group, due to its learning feature. The learning unit of our system is a self-organizing map (SOM), trained by the contour functions of the motives (Kohonen, 1995; Toiviainen, 2000). The optimal segmentation of a given melody phrase is determined by a minimal average Euclidean distance between the corresponding motive contours and the motive type contours located at the lattice points of the SOM. Thus, our system characterizes the distance of the contours by a Euclidean measure.

Automatic estimation of melodic similarity and classification of large music corpora is a crucial problem of music retrieval, computer-aided folk-song research and music cognition analysis. The idea of a systematic, computer-aided study arises from the 1950s (Freeman & Merriam, 1956). Kodály initiated a project to develop a large digitized European folk-song database and a program system for a comparative analysis in the early 1960s (Csébfalvy et al., 1965). Current interdisciplinary research, based on the cooperation of musicology, artificial intelligence research and data mining, focuses on automatic similarity measurement, segmentation, contour analysis and classification using different statistical characteristics, e.g. pitch-interval or rhythm distribution. The large folk-song corpora of the Essen collection were classified by determining typical contour averages (Huron, 1996). A systematic description of Hungarian folk music in a multidimensional space was also based on the study of the contour, using principal component analysis (Juhász, 2002). A very widely used kind of artificial neural networks, the so-called self-organizing map (SOM) was used for classification of certain musical statistics of Finnish folk music (Kohonen, 1995; Toiviainen, 1996, 2000; Krumhansl et al., 1999), as well as the musical timbre (De Poli & Prandoni, 1996). Self-organizing maps also were applied in a complex model of human music cognition (Leman, 2000).

A wide scale of methods has been published in the field of melody similarity measurement, focusing on the pitch or on the rhythm (Antonopoulos et al., 2007; Chordia, 2007; Volk et al., 2007), applying string edit distance (Orpen & Huron, 1992), earth mover's distance algorithm (Schmuckler, 1999; Garbens et al., 2007), spectrum analysis of a graph representing conditions of consecutive notes in a melody (Pinto et al., 2007), measuring the distance of the contour by Euclidean distance (Juhász, 2006), or dynamic time warping (Juhász, 2007). A SOM-based system has been elaborated for simultaneous analysis of the contour as well as the pitch, interval and duration distributions (Toiviainen & Eerola, 2001, 2002). The general problem of an adequate data representation, similarity measurement and search strategy for music was also discussed (Selfridge-Field, 1999; Schmuckler, 1999; Müllensiefen & Freiler, 2007).

In this paper, we describe a method that measures the melodic similarity in two steps: firstly, the motives of the melodies are assigned to their closest relatives, and secondly, the similarity is calculated as the average of the distances of the motive-pairs.

Although the SOM results in a topographical map, that is, similar contours are located at close lattice points, it is not easy to understand the systematic connections between the motive types on the basis of this topography. In a previous work, we have shown that a mapping of the contour vectors into points of a multidimensional space results in a very visual and clear spatial structure, and the clusters of the resulting point system can be traced back to clear musical principles (Juhász, 2002). To give a visual overview about the characteristic motive types of the eleven cultures, we mapped the motive contour vectors determined by the SOM into this multidimensional 'melody space'.

2. Melody corpora

The basis of our digital melody collections arises from the early 1960s, when representative European folk song collections have been digitized in Hungary, due to the initiation of Zoltán Kodály (Csébfalvy et al., 1965). As a result of this work, representative digital data bases of Slovak (1940 melodies), French (1480 melodies), Sicilian (1380 melodies), Bulgarian (1040 melodies) and Appalachian folk songs (880 melodies) have been generated. The literal sources of the above data bases are representative collections themselves (Stoin, 1931; Sharp, 1932; Slovenské Ludové Piesne, 1950; Canteloube, 1951; D'Harcourt & D'Harcourt, 1956; Favara, 1957). Further folk-music collections have also been digitized, but we focused our work on the largest data bases, counting more melodies than an appropriate limit (1100). To reach this number, we completed the Appalachian collection by Irish and Scottish songs from New Ireland and Nova Scotia from the HUMDRUM collection (HUMDRUM). The Hungarian digital data base was built in the 1990s, based on the work of Dobszay and Szendrei (1992), characterizing Hungarian folk music using 2323 'Old Hungarian Folksong Types'.

We extended the above corpora by collections of Finnish, German, Anatolian and Chinese melodies, as well as a regional collection of Mari, Chuvash, Tatar and Votyak folk songs. The Finnish database contains 2400 songs of the Digital Archive of Finnish Folk Tunes, which counts more than 8000 pieces of spiritual folk songs, folk songs, rune songs, dance tunes and instrumental kantele and jouhikko melodies (Eerola & Toiviainen, 2004). The German collection is constructed by 2220 melodies of the Erk and Boehme databases of the Essen Folksong Collection (Schaffrath, 1995; ESAC). The source of the 2200 songs of our Chinese melody corpus is also the Essen Folksong Collection. The melodies arise from the Han nation, as well as from the Shanxi district of China (Schaffrath, 1995; ESAC). The database of 2200 Anatolian folk songs has been assembled by János Sipos from his private collection and from the Archive of the Turkish Radio (Sipos, 2000, 2006). The folk songs of the nations living in the Volga Region (Mari, Chuvash, Tatar and Votyak) has been studied and collected by László Vikár and Gábor Bereczki (1971, 1979, 1989, 1999). We digitized their collection to build a database of 1670 melodies representing the region as a whole.

Certain problems arise from the fact that the principles of the transcriptions are very different in our sources. Some of them, focusing on the invariable main notes, ignore the ornamentations, but others are considered as accurate transcriptions of the recordings. A further problem emerges from the fact that some collections apply a general, others different ending notes. To assure uniform conditions, our method focuses on the main movement of the contour, and transposes all melodies to the common final tone G automatically.

3. The segmentation algorithm

The operation of our segmentation algorithm is demonstrated in Figure 1. The input of the algorithm is a melody phrase selected randomly from the data base. At the beginning of the process, the D-dimensional motive type contour vectors assigned to the lattice points of the SOM, $\underline{w}_{i,j}$ are filled by random numbers. The choice of $D = 32$ proved to be sufficient for our data base. The processing is constructed by the steps as follows:

1. In the first step, all possible segmentations of the phrase are produced, with the following constraints:
   - The number of the motives in a phrase ranges between 1 and 6. Thus, the case that the whole phrase is equivalent to one typical motive is also possible. The maximal number of the motives is optional, but we found that 6 is a sufficient value for our data base.
   - A further constraint prescribes that a valid segmentation must have at least one motive which contains at least 4 notes, where the value 4 is also a result of practical considerations. Note that this constraint allows also motives of 2 or 3 notes but prohibits segmentations composed solely by such short motives, to avoid the trivial result that all melodies can be divided into consecutive intervals.
2. The time duration of each possible motive is divided into D parts, and the pitch values belonging to the subsequent time intervals are stored in a D-dimensional vector $\underline{x}$, according to Figure 2. Using this method, the motives are mapped into a D-dimensional space, independently of their actual time duration. This 'linear time warping' of the motives to the same length allows us to use the Euclidean distance as a similarity measure in the algorithm.
3. The optimal segmentation is determined, on the basis of the current estimates of the most typical motive types assigned to the lattice points of the SOM. Let $\underline{x}_{s,k}$ denote the contour vector belonging to the kth motive of the sth possible segmentation, and $\underline{w}_{i,j}$ the current estimate of the motive contour type belonging to the lattice point with the coordinates (i, j). The Euclidean distance between the motive contour $\underline{x}_{s,k}$ and the motive type contour $\underline{w}_{i,j}$ is

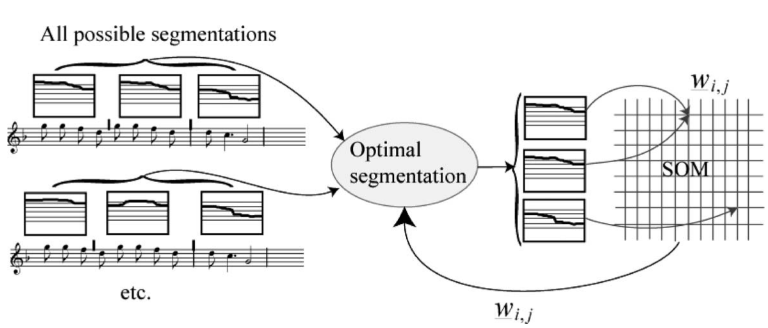


Fig. 1. The learning system for segmentation of folk songs.

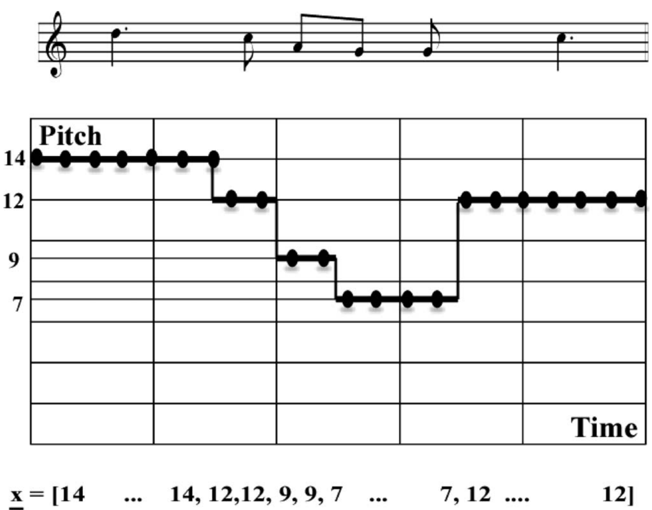


Fig. 2. The generation of the contour vector $\underline{x}$ from the notation. The pitch function increases by 1 with each semitone, with the zero level set to C.

$$\delta_{s,k}(i,j) = \left[\left(\underline{x}_{s,k} - \underline{w}_{i,j}\right)^{\mathrm{T}}\left(\underline{x}_{s,k} - \underline{w}_{i,j}\right)\right]^{1/2}. \qquad (1)$$

The motive contour vector $\underline{x}_{s,k}$ is assigned to the lattice point of the SOM where the distance value of Equation 1 is minimal:

$$d_{s,k} = \min_{(i,j)}\left(\delta_{s,k}(i,j)\right). \qquad (2)$$

Thus, each motive of the segmentation is assigned to the most similar motive contour type vector of the SOM. The error of the *s*th segmentation is defined as the average of the distances between the motive and the corresponding motive type contours. Supposing that the *s*th segmentation is built up by K motives, the error of the *s*th segmentation is

$$\Delta_s = \frac{1}{K}\sum_{k=1}^{K} d_{s,k}. \qquad (3)$$

The criterion to select the optimal segmentation S is defined by the minimal error

$$\Delta_S = \min_{(s)}\left(\Delta_s\right) \qquad (4)$$

4. The SOM is trained with the motive contour vectors of the optimal segmentation $\underline{x}_{S,k}$, using the well-known algorithm. Each $\underline{x}_{S,k}$ determines a 'winner' motive type contour on the SOM according to Equation 2, and the winner vectors $\underline{w}_{m,n}$ are modified towards the corresponding motive contour (denoting a winner position by (m, n) on the SOM). The motive type vectors located in the surroundings of a winner are also modified, while the radius defining the surroundings decreases during the training steps (Kohonen, 1995).

Repeating the above algorithm many times, the mean error of the optimal segmentations Δ_S tends to a minimal value, while the motive type vectors converge to very clear elementary melody contours. After training, any melody phrase can be segmented using the resulting motive contour vectors and Equations 1–4. The melodies in Example 1 illustrate the efficiency of the method, since the segmentation boundaries of the first and second phrases were determined using the above-described algorithm. The coupled numbers refer to coordinates of the SOM accomplishing the segmentation.

The input data vectors are usually invariable during the training process of self-organizing maps. In our system, however, they are variable, because the choice of the optimal segmentation using Equations 1–4 depends on the current state of the motive type vectors $\underline{w}_{i,j}$ [see Equation 1]. Since $\underline{w}_{i,j}$ are modified during the learning process, the optimal segmentation itself depends on the current state of the SOM. In other words, there exists a feedback between the segmentation and the learning algorithm, thus, our system converges to an optimal training and feature vector set in parallel.

4. Motive type maps of eleven cultures

We determined the motive type maps for the eleven cultures one by one, training the system with the first and second melody phrases. The size of the SOMs was fitted to each data base systematically. We defined a common threshold value for the optimal segmentation error Δ_S, to decide whether the segmentation was successful or not. The sizes of the squared maps were increased while the increase of the ratio of successfully segmented melodies exceeded a critical value. We found that the resulting map sizes varied in rather weak limits, thus, we used map sizes of 18*18–20*20 lattice points for all cultures.

The training of 11 SOMs by melodies of 11 cultures yields 11 national sets of motive contour type vectors. Figure 3 shows some examples for motive type vectors determined on different national maps. The vectors are represented in 8 groups in the figure, each of them containing very similar contours. Since all of the 8 groups contain at least 5 variants of different cultures, the similarity of the curves refers to a rather international system of the motives. Example 1 shows representations for motive types denoted by A and B in Figure 3, too. Thus, the motives attributed to Hungarian melody parts in Example 1 are also typical in several other cultures.

The contour groups describe rather simple musical movements in Figure 3. The horizontal couples of contour types correspond to descending as well as ascending movements between 5–8, 5–b3, 5–1 and V–1,

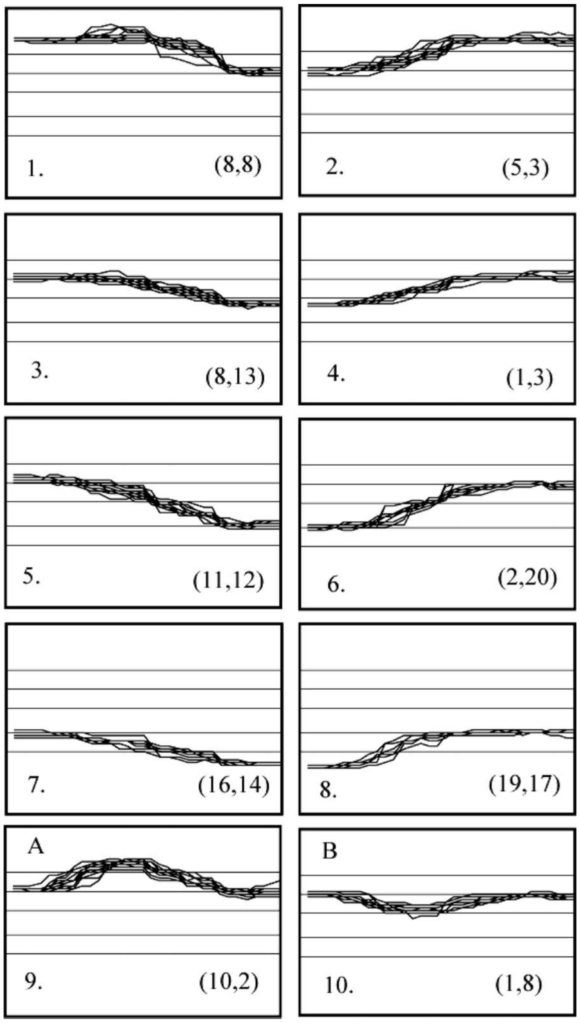


Fig. 3. Some typical motive contours determined by the SOMs of 11 cultures. Horizontal couples show descending-ascending couples between the same beginning and ending notes. Each group contains motive types of at least 5 different cultures. The coordinates refer to the location of the motives on a common SOM.

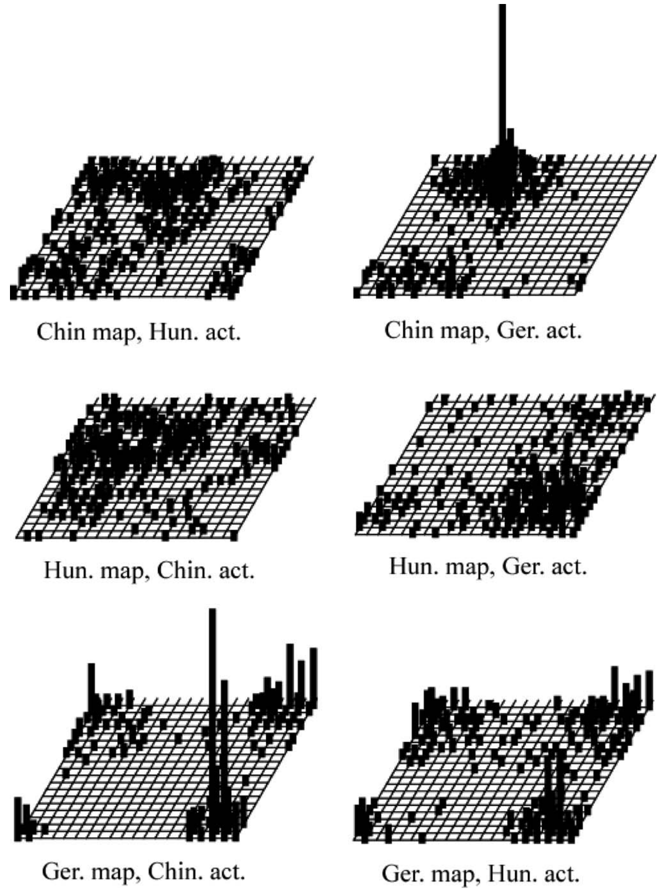


Fig. 4. Cross-activation of Chinese, Hungarian and German maps of motive types. The height of the columns is proportional to the number of foreign motive types activating the corresponding motive type of the map.

denoting the tonic by 1, the fifth by 5, the minor third by b3, etc. Motive types A and B describe more complex 'convex' and 'concave' domed structures, beginning and ending at the fifth.

After training, any 'input' vector of dimension D can be assigned to the most similar motive type vector of the SOM, according to Equation 2. Using this process, being often called 'classification', the motive contour types of a given culture can also be classified on a SOM of any foreign culture. We call the process classifying all contour type vectors of culture A on the SOM of culture B an 'activation of map B by culture A'. We indicated the number of contour type vectors of the activating culture A assigned to a lattice point of the activated map B by columns in Figure 4. This technique activates special patterns on the SOM, visually demonstrating the part of map B being in contact with culture A. For example, Figure 4 shows that the pattern activated on the Chinese map by German motives is a subset of the Hungarian pattern. The activation of the German map by Chinese and Hungarian motives yields nearly the same patterns, while Chinese and German motives activate nearly complementary parts of the Hungarian map. The lesson of these experiences could be summarized in the statement 'common motives of Chinese and German music are also found in Hungarian music, but most of the common motives of Hungarian and German, as well as Hungarian and Chinese cultures are separated into two groups'.

This example shows that an analysis of cross-activated patterns may yield valuable statements about musical relations, however, it is rather complicated to draw a general picture about the contacts of eleven cultures by this way. Therefore, we focus on a very simple feature of cross-activated maps, and characterize the intensity of the contacts between cultures A and B by the number of lattice points activated by motive types A on the SOM of B, divided by the total number of the lattice points of map B. In other words, we define the similarity of cultures by the ratio of the area activated by the motives of culture A on the map of culture B to the total area of map B. We accomplished this measurement for all couples of the studied cultures, independently of any musicological, historical or ethnological precognition or hypothesis.

The diagram of Figure 5 shows the results of the above similarity measurement of cultures. The order of the columns in a group corresponds to that of the group itself. For example, the strongest activator of the Anatolian map is the Anatolian motive set itself (~ 83%), while the second, third and fourth activators are the Sicilian, Hungarian and Slovak motives. The edges on the graph in Figure 5 are drawn in the same way: an edge is drawn between node A and B, if A is one of the 3 most activating foreign cultures on the SOM of B. According to the above example, nodes 'Sic', 'Hu' and 'Sl' are connected to the node 'Anat', and the arrows distinguish between activated and activator nodes. It makes the importance of this distinction very clear that most of the nodes have more than 3 edges, but exactly 3 arrows point to any node. For example, the remaining edges of node 'Hu' indicate that Hungarian motives belong to the 3 strongest activators of Chinese, Finnish and Appalachian maps. The arrows on both ends of edges pointing to nodes 'Sic', 'Vol' and 'Sl' indicate a mutually strong activation between Hungarian as well as Sicilian, Volga and Slovak motives.

The graph constructed on the basis of the above procedure shows a very illuminating picture. The upper part, containing Chinese, Volga, Sicilian and Anatolian nodes would be completely distinct from the lower part containing German, Finnish, Bulgarian, Appalachian and French nodes in the absence of Hungarian and Slovak vertices. In other words, the unbroken structure of the 'Eastern' (Chinese, Volga, Sicilian, Anatolian) and the 'Western' (German, Finnish, Bulgarian, Appalachian, French) sub-graphs is assured merely by the sub-graph of the Carpathian Basin (Hungarian and Slovak). The special role of these latter nodes is sustained by the fact that an omission of any other couple of nodes would not sunder the unbroken structure.

To draw a general picture about the conditions shown in the graph, we trained a SOM with the unified set of all national motive type contours. Since many of the

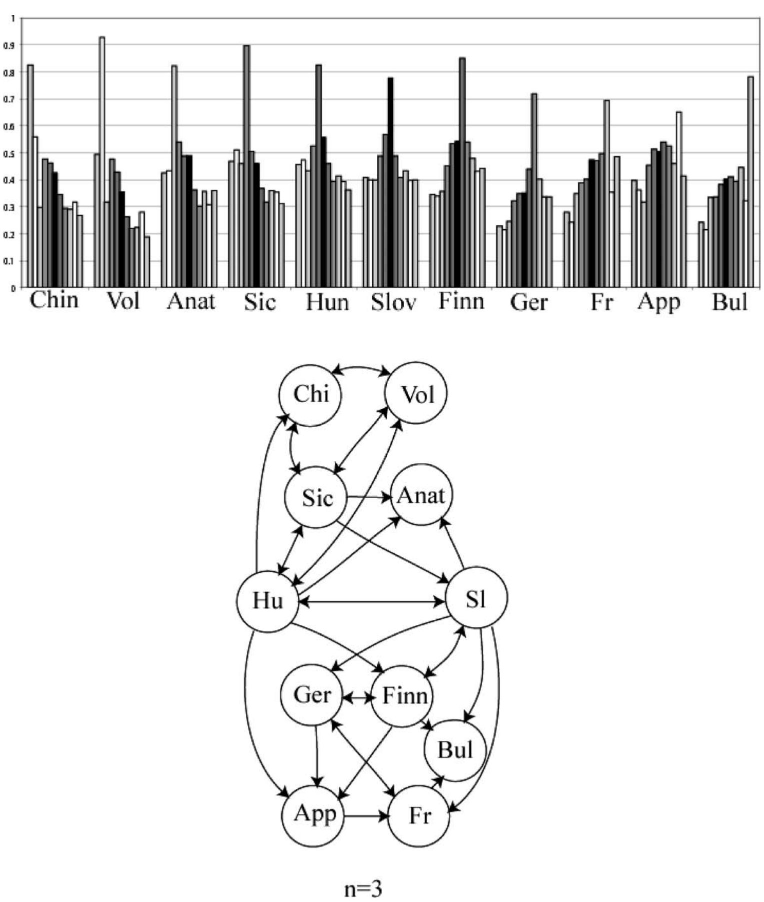


Fig. 5. The upper diagram shows the ratio of the number of activated lattice points of the national SOMs to the total map size, after the effect of different motive type collections. The graph below, showing the most intensive contacts between the cultures, was drawn using the data of the diagram. The arrows trending towards a node indicate the $n = 3$ most activating foreign cultures of it.

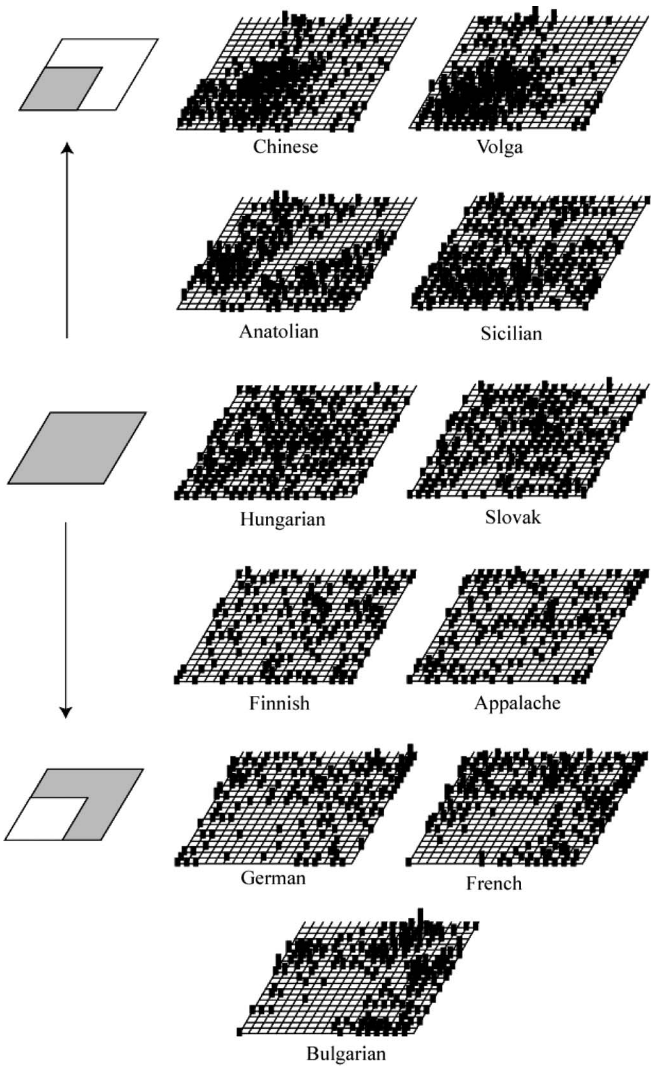


Fig. 6. Patterns due to the activation of the common motive type map by national motive type collections. The sketchy diagrams show a simplification of the Eastern, Western, and Carpathian topographies.

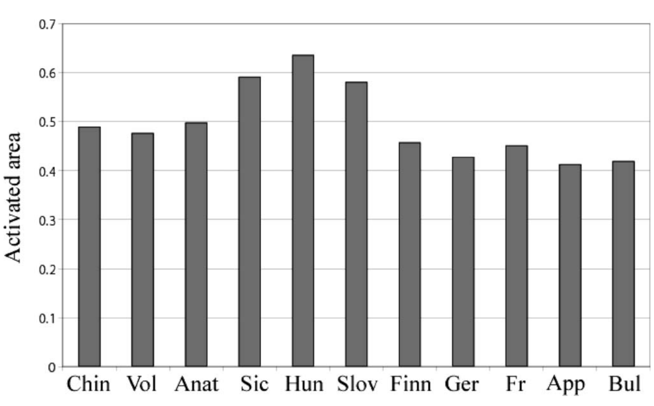


Fig. 7. The ratio of activated common motive types to the total map size, due to the activation by national motive type collections.

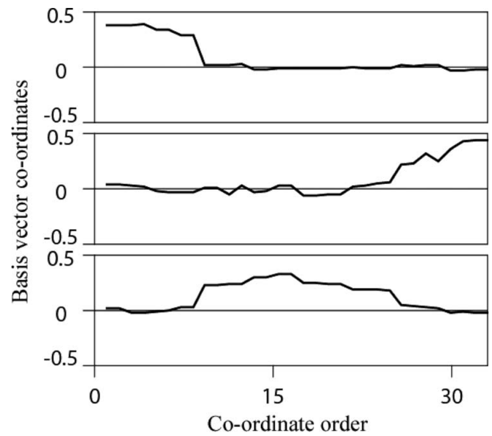


Fig. 8. $D = 32$ dimensional basis vectors of a three-dimensional subspace in the $D = 32$ dimensional space of the contour vectors. Coordinates belonging to the vectors refer to the mean pitch height at the beginning, ending and central part of the contours.

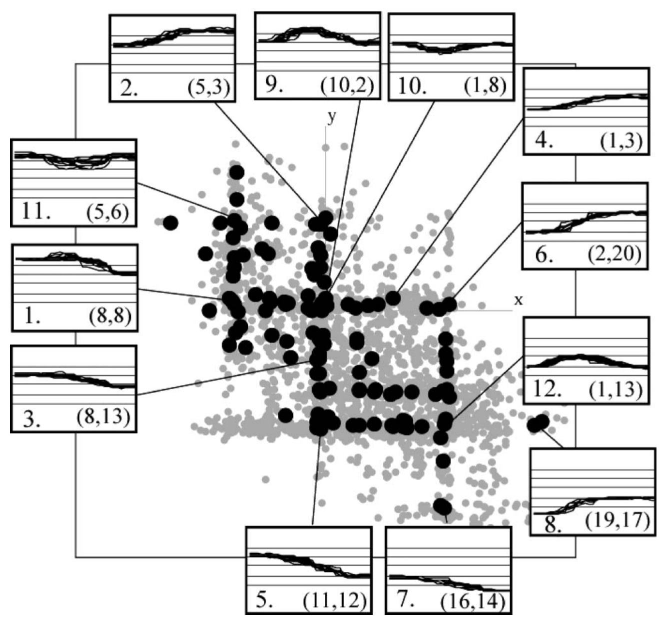


Fig. 9. The map of the most important common motive type contours activated by Hungarian motive types. Grey points in the background correspond to first phrases of Hungarian melodies. The pitch height of the beginning and the ending parts increases from the right to the left, as well as from the bottom to the top. The location of the motive types in Figure 3 is also indicated.

national motive types have very close variants in foreign cultures (as we have demonstrated in Figure 3), a map size of 20*20 lattice points proved to be sufficient to represent all main forms of the eleven cultures. Due to the same reason, the resulting common motive type vectors also have many national variants.

The national cultures activate different areas of the common map. Figure 6 shows that the patterns can be ordered into a system which corresponds to the graph in Figure 5.

Chinese and Volga motive types activate a very well distinguishable square on the common map, while German, French and Bulgarian cultures activate practically just the complementary part. The activation by Hungarian and Slovak motive types yields an even distribution, practically on the whole surface. Finnish and Appalachian patterns seem to form a transition between Hungarian–Slovak as well as German–French–Bulgarian structures. Similarly, the Sicilian pattern can be considered as a transitional form between the Hungarian–Slovak as well as Chinese–Volga structures, while Anatolian distribution seems not completely fitting into this systematization. The ratio of the common motive types activated by national cultures is represented in Figure 7. The diagram supports a central position of Sicilian, Hungarian and Slovak motive types.

5. The musical system of common motive types

What are the concrete musical forms contacting the cultures to each other in the graph of Figure 5? A detailed answer to this question exceeds the frames of this work, but the main features of the common motive types appearing in different national cultures can be characterized using a technique described in a previous article (Juhász, 2002): the *D*-dimensional contour vectors determine points in a *D*-dimensional 'melody' space. The principal component analysis of the point system of melody phrases led to the conclusion that most contours

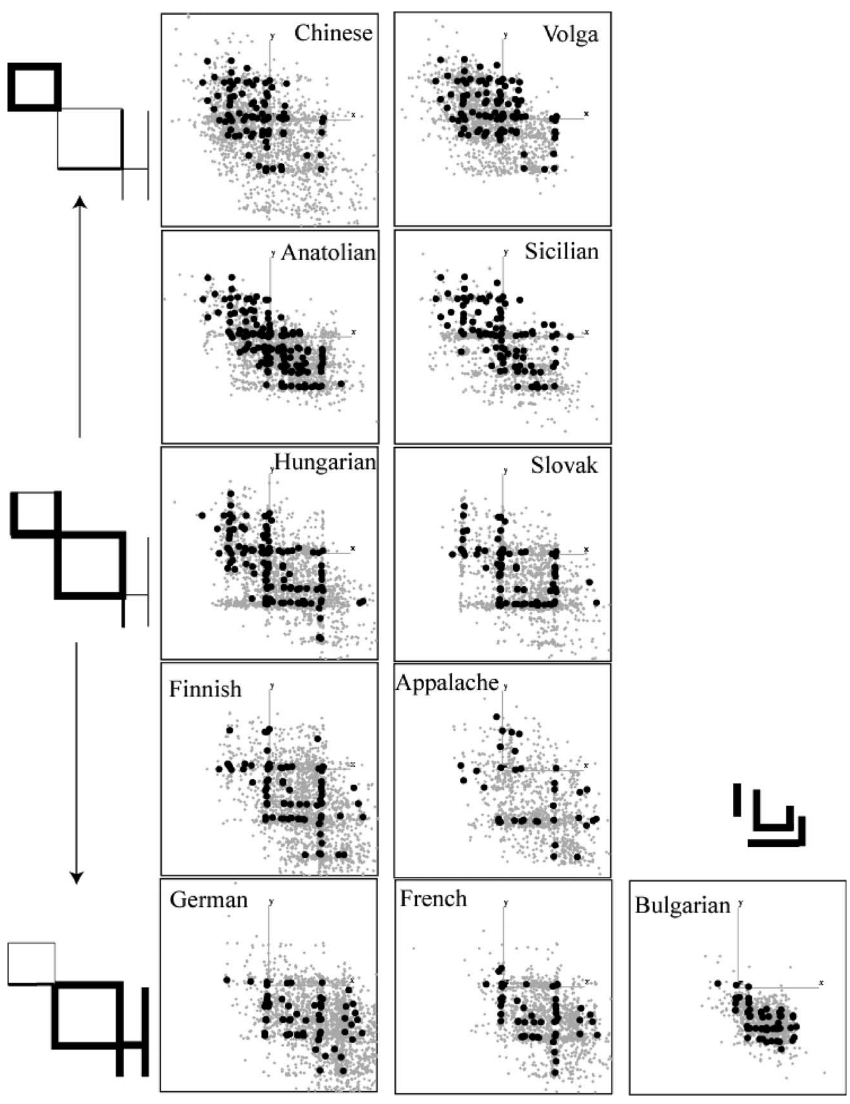
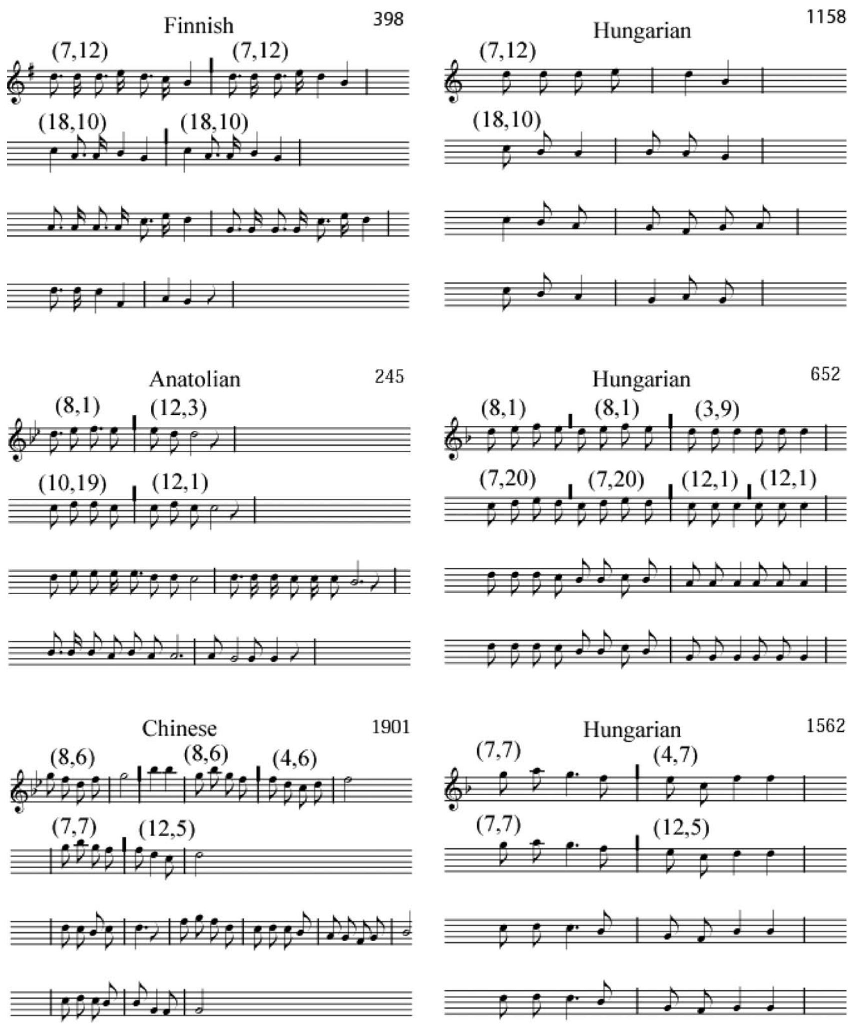


Fig. 10. Maps of the most important common motive type contours activated by national motive types. Grey points in the background correspond to first phrases of national melodies. The sketchy diagrams are simplified versions of the typical point systems.

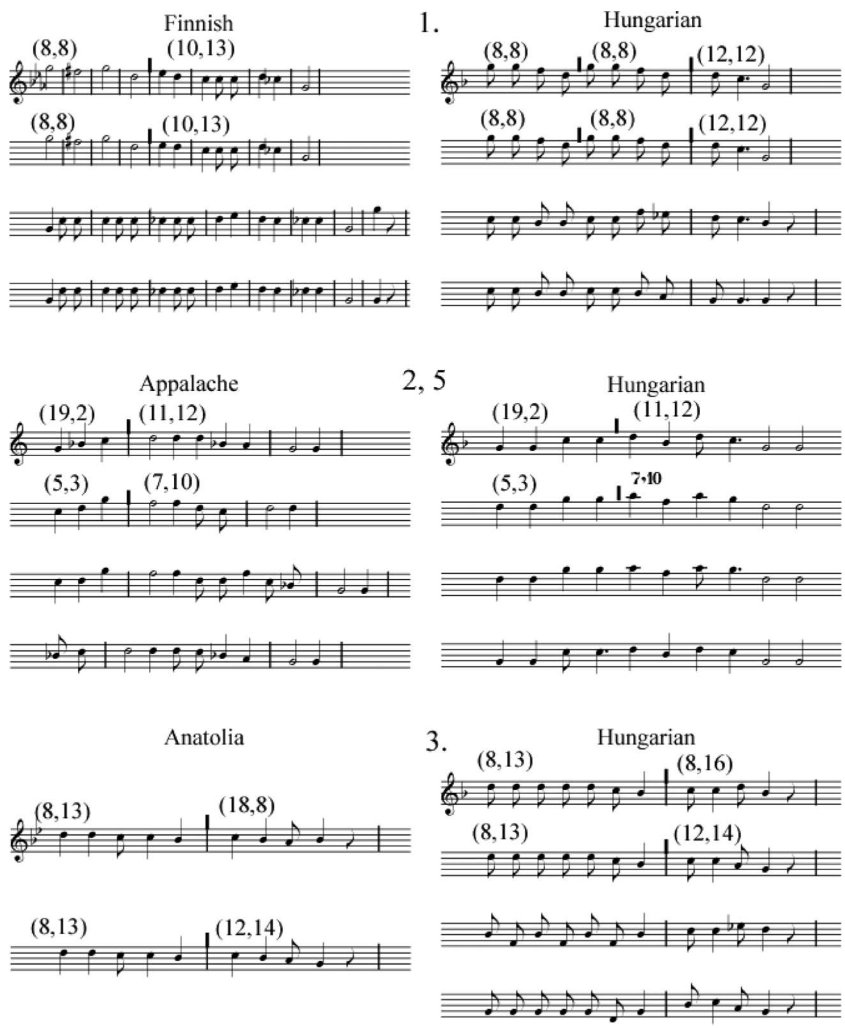
of folk song phrases (and motives, as well) can be well approximated in a three-dimensional subspace of this *D*-dimensional space. The axes of the subspace divide the melody motives into three consecutive parts, therefore the corresponding coordinates describe the mean pitch height at the beginning, ending, and central parts of the motives (see Figure 8).

The contours of the SOM of common motives activated by the highest number of Hungarian motive types are represented in the above subspace in Figure 9. The point system of the first melody phrases of 2323 Hungarian folk songs is also indicated in the figure, by grey points. The pitch heights of the beginning as well as ending parts of the motives increase along the horizontal as well as the vertical axes. The third axis, standing for the central part of the contours is perpendicular to the plane of the figure.

The points corresponding to the contours 9–10 in Figure 3 are also indicated with two additional contour examples (11 and 12) in Figure 9. The musical background of the regularities of the point system can be interpreted with the aid of these contour examples. Comparing the starting and ending notes of motives 2 (5–8), 9 (5–5), 10 (5–5) and 5 (5–1), the common musical feature characterizing the large vertical cluster (represented just by motives 2, 10 and 5) can be easily recognized—the contours mapped into this vertical cluster have a beginning at the fifth. On the other hand, the comparison of motives 1 (8–5), 9 (5–5), 10 (5–5) and 6 (1–5) leads to the recognition that the large horizontal cluster in the middle of the point system contains contours of an ending note of the fifth. By a similar way, contours 1 and 11 as well as 6, 12 and 7 show that the corresponding vertical clusters contain contours starting at the octave as well as at the tonic. Contours 5, 7 and 12 verify that the corresponding vertical cluster is constructed by contours ending at the tonic. The plagal motive contours



Example 2. Melody relations appearing in the similarities of motives. Related melodies are ordered into horizontal couples. The corresponding motive boundaries are marked by short thick lines on the scores.



Example 3. Related melody pairs containing similar motives in their first couples of phrases. The motive boundaries are marked by short thick lines on the scores. The corresponding motive contour types are shown in groups 1, 2, 3 and 5 in Figure 3.

7 (1–V) and 8 (V–1) are rather lonely but characteristic elements of the system.

As a summary of the above experiences, we can say that the large vertical and horizontal clusters contain contours of common beginning as well as ending notes, determined by the tonic, fifth and octave. Contours of high beginning and ending pitch level are mapped into the left upper, while those of having a low beginning and ending note are mapped to the right bottom region of the structure. The system can be simplified to two squares: the larger one, situated in the central part of the figure, is determined by the contours in its vertices, starting or ending at the tonic or the fifth ((1–1), (1–5), (5–5), (5–1)), while the smaller one, situated in the upper left part, is determined by the fifth and the octave ((5–5), (8–5), (8–8), (5–8)). The squares contact on vertex (5–5).

This schematic system is shown in Figure 10 in the neighbourhood of the Hungarian point system. Casting a glance on the German and French point systems in Figure 10, we can see that the smaller square disappears, but the large square determined by the tonic and fifth is completed by additional large clusters containing plagal contours starting or ending at 1 and V. At the same time, the large square determined by the tonic and the octave becomes fragmentary in the Chinese and Volga point systems, while the upper square, determined by the fifth and octave becomes emphasized. In the Volga system, the small square is extended by motive contours starting or ending at the fourth or octave, forming a larger square around the smaller one.

Thus, the ‘Eastern’, ‘Western’ and ‘Carpathian’ systems of motive contours can be derived as different subsets of a hypothetic common system which involves all characteristic motive contour types. The regional differences arise from the dominance of different parts of this large system. The dominant and fragmentary parts of the common motive type system are drawn by thick and thin lines for ‘Eastern’, ‘Western’ and ‘Carpathian’ cases in Figure 10.

6. Melody similarity measurement using segmentation

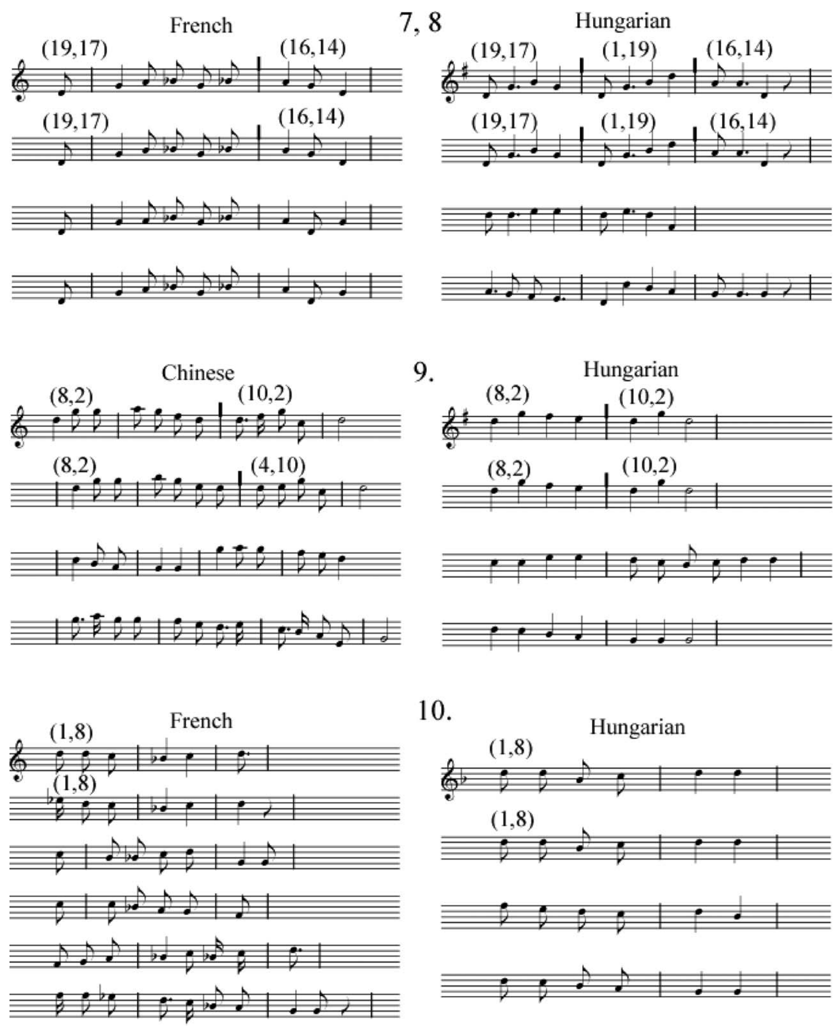
The rigorous requirement of the similarity along the whole melodies may hide certain important musical relations. We illustrated some cases of musical contacts without the similarity of entire contours in Example 2. The segmentation of the melodies is a product of our algorithm, on the basis of the SOM of the common motive types.

The first couple of phrases in the Finnish and the corresponding Hungarian melodies are constructed by the same motives, located at coordinates (7,12) and (18,10). Since the phrases of the Finnish song contain repetitions of these motives, while the Hungarian phrases are identical to the motives themselves, the phrase contours are not similar to each other although the musical contact is evident.

In the second couple of examples, the first phrase of the Anatolian melody is constructed by motives located at the lattice points (8,1) and (12,3). The corresponding Hungarian melody phrase contains a repetition of motive (8,1), and a further one with coordinates (3,9). The contour of this latter motive is very similar to the second one of the Anatolian song, although their locations are different on the SOM. A very similar situation can be recognized in the second phrases, where the first Anatolian and Hungarian motives are close variants with coordinates (10,19) and (7,20). This example shows that the SOM may contain similar motive contours at different lattice points, therefore, the similarity test cannot be restricted to a simple question whether the compared motives belong to the same lattice point or not.

In the third couple of examples, the second and third motives of the Chinese song are located in the neighbourhood of the motives of the corresponding Hungarian melody phrase ((8,6)–(7,7) and (4,6)–(4,7)). The first motive of the Chinese song has no Hungarian parallel. The motives of the second phrases are very

Example 4. Related melody pairs containing similar motives in their first couples of phrases. The motive boundaries are marked by short thick lines on the scores. The corresponding motive contour types are shown in groups 4, 5 and 6 in Figure 3.



Example 5. Related melody pairs containing similar motives in their first couples of phrases. The motive boundaries are marked by short thick lines on the scores. The corresponding motive contour types are shown in groups 7, 8, 9 and 10 in Figure 3.

similar and their locations are also identical. It is also interesting that the third phrase of the Hungarian song can be found in the last 3 bars of the corresponding Chinese melody, and the last phrases are also very close variants. This example shows that the unambiguous relation between all motives is also not a necessary requirement of musical relation.

Based on the lessons of such examples, we defined a similarity measurement method which orders the closest motives of the compared melodies to each other independently of their position in the melodies, and determines the distance of the melodies as the mean of these local motive distances.

Let us denote the motive contour vectors of melody A and B by $\underline{x}_{A,k}$ and $\underline{x}_{B,k}$, respectively. We suppose that the number of motives are different in the melodies, therefore, we denote the motive numbers of melody A and B by M and N, respectively. We calculate the Euclidean distance of the kth motive contour of melody A from all motive contours of melody B and define its local distance as its distance from the closest motive contour of melody B:

$$\Delta_{A,k} = \min_{(m)} \left(\left[\left(\underline{x}_{A,k} - \underline{x}_{B,m} \right)^{\mathrm{T}} \left(\underline{x}_{A,k} - \underline{x}_{B,m} \right) \right]^{1/2} \right), \quad k = 1, \ldots, M. \tag{5}$$

Similarly, we define all local motive distances of melody B from melody A:

$$\Delta_{B,k} = \min_{(k)} \left(\left[\left(\underline{x}_{A,k} - \underline{x}_{B,m} \right)^{\mathrm{T}} \left(\underline{x}_{A,k} - \underline{x}_{B,m} \right) \right]^{1/2} \right), \quad k = 1, \ldots, N. \tag{6}$$

The distance of melodies A and B is defined as the mean of the M local motive distances of A and N local motive distances of B:

$$\Delta_{A,B} = \frac{1}{M+N} \left(\sum_{k=1}^{M} \Delta_{A,k} + \sum_{k=1}^{N} \Delta_{B,k} \right). \tag{7}$$

The contacts between the horizontal couples of melodies in Examples 2–5 were determined using the above distance measure, for the first couple of phrases. The second member of the examples is always a Hungarian song, because this reference gives the best chance for the author to control the results as a musician expert.

7. Summary

We have described a data-based system for automatic segmentation of large folk song corpora, based on a SOM that learns the most typical contour vectors of the motives. The feedback between the learning element and the segmentation unit of the system causes a simultaneous learning of the data and feature vectors. Based on this segmentation algorithm, we also defined a melody similarity measure, determining local similarities between the closest motives. This method is able to indicate musical relations even if they appear in some local contacts of motives, instead of a global contact of entire contours. The efficiency of the method is demonstrated in Examples 1–5. The performance of the segmentation algorithm itself could be verified more exactly by a comparison to other data-based and rule-based techniques.

Using this system, the typical motive collections of eleven cultures were determined as the feature vector set of the SOMs trained by data sets of the eleven cultures one by one. The measurement of the cross-activations of the resulting 11 SOMs by foreign cultures allowed us to draw the graph of connections, defining the edges by the most intensive activators of the cultures. Although this method is completely unrelated to any ethnological, geographical or historical assumption, the graph reflects the main geographical conditions of the studied cultures, distinguishing between an 'Eastern' and a 'Western' group, and it attributes a central role to the Carpathian Basin alone assuring the unbroken structure of the connections.

The impressive similarities of the national motives led to the idea of a common motive contour map. The motive contours were mapped to points of a three-dimensional space which has been derived from a multidimensional space of the melody contour vectors. The very regular structure of the resulting point system opened the possibility to analyse the musical structures of different cultures as different manifestations of a common motive set. This analysis clarified that 'Eastern' cultures prefer motives in their first phrases that move between the octave and the fifth as well as fourth, while the first phrases of 'Western' melodies prefer motives connecting the tonic to the fifth or to a fourth beyond the tonic. The motive system of the Carpathian Basin shows a balance of the above specific systems.

References

Antonopoulos, I., Pikrakis, A., Theodoridis, S., Cornelis, O., Moelants, D. & Leman, M. (2007). Music retrieval by rhythmic similarity applied on Greek and African traditional music research. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

Bartók, B. (1949). On Collecting Folk Songs in Turkey. *Tempo*, New Ser., *13*, Bartok Number, 15–19 + 38.

Bod, R. (2001a). Memory-based models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research*, *31*, 27–37.

Bod, R. (2001b). In *Probabilistic Grammars for Music Proceedings BNAIC 2001*, Amsterdam, The Netherlands.

Cambouropoulos, E. (1996). A formal theory for the discovery of local boundaries in a melodic surface. In *Proceedings of the Troisiémes Journées d'Informatique Musicale (JIM-96)*, Caen, France.

Cambouropoulos, E. (1998). Musical parallelism and melodic segmentation. In *Proceedings XII Colloquium on Musical Informatics*, Gorizia, Italy.

Csébfalvy, K., Havass, M., Járdányi, P. & Vargyas, L. (1965). Systematization of tunes by computers. *Studia Musicologica*, *VII*, 253–257.

Charniak, E. (1996). Tree-bank grammars. In *Proceedings AAAI-96*, Menlo Park, CA.

Charniak, E. (2000). A maximum-entropy-inspired parser. In *Proceedings ANLP-NAACL'2000*, Seattle, Washington.

Chordia, P. (2007). Segmentation and recognition of tabla strokes. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

De Poli, G. & Prandoni, P. (1997). Sonological models for timbre characterization. *Journal of New Music Research*, *26*(2), 170–197.

Freeman, L. & Merriam, A. (1956). Statistical classification in anthropology: An application to ethnomusicology. *American Anthropologist*, *58*, 464–472.

Garbens, J., Kranenburg, P., Volk, A., Wiering, F., Veltcamp, R. & Grijp, L. (2007). Using pitch stability among a group of aligned query melodies to retrieve unidentified variant melodies. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

Huron, D. (1996). The melodic arch in Western folksongs. *Computing in Musicology*, *10*, 3–23.

Juhász, Z. (2002). The structure of an oral tradition – mapping of hungarian folk music to a metric space. *Journal of New Music Research*, *31*(4), 295–310.

Juhász, Z. (2004). Segmentation of Hungarian folk songs using an entropy-based learning system. *Journal of New Music Research*, *33*(1), 5–15.

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

Juhász, Z. (2007). Analysis of melody roots in Hungarian folk music using self-organizing maps with adaptively weighted dynamic time warping. *Applied Artificial Intelligence*, *21*(1), 35–55.

Kodály, Z. (1971). *Folk music of Hungary*. Budapest: Corvina.

Kohonen, T. (1995). *Self-organising Maps*. Berlin: Springer-Verlag.

Krumhansl, C.L., Louhivouri, Y., Toiviainen, P., Jarvinen, T. & Eerola, T. (1999). Melodic expectation in Finnish spiritual hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception*, *17*(2), 151–195.

Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, *17*(4), 481–509.

Müllensiefen, D. & Freiler, K. (2007). Optimizing measures of melodic similarity for the exploration of a large folk song database. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

Orpen, K. & Huron, D. (1992). The measurement of similarity in music: A quantitative approach for non-parametric representations. *Computers in Music Research*, *4*, 1–44.

Pinto, A., Leuken, R., Demirci, M., Wiering, F. & Veltcamp, R. (2007). Indexing music collections through graph spectra. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

Schmuckler, M.A. (1999). Testing models of melodic contour similarity. *Music Perception*, *16*(3), 109–150.

Selfridge-Field, E. (1999). Concepts and procedures. In W.B. Hewlett & E. Selfridge-Field (Eds.), *Melodic Similarity: Concepts, Procedures, and Applications* (pp. 187–209). Cambridge, MA: MIT Press.

Seneff, S. (1992). TINA: A natural language system for spoken language applications. *Computational Linguistics*, *18*(1), 61–86.

Singer, J. (2004). Creating a nested melodic representation: Competition and cooperation among bottom-up and top-down Gestalt principles. In *ISMIR 2004*, Barcelona, Spain.

Sipos, J. (2000). *In the Wake of Bartók in Anatolia*. Budapest: European Folklore Institute (Bibliotheca Traditionis Europeae 2).

Sipos, J. (2006). *Comparative Analysis of Hungarian and Turkic Folk Music – Türk-Macar Halk Müziğinin Karş ilaşt irmal i Araşt irmas i*. (TIKA [Türk İşbirliği ve Kalk inma İdaresi Başkanl iğ i] & the Embassy of Hungary in Turkey 2006, Eds.). Ankara.

Toiviainen, P. (1996). Optimizing auditory images and distance metrics for self-organizing timbre maps. *Journal of New Music Research*, *25*, 1–30.

Toiviainen, P. (2000). Symbolic AI versus connectionism in music research. In E. Mirinda (Ed.), *Readings in Music and Artificial Intelligence*. Amsterdam: Harwood Academic Publishers.

Toiviainen, P. & Eerola, T. (2001). A method for comparative analysis of folk music based on musical feature extraction and neural networks. In *Proceedings of the VII International Symposium on Systematic and Comparative Musicology III International Conference on Cognitive Musicology*, University of Jyvaskyla, Finland, pp. 41–45.

Toiviainen, P. & Eerola, T. (2002). A computational model of melodic similarity based on multiple representations and self-organizing maps. In C. Stevens, D. Burham, G. McPherson, E. Schubert & J. Rewick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 236–239). Sidney, Adelaide: Causal Productions.

Volk, A., Garbens, J., Kranenburg, P., Wiering, F., Grijp, L. & Veltcamp, R. (2007). Comparing computational approaches to rhythmic and melodic similarity in folksong research. In *Proceedings of the ISMIR 2007*, Vienna, Austria.

Melody sources

Canteloube, Joseph. (1951). *Anthologie des Chants Populaires Francais* I–IV. Paris: Durand & C.

D'Harcourt, Marguirete & D'Harcourt, Raoul. (1956). *Chansons Folcloriques Francaises au Canada*. Québec: Paris.

Dobszay, L. & Szendrei, J. (1992). Catalogue of Hungarian Folksong Types. Budapest.

Eerola, T., & Toiviainen, P. (2004). Suomen Kansan eSävelmät. *Finnish Folk Song Database*. Retrieved March 11, 2004, from http://www.jyu.fi/musica/sks/

ESAC (Essen Collection). Retrieved November 11, 2007, from http://www.esac-data.org/data/

Favara, Alberto. (1957). *Corpus di Musiche Popolari Siciliane* (O. Tiby, Ed.) (Vols. 1–2). Palermo: Academic di Scienza Lettere ed Arti.

HUMDRUM. Retrieved November 11, 2007, from http://www.esac-data.org/data/

Schaffrath, H. (1995). *The Essen Folksong Collection in Kern Format*. [Computer database]. Menlo Park, CA: Centre for Computer Assisted Research in the Humanities.

Sharp, C.J. (1932). *English Folk Songs from the Appalachians collected by Cecil J. Sharp* (Vols. 1–2). London: Oxford University Press.

Sipos, J. (1994). *Török népzene*. Budapest: MTA ZTI.

Slovenské L'udové Piesne. (1950). Bratislava.

Stoin, Vassil. (1931). *Chants Populaires de la Partie Centrale de la Bulgarie du Nord*. Sophia.

Vikár L. & Bereczki, G. (1971). *Cheremis Folksongs*. Budapest: Akadémiai Kiadó.

Vikár L. & Bereczki, G. (1979). *Chuvashs Folksongs*. Budapest: Akadémiai Kiadó.

Vikár L. & Bereczki, G. (1989). *Votyak Folksongs*. Budapest: Akadémiai Kiadó.

Vikár L. & Bereczki, G. (1999). *Tatar Folksongs*. Budapest: Akadémiai Kiadó.