An Automatic Pitch Analysis Method for Turkish Maqam Music

Barış Bozkurt

Published online: 20 Aug 2008.

Submit your article to this journal

View related articles

Routledge
Taylor & Francis Group

An Automatic Pitch Analysis Method for Turkish Maqam Music

Barış Bozkurt

Izmir Institute of Technology, Turkey

Abstract

Automatic pitch analysis of large audio databases is essential for studies on music information retrieval and developing a pitch scale theory for Turkish maqam music. However no such study is available. In this article, we first determine the main obstacle as the alignment of frequency analysis results from multiple files. We then propose a new method to automatically detect the tonic of a recording, align the data, and estimate overall frequency histograms from large databases. We show that such histograms can be successfully used for pitch scale (tuning) studies on the recordings of Tanburi Cemil Bey, an undisputed master of the genre.

1. Introduction

Although the "maqam world" corresponds to a regionally very large multicultural area (Touma, 1971; Powers, 1988; Zannos, 1990), computational methods aiming to process maqam music is very limited in number. This study targets development of fully automatic methods for frequency analysis of large audio databases of maqam music. In this study, for practical reasons, the discussions and data are limited to Turkish maqam music. However the basic methods presented are potentially applicable to other maqam music traditions.

Development of computerized music analysis methods necessitates use of music theory. Similarly, music theory research can also profit highly from use of computer-based analysis especially for traditional music (Tzanetakis et al., 2007). Today, for Turkish maqam music studies, it appears that computerized analysis can be a very important tool for progress in music theory research.

A maqam generally implies a set of rules for composition and improvisation. These rules are defined in various dimensions: the pitch scale, melodic progression defining overall ascending–descending characteristics, temporary tonics, possible modulations to other maqams, etc. For an in-depth analysis of maqam music, computerized methods are needed that can provide a large set of data for each of these dimensions. This study aims at providing automatic analysis tools only for the pitch scale dimension of the problem.

It is well known that developing the pitch scale theory for Turkish maqam music is still an open issue. Many studies mention that there exist mismatches between the intervals measured on recordings from master musicians and those specified in the Arel–Ezgi–Uzdilek (AEU) theory (Arel, 1930) although this theory is widely accepted and taught.[1] New studies continue to emerge which compare old scales with actual intervals being played and propose new formulations for more efficient representations of pitch scales (i.e. maqam music notes called "*perde*"s) (Karaosmanoğlu & Akkoç, 2003; Tulgan, 2007; Yarman, 2008).

It is clear that collecting large audio databases and analysing them in a systematic way is very crucial for such studies. However the common approach used in most of the studies is to analyse a very limited number of examples from well-known musicians and draw general conclusions from these limited data. This choice in methodology stems from some difficulties related to the nature of maqam music: it is well known in maqam music theory that notes do not correspond to standardized single fixed frequencies. In addition to the fact that there are 12 possible diapasons called

---

[1] A recent congress was completely dedicated to this topic: "Theory-application mismatch for Turkish Music: Problems and Solutions", organized by Istanbul Technical University, State Conservatory for Turkish Music, 3–6 March 2008, İstanbul.

"*ahenk*"s[2] and dozens of maqams employing a large variety of intervals, musicians may differently interpret certain degrees of some scales.[3] Therefore, it is not at all trivial to develop computerized methods that can perform automatic analysis and gather results from large databases. The lack of such methods continues to hamper research into the actual tuning of Turkish maqam music and the development of a theory conforming to practice.

Furthermore, a literature study of maqam music shows that music information retrieval (MIR) techniques developed for maqam music is very much limited in number. The problems mentioned above stays at the heart of many issues in MIR studies.

This study proposes a novel method for aligning pitch frequencies computed on various recordings in order to facilitate automatic analysis of large databases. The alignment is performed via matching the tonics of each recording. To achieve this, a new tonic detection algorithm for maqam music is presented.

It should be noted that for direct alignment with tonic as reference, the recordings should be from the same class of maqams that have the same note as tonic. In terms of their tonics, maqams can be classified into several broad classes, those with the tonic note being: *rast*, *dügah*, *segah*, *ırak*, *yegah*, *acem aşiran*, *hüseyni aşiran*, *geveşt*, *buselik*, *çargah*[4] (Karadeniz, 1965). The largest set is that of maqams with *dügah* as tonic, followed by *rast* and *segah*. From tables provided in Çevikoğlu (2007), the percentage of songs in the repertory of TRT (Turkish Radio Television Broadcasting) can roughly (by computing percentage on the list provided for 72% of the whole repertory) be calculated as: 44.7% for maqams with *dügah* tonic, 39.2% for maqams with *rast* tonic and 11.8% for maqams with *segah* tonic. This means that, if recordings from these three classes (maqams with tonic as *dügah*, *rast* or *segah*) can be aligned and processed together, one could cover more than 90% of recordings. To avoid confusion and save space, we have chosen recordings only from maqams with *dügah* tonic as test

data in the application part of this study. But the system can be successfully applied to other classes in a straightforward way.

To demonstrate the effectiveness of the method, we present an application of the methods to recordings of Tanburi Cemil Bey (1871–1916). There are various reasons for this choice. Firstly, Tanburi Cemil Bey is the most commonly referred musician for the correctness and precision of his pitch intervals. Many talented musicians claim that his recordings are the most valuable source of information for maqam music (Tanrıkorur, 2004). He has mastered many instruments including tanbur, kemençe, lute, tar, cello, violin, kanun, clarinet, zurna and bağlama. He is known as the creator of the bowed tanbur (yaylı tanbur) and was an indisputable virtuoso of kemençe and tanbur. Although he was mainly known as a performer he was an important composer too. His compositions are still being played not only in Turkey but also in Iran, Iraq, Syria, Lebanon, Egypt, Tunisia, Greece and the Balkans. For further information the reader is referred to Cemil (1947).

In addition, the author thinks that Tanburi Cemil Bey's recordings, being among the rather old recordings of maqam music, represent the traditional pitch frequencies taught through oral education without much influence from the official Arel–Ezgi–Uzdilek (AEU) theory (Arel, 1930). It can be easily seen in recent instrumental methods that tables on fretting reflect more the dictates of the AEU tuning than the demands of actual practice. A similar predicament is also true for formal education; the conservatories teach the AEU system. However, mismatches between theory and actual practice are also acknowledged. Through the tests performed on Tanburi Cemil Bey's recordings for our proposed methods, we show that "theory–practice" mismatches can be directly observed and scrutinized on the average histograms automatically obtained.

Throughout the study, the YIN algorithm (de Cheveigne & Kawahara, 2002) is used for fundamental frequency (f0) estimation together with some post-filters designed specially for Turkish maqam music. These post-filters are explained in Section 2. All recordings used are monophonic to avoid the complex multi-pitch estimation problem since it is not the main issue in our study. In Section 3, we discuss f0 histogram computation. It is common practice to use Holdrian comma (Hc) as the smallest intervallic unit in Turkish maqam music theory. To facilitate comparisons with other studies, we also use the Holdrian comma unit in most of our figures and tables. Section 4 is dedicated to the presentation of our proposed method for combining histograms of multiple recordings and our tonic detection algorithm. Section 5 includes the discussions and conclusions. In addition, the AEU system which is also explained in Yarman (2007) to be equivalent to the Yekta-24 system in terms of intervals used is very briefly explained in Appendix A. The audio

---

[2]*Ahenk* is synonymous with key transposition or diapason. Due to the fact that maqam music notes ("*perde*"s) are named by reference to their relative positions, the same *perde* corresponds to different pitches on different sizes of instruments. The reader is referred to Erguner (2007) and Appendix B of Yarman (2008) for detailed reviews on *ahenk*s, ney fingerings, location of holes on neys with various lengths and the corresponding pitch frequencies produced.

[3]For example for fretted instruments there is a lack of standards and defining appropriate fret locations is an open topic of research. It is observed that the number of frets and their locations vary regionally or due to personal choice.

[4]It should be noted that some names are used both for makams and notes("*perde*"s), like: *Hicaz*, *Buselik*, *Kürdi*, *Segah*, *Mahur*, etc. For discriminating the two, maqam names are written with the first letter being capital.

material used is referred to at various steps in plots and discussions. We preferred to present the brief description of the recording database used in Appendix B so as not to interrupt the natural flow of the text.

2. F0 post-filtering for Turkish maqam music recordings

Fundamental frequency estimation methods are not free of errors. Depending on the acoustic characteristics, various methods have various error rates. It is common practice to design post-filtering methods to correct some of these errors. For some special cases as ours, post-filters may be designed considering the melodic characteristics of a particular type of music. For Turkish maqam music we have verified that the post-filters below can be effectively used.

Filters which impose continuity with the assumption that the following are not likely in Turkish maqam music:

(1) low signal energy and short duration pitch chunks with boundary intervals larger than a perfect fifth;
(2) relatively short duration chunk at an integer multiple octave higher or lower in pitch compared to a long duration previous or post chunk;
(3) melodic dynamic range being larger than four octaves.

The first two properties are implemented via conditional code lines. The last one is realized as a filter by: computing the mean of frequencies measured for a recording, defining maximum and minimum possible frequencies as two octaves higher and lower than the mean for that specific recording and filtering out the

estimates falling outside this range. It is observed that such simple post-filters can correct an important portion of errors for some recordings. Two examples are provided in the Figure 1.

These filters are actually very much needed for old recordings like Tanburi Cemil Bey's due to the high level of background noise causing relatively high f0 estimation errors.

3. Bin-width selection for f0 histogram analysis

Use of histograms for analysis of pitch frequencies is a common practice (Akkoç, 2002; Karaosmanoğlu & Akkoç, 2003; Zeren, 2003; Karaosmanoğlu, 2004). The bin-width selection problem, although very rarely mentioned in the literature, appears to be a critical issue for computer based analysis of f0 histograms that may include peak detection.

A pitch histogram, $Hf_0[n]$, is a mapping that corresponds to the number of f0 values that fall into various disjoint categories (known as bins):

$$Hf_0[n] = \sum_{k=1}^{K} m_k,$$
$$m_k = 1, \quad f_n \le f_0[k] < f_{n+1},$$
$$m_k = 0, \quad \text{otherwise,}$$

where (f_n, f_{n+1}) are boundary values defining the f0 range for the nth bin.

The choice of the bin-width $(f_{n+1} - f_n)$, the width of each category, defines the resolution of the histogram. In theoretical pitch scale studies, it is common practice to use uniform sampling of the whole f0 range. Given the
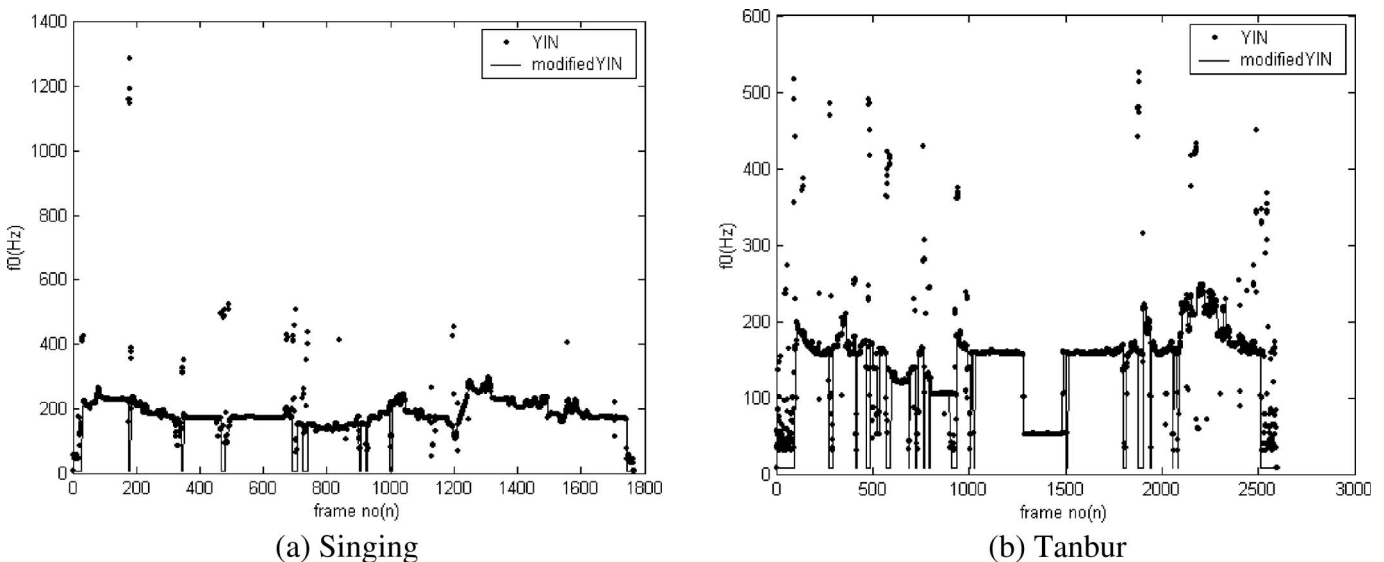

(a) Singing

(b) Tanbur

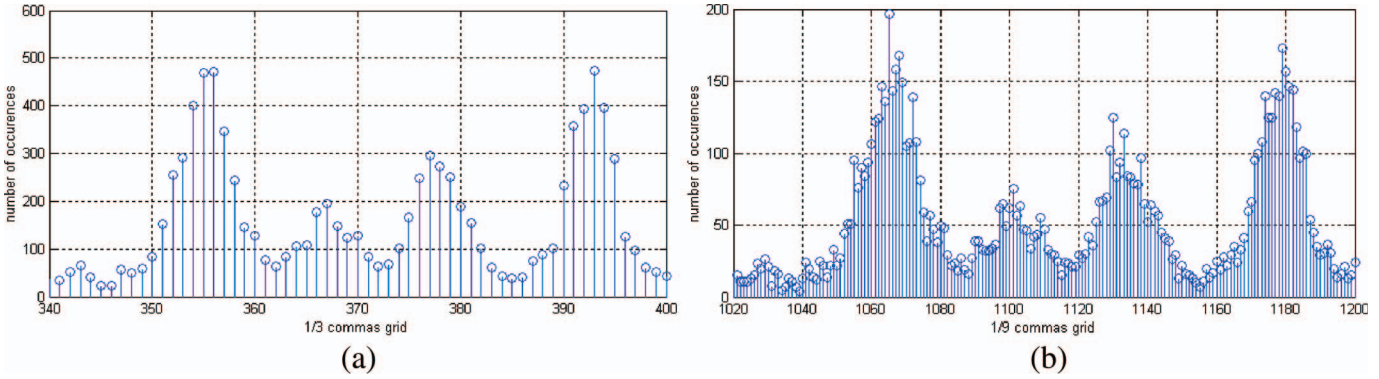Fig. 1. Post-filtering examples: (a) singing, (b) tanbur.

Fig. 2. Histograms obtained with (a) 1/3 Hc grid and (b) 1/9 Hc grid.

number of bins, N, and f0 range (f_{0max} and f_{0min}) bin-width, W_b, is obtained simply by the following:

$$W_b = \frac{f_{0max} - f_{0min}}{N},$$
$$f_n = f_{0min} + (n-1)W_b.$$

One of the critical choices made in histogram computation is the decision of bin-width when automatic methods are concerned. In automatic processing of histograms, peak detection is one of the basic operations; therefore our choice targets facilitation of the detection of note peaks.

A fine grid, i.e. small bin-width, has an advantage in terms of precision but is disadvantageous for automatic peak picking since spurious peaks are produced. This is vice versa for a coarse grid, i.e. large bin-width. As an example we present, in Figure 2, two histograms with two different bin-widths on the same f0 data.

As can be seen on Figure 2, a simple peak picking algorithm would find more than twice as many peaks in the 1/9 Hc grid compared to the 1/3 Hc grid. As a result of empirical tests with various grid sizes, we decided to use the 1/3 Hc resolution, a value that optimizes smoothness and precision of pitch histograms. In addition, this is the highest precision we could find in theoretical pitch scale studies (as used in Yarman (2008)). In all the pitch histograms used in this study, the pitch frequencies measured are indicated as intervals in Holdrian commas with 1/3 Hc precision with respect to the tonic.

4. A method for combining histograms of multiple files

In studies concerning pitch scales of Turkish maqam music, each recording's histogram is studied separately and manually due to the difficulty of gathering analysis results in a systematic way as mentioned previously. In the literature, we could only find one software that facilitates the user to combine results from more than one

recording: "*İcra Analizi*".[5] In "*İcra Analizi*", the user can manually input "calibration" values to align the results from various recordings and derive histograms of played intervals for more than a single recording. "*İcra analizi*" is no doubt an important step towards tackling the problem. However it is clear that processing large databases necessitates manual work (finding the correct calibration value and typing in the system) which is very cumbersome.

Due to the tuning problem mentioned in the introduction, it is more convenient to study intervals rather than exact pitches as preferred in various studies. Two types of intervals typically are used: interval with respect to the previous note played and interval with respect to the tonic of the maqam. The first one has the potential to capture musical context dependent intervals played: the intervals played in descending and ascending melodic lines can be segregated and studied. However it is practically difficult to implement such automatic algorithms robust to ornamentations like glissandos and vibratos frequently used in maqam music (see Figure 12, Appendix A, for example). Due to practical reasons, we use the second as a means to "align" pitch histograms, i.e. the tonic of the maqam serves as an alignment point for two pitch histograms to be combined. All pitches are computed as intervals with respect to tonic and counted independent of the time of occurrence or musical context to form histograms that can be further combined. The payback is the loss of the time dimension and the link of intervals to the musical context.

For this process to be automatic, a tonic detection algorithm is needed. In the next section we present our novel tonic detection algorithm.

4.1 Automatic tonic detection

For recordings with high signal-to-noise ratio, detection of the tonic is very trivial: in theory, a recording in a

[5]http://www.musiki.org/icra_analizi.htm

specific maqam always ends at the tonic as the last note (Akdoğu, 1989). We have implemented a simple algorithm to capture the last note and tested on various recordings. This approach works quite well when the background noise is rather low.

However in old recordings such as Tanburi Cemil Bey's recordings, the energy of background noise sometimes exceeds the energy of the musical signal and the estimated frequencies become unreliable especially towards the end. For this reason a more robust method is developed for tonic detection with the assumption that the maqam of the recording is known (either from the tags or track names since it is common practice to name tracks with the maqam name: "*Hicaz taksim*", for example). For tonic detection in such adverse situations, maqam histogram templates can be utilized as reference. Briefly, our algorithm detects the tonic of a given recording by aligning its histogram with its maqam histogram template which is initiated as a Gaussian mixture from theoretical intervals specified in the AEU system and updated recursively as multiple recordings are aligned. We present the complete process flow (starting from wave files), the tonic detection algorithm and the histogram template construction algorithm in Figure 3.

Given the intervals defined in the AEU system for a specific maqam, a simple theoretical histogram template in the Holdrian commas scale, $H(f_0)$, for the given maqam can be constructed as a mixture of Gaussian functions, $G_k(f_0)$,

$$H(f_0) = \sum_{k=1}^{K} a_k G_k(f_0),$$

$$G_k(f_0) = \exp\left[-0.5\left(\alpha\frac{f_0 - f_k}{W_b/2}\right)^2\right],$$

where f_k is the centre frequency of the Gaussian function defined by the kth degree's interval with respect to tonic.

α is the reciprocal of the standard deviation. f_k, f_0 and W_b are discrete frequency values in Hc with 1/3 Hc resolution.

Assigning the tonic to 0 Hc, the intervals specify the centre frequencies of the notes, f_k, which correspond to the mean of each Gaussian distribution function. For simplicity the weights, a_k, of each distribution are assigned to be equal and this value is set to the maximum value of the histogram of a given recording for which the tonic will be found. The variances of each distribution are also set to be equal and it is a user-defined value to set the width of each Gaussian.

As an example, we present a histogram template for maqam *Hicaz* from the intervals defined in the AEU system in Figure 4.

We have chosen the AEU system intervals for template construction since it is "the official/standardized system" despite its errors in representing the practice. We leave the comparison to use of other tunings for this specific application to future work. Only the initial histogram template is constructed theoretically. The template is updated within a loop of optimization and all templates except the initial are obtained by averaging the real recordings' pitch histograms after alignment. The use of histogram averaging results in discarding relatively rare musical events and construction of a representation of what is common as pitch intervals in a large database. This is convenient for our target but causes a loss of fine details which may be critical in characterizing a given maqam. Certain degrees of certain maqams' pitch scales vary in pitch depending on the melodic line being ascended or descended. Such variations are observed to be relatively small (within less than 1.5 Hc range) and in average histograms these variations result in a widening of a peak width instead of creating additional peaks. On the other hand, it is preferable to filter out some of the details like small peaks due to temporary modulations to other maqams
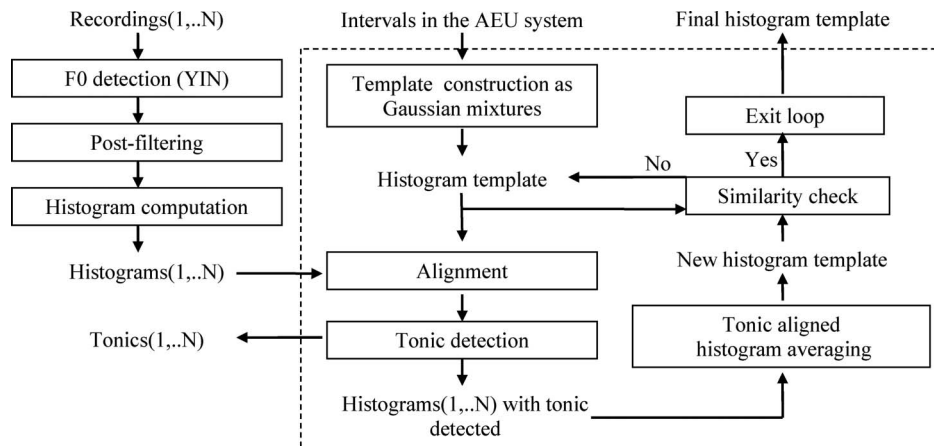

Fig. 3. Tonic detection and histogram template construction algorithm (box indicated with dashed lines) and the overall analysis process. All recordings should be in a given maqam which also specifies the intervals in the AEU system.
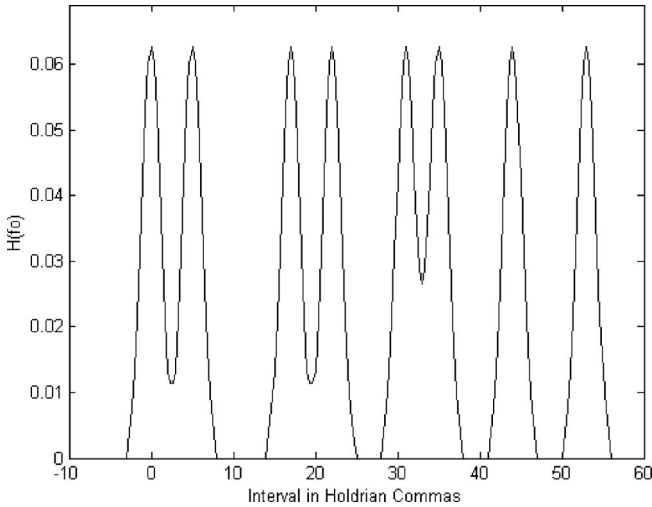
Fig. 4. Theoretical *Hicaz* maqam histogram template. The AEU system scale intervals: [0 5 17 22 31 35 44 53] Hc (Table 1, Appendix A).
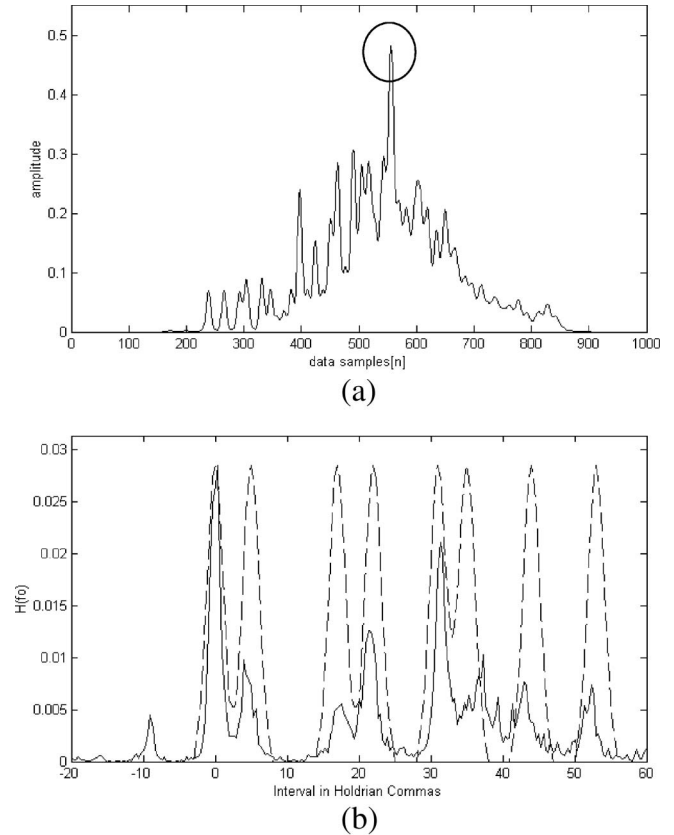

Fig. 5. Aligning the theoretical *Hicaz* histogram template (dashed lines) and the pitch histogram of a recording (Vol. 2, Track 16, *Hicaz Taksim*). (a) Cross-correlation function of the theoretical histogram template and the pitch histogram of the recording. (b) Sychronized plot for the theoretical histogram template (dashed lines) and the pitch histogram of the recording (solid lines).

which may lead to confusion in the pitch scale analysis of a specific maqam. This is especially problematic when automatic peak detection algorithms are used for analysis of the histograms.

Once the initial template of the maqam is available, automatic tonic detection of a given recording is achieved by:

– computing the cross-correlation function between the template and the pitch histogram of the recording;
– finding the maximum correlation point;
– aligning the template and the actual histogram;
– assigning the first peak to tonic.

These steps are represented as two blocks (Alignment, Tonic Detection) in Figure 3.

The cross-correlation function, $c[n]$, for given two real valued signals $x[n]$ and $y[n]$ can simply be computed using the equation:

$$c[n] = \frac{1}{K}\sum_{k=0}^{K-1} x[k]y[n+k].$$

Computing $c[n]$ and finding the index of the maximum peak, one can estimate the optimum amount of shift to be applied on one of the signals such that the two signal waveforms match the most. This is illustrated in Figure 5 for the track: Vol. 2, Track 16, *Hicaz Taksim*.

The index of the maximum of cross-correlation function (marked with a circle in Figure 5(a)) gives us the amount of shift to be applied to the template signal for alignment of the two histogram signals as in

Figure 5(b). The location of the first Gaussian's peak is labelled as tonic and mapped to 0 Hc.

The algorithm described is applied on the recordings described in Section 5 and compared with manually labelled tonics. We present four additional samples in Figure 6.

Even for maqams for which the theory is criticized (examples can be found in Tulgan (2007)) for not matching exactly the intervals of certain degrees (like maqam Saba, Figure 6(d)), tonic detection is successfully performed. The degrees misrepresented in theory are clearly seen on histograms and support the criticism. For example in Figure 6(d), the 4 degree note peak does not exactly match with the theoretical Saba maqam histogram template's 4 degree's peak (corresponding to the note "*hicaz*") as stated in Tulgan (2007).

The tonic detection algorithm is applied on 67 recordings and results for five recordings were problematic. In all other recordings, the tonic was successfully found within ± 2/3 Hc precision. For those five recordings, it has been observed that the repetitions in intervals
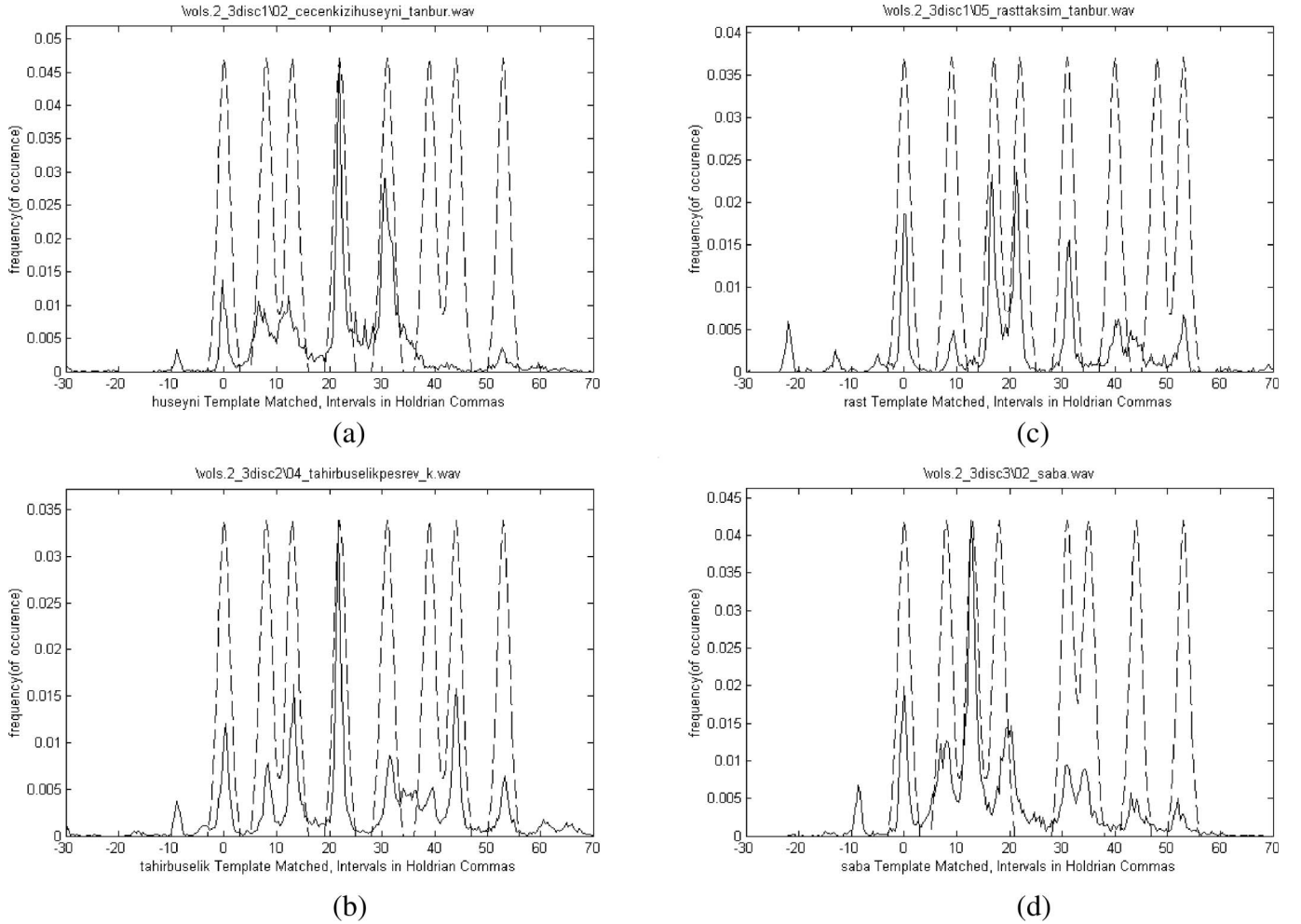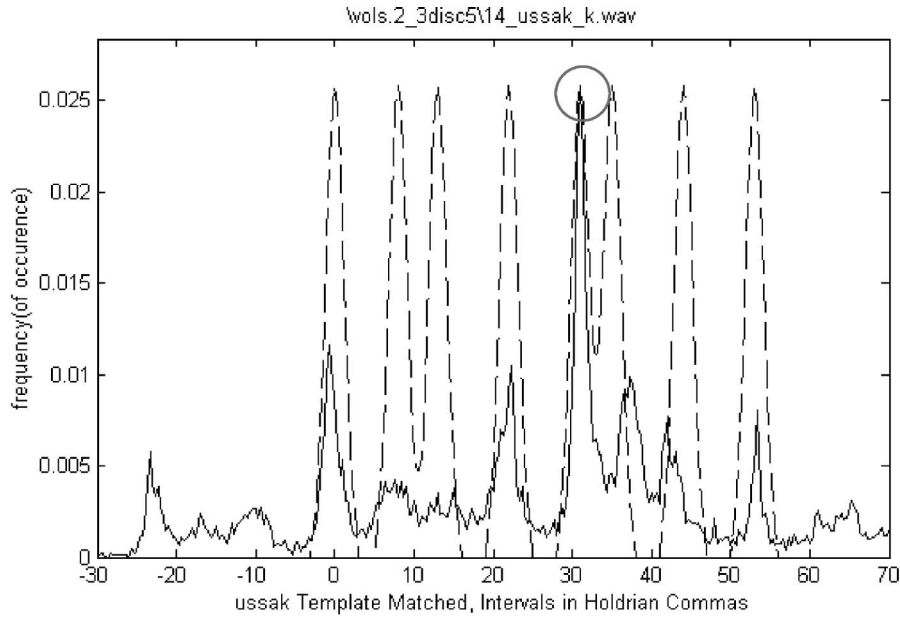
Fig. 6. Histogram matching examples for tonic detection. Maqams of the recordings are: (a) *Hüseyni*, (b) *Tahir Buselik*, (c) *Rast*, (d) *Saba*.

resulted in wrong tonic detection. In addition, these recordings included solos from different instruments (singing and tanbur interchanging solo sequences) which result in addition of f0-regions that intersect but not match exactly in histograms. This typically causes changes in highest peak location and f0-data being concentrated on a rather larger f0-range. For this reason, we think that the algorithm is more reliable when applied to solo recordings of a single instrument such as taksims (which is in fact a very common improvisational form of Turkish maqam music; taksims are also the main material used in music theory research). We leave comprehensive testing for various forms to our future work.
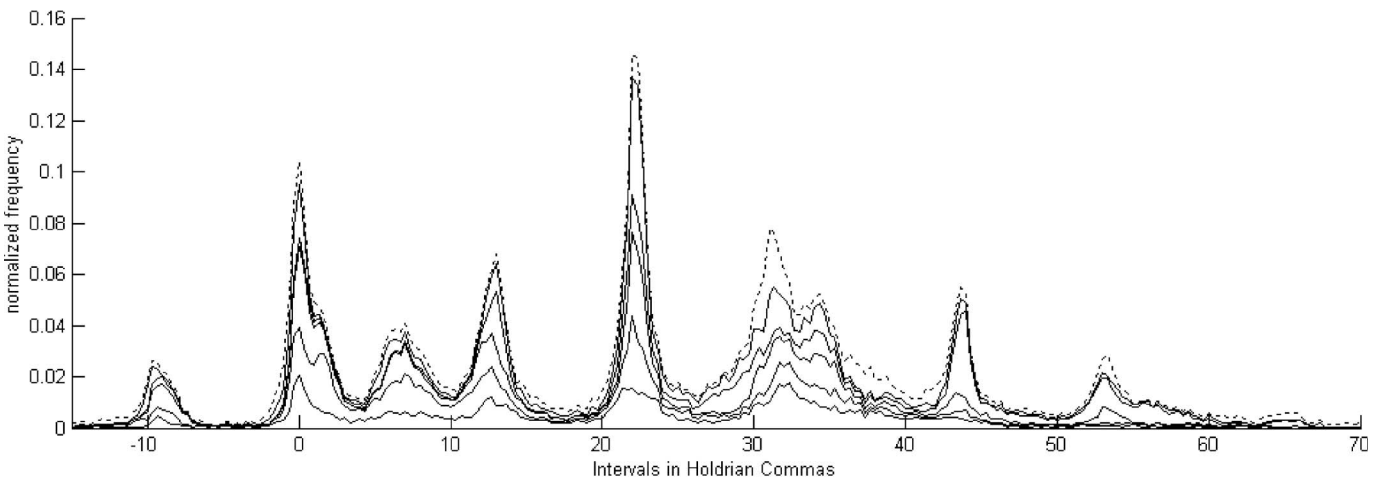
Once tonics are detected, histograms of all recordings are averaged to obtain a new histogram template (*n*th iteration). Fortunately, the contribution of rare tonic detection errors in the template histogram update in the loop is rather negligible. This is illustrated in Figure 7, an example where the algorithm is applied to six recordings

in maqam Uşşak. For the last recording, the tonic is misdetected (Figure 7(a)) and the *n*th template is computed accumulatively for demonstration, the last addition being the misdetected example's histogram.

In Figure 7(a), we observe that the tonic is located close to the middle of the most frequent f0-range spanned which is not very likely for single instrument playing. This example is a singing–tanbur duo where each instrument plays solo in turn. The tanbur's tonic is the labelled peak and the singer's tonic is the left-most peak where a misdetected tonic is located at 0 Hc. In Figure 7(b), we present the accumulation of the new histogram template from recording 1 to 6. The last addition is performed with the example presented in Figure 7(b) and its contribution is plotted with dotted lines. It is clear that the general shape is not much altered, the main change being a sharpening of the peak at 31 Hc since it is the largest peak in the added histogram. The template histogram can be used in theoretical studies on pitch scales perfectly.

(a)

(b)

Fig. 7. (a) Tonic misdetection example (Vol. 5, Track 14, *Uşşak*) (correct tonic is labelled with a circle) and (b) the 2nd template computed accumulatively where the last contribution in dots come from the histogram in (a).

It is shown in Figure 3 that the algorithm updates the template histogram in a loop: all recordings are aligned with a given template to form a new template by averaging, and the new template is again used to align histograms. The loop continues until the template computed in the $(n-1)$th cycle is the same with the template computed in the nth cycle. Since histograms are discrete signals with 1/3 Hc resolution, a few loops are sufficient to reach the termination. In Figure 8, the templates computed for the six recordings in maqam Uşşak are shown (iteration starts on the theoretical template with the dashed lines).

For this example, there is a mismatch between theoretical template peaks and estimated final template:

the peak just after the tonic at 0 Hc. All studies discussing the weak points of the AEU theory, without any exception, state the mismatch of the theoretical and performed intervals for this second degree in Uşşak maqam's pitch scale. Since the final template is obtained as the mean of aligned histograms, the error in theory is not reflected in the final histogram. Theory only serves for initial alignment of histograms. For studies targeting construction of models from the data (i.e. using the analysis results of actual practice as the main reference in building the model), this is very crucial. One alternative to using theoretical intervals for constructing an initial histogram template is to provide the system a real recording histogram example with manually

labelled tonic. This would need to be done once for each maqam.

4.2 Aligning histograms of different maqams

In the previous section we have explained an automatic algorithm to compute histogram templates from recordings for a given maqam. The templates computed can be successfully used in analysing intervals used in a specific maqam "in practice".

As we have mentioned in the introduction, tonic can successfully serve as a reference point for alignment of recordings from different maqams given that the tonic is the same. In Figure 9, we present aligned histograms of five maqams (with tonic *dügah*) from 17 recordings of Tanburi Cemil Bey.

In Figure 9, we observe that most of the peaks match theoretical intervals. However for two peaks this is not the case. A peak observed at 6.66 Hc for the second degree of maqam *Uşşak* does not match with the theoretical interval specified to be 8 Hc. In many studies (Özkan, 1998; Çakar, 2004; Erguner, 2007; Tulgan, 2007) we find the following explanation for this second degree: the second degree of maqam *Uşşak*, the *segah* note, is played 1–1.5 Hc lower than indicated
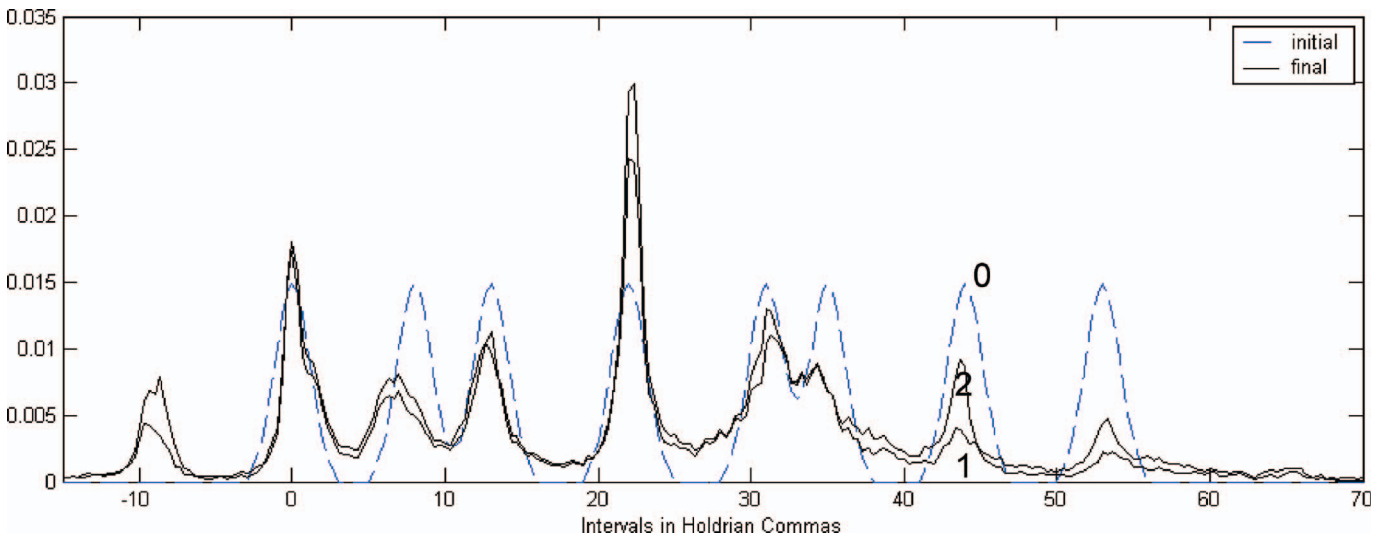

Fig. 8. The templates computed iteratively by averaging N ($=6$) files in each iteration. Iteration number indicated on the resultant template.
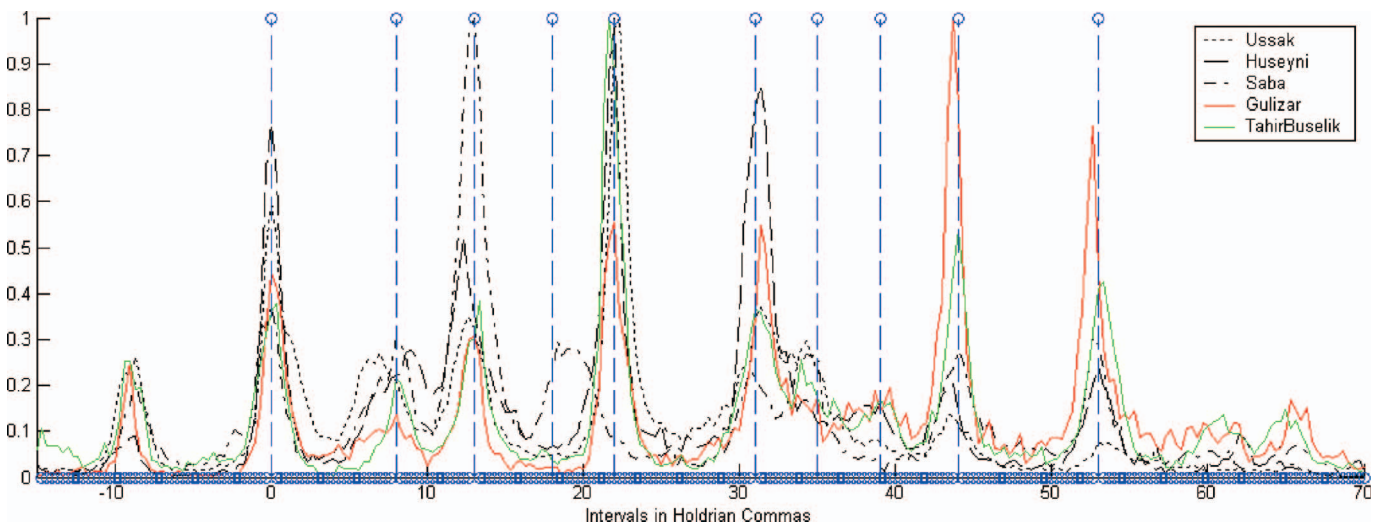

Fig. 9. Tonic aligned and normalized histograms for recordings with *dügah* as tonic: 5 maqam template histograms superimposed (5 histograms computed from 17 recordings). Vertical dashed lines indicate the intervals specified in the AEU system (Table 1, Appendix A).

on the staff notation (which refers to 8 Hc as the interval).

A similar observation can be made on the fourth degree of maqam *Saba*, namely the note *hicaz*. We observe in Figure 9 that the fourth peak of the maqam *Saba* template at 19 Hc does not match the theoretical interval specified. In Özkan (1998), Çakar (2004) and Tulgan (2007) the corresponding explanation is: the fourth degree of maqam *Saba*, the *hicaz* note, is performed at higher pitch than indicated on the staff notation.

The mismatching intervals for the given maqams support the criticism. Other degrees' peaks match the theoretical intervals with 1/3 Hc precision. This example shows that the constructed histogram templates can be successfully used for theoretical studies. Due to space considerations, results from other maqams (which lead to the same conclusions) are not discussed here.

5. Discussions and conclusions

This paper presented a new method to align histograms of different recordings of Turkish maqam music. Such a method is extremely useful for automatic processing of multiple files which was not available until now in the domain of Turkish maqam music research. There are various applications for such automatic processing. To name a few: music information retrieval applications like automatic maqam classification and automatic transcription, and tuning research. For example, in studies aiming at detection of the whole set of pitches for maqam music, such a tool is potentially very useful. In Turkish maqam music theory, how many tones in an octave is needed to represent the maqam scale theory properly is still an open question. In the literature various propositions can be found varying from 17 to 79 tones in an octave (Yarman, 2007, 2008). With the method presented, large databases including maqam recordings with the same tonic can be aligned and viewed together with intervals specified in various theories. It is one of our future goals to study results from recordings of Tanburi Cemil Bey and compare in detail the histogram templates obtained with the intervals specified in several other theories like the Töre–Karadeniz system (Karadeniz, 1965) in comparison to the AEU system. Since average histograms carry information about commonly played pitches rather than details, they can be successfully used for automatic classification problems. In Gedik and Bozkurt (2008) it has been shown that high classification rates are obtained by simply using average histograms in a template matching framework for automatic classification. It can be misleading to process fine details of the average histograms and draw general conclusions or expect that such histograms reveal details about the nature of a given maqam.

Our proposed method involved a novel automatic tonic detection algorithm. The tonic detection algorithm was tested on 67 recordings of Tanburi Cemil Bey and 62 files' tonics were correctly identified. The tests can be considered demonstration of the potential at this level. The algorithm needs to be evaluated through testing on large databases including recordings with various acoustic conditions, various musicians and various forms. We leave such heavy testing to future work. It is clear that f0 estimation errors contribute to errors but the methodology we propose is independent of a specific f0 estimation algorithm. Therefore the contribution of f0 estimation errors in the average histogram computation are not studied in detail either. Efficient post-processing techniques for reduction of f0 estimation errors are also proposed in the first sections of this paper.

In addition, there are certain deficiencies of the study to be mentioned. Our approach discards the time and melodic context varying characteristics which are essential for maqam music. Despite these deficiencies, we think that the tools provided will be of high value for at least one of the most important dimensions of maqam music: the pitch scale theory.

Finally, most of the techniques proposed in this study are not limited to Turkish music. We think they are potentially applicable to other traditional (modal) music recordings from Pakistan, Afghanistan, Iran, Egypt, Morocco, Tunisia, Algeria, Azerbaijan, etc.

Acknowledgement

References

Akdoğu, O. (1989). *Taksim Nedir, Nasıl Yapılır?*. İzmir.

Akkoç, C. (2002). Non-deterministic scales used in traditional Turkish music. *Journal of New Music Research*, *31*(4), 285–293.

Arel, H.S. (1930). *Türk Musikisi Nazariyatı Dersleri*. İstanbul: Hüsnütabiat Matbaası (reprint: 1968).

Çakar, Ş.Ş. (2004). *Türk müziği teorisi ve makamlar*. Ankara: Milli Eğitim Bakanlığı Yayınları.

Cemil, M. (1947). *Tanburi Cemil'in Hayatı*. Ankara: Sakarya Yayınevi.

Çevikoğlu, T. (2007). Klasik Türk Müziğinin Bugünkü Sorunları. In *Proceedings from International Congress of Asian and North African Studies (İcanas 38')*, Ankara.

de Cheveigne, A. & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, *111*(4), 1917–1930.

Erguner, S. (2007). *Ney, 'metod'*. İstanbul.

Ezgi, S.Z. (1933). *Nazarî ve Amelî Türk Mûsıkîsi. Vol. I.* (pp. 8–29). İstanbul: Milli Mecmua Matbaası.

Gedik, A.C. & Bozkurt, B. (2008). Automatic classification of taksim recordings in Turkish makam music. In *Conference on Interdisciplinary Musicology*, 2–6 July, Thessaloniki, Greece.

Karadeniz, M.E. (1965). *Türk Musıkisinin Nazariye ve Esasları*. İstanbul: İş Bankası Yayınları.

Karaosmanoğlu, M.K. (2004). Türk musikisi perdelerini ölçüm, analiz ve test teknikleri. In *Proceedings from Yıldız Teknik Üniversitesi Müzik Konferansı*, İstanbul, Turkey.

Karaosmanoğlu, M.K. & Akkoç, C. (2003). Türk musiki-sinde icra – teori birliğini sağlama yolunda bir girişim. In *Proceedings from 10th Müzdak Symposium*, İstanbul, Turkey.

Özkan, İ.H. (1998). *Türk Musikisi Kuramsalı ve Usûlleri – Kudüm Velveleleri*. İstanbul: Ötüken Neşriyat.

Öztuna, Y. (2006). *Türk Musikisi: Akademik Klasik Türk San'at Musikisi'nin Ansiklopedik Sözlüğü, Vol. 2.* İstanbul: Orient Press.

Powers, H. (1988). First meeting of the ICTM Study Group on maqam. *Yearbook for Traditional Music*, *20*, 199–218.

Tanrıkorur, C. (2004). *Türk Müziği Kimliği*. İstanbul: Dergah Yayınları.

Touma, H.H. (1971). The maqam phenomenon: an im-provisation technique in the music of the Middle East. *Ethnomusicology*, *15*(1), 38–48.

Tulgan, Ö. (2007). Makam musiki perdelerinin sırrı, dedüktif bir deneme. *Müzik ve Bilim*, *7*, http://www.musikbilim.com/7m_2007/tulgan_o.html

Tzanetakis, G., Kapur, A., Schloss, W.A. & Wright, M. (2007). Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, *1*(2), 1–24.

Yarman, O. (2007). A comparative evaluation of pitch notations in Turkish makam music. *Journal of Inter-disciplinary Music Studies*, *1*(2), 43–61.

Yarman, O. (2008). 79-tone tuning & theory for Turkish maqam music. PhD thesis, İstanbul Technical University, Social Sciences Institute, İstanbul.

Yekta, R. (1922). *Türk Musikisi*. Transl. O. Nasuhioğlu. (pp. 6–16). İstanbul: Pan Yayıncılık (reprinted: 1986).

Zannos, I. (1990). Intonation in theory and practice of Greek and Turkish music. *Yearbook for Traditional Music*, *22*, 42–59.

Zeren, A. (2003). *Müzik sorunlarımız üzerine araştırmalar*. İstanbul: Pan Yayıncılık.

Appendix A: Pitch scale intervals in the Arel–Ezgi–Uzdilek (AEU)

The most commonly referred system for Turkish maqam music is the Arel–Ezgi–Uzdilek (AEU) system (Arel, 1930; Ezgi, 1933) which is considered as the "official theory". Most of the available theory and instrument method books use AEU as the basic system. It is shown

in Yarman (2007) that the AEU system is also equivalent to Yekta's (1922) system in terms of the intervals used. In the AEU system, the following basic intervals (in terms of Holdrian commas (Hc)) are used:

Bakiye (B):	4 Hc
Küçük Mücennep (S):	5 Hc
Büyük Mücennep (K):	8 Hc
Tanini (T):	9 Hc
Artık ikili (A):	12 or 13 Hc

The accidentals used in the AEU system are summarized in Figure 10. Each step corresponds to 1 Hc.

For each maqam, a scale is presented in Arel (1930) and the intervals are provided in terms of the basic intervals listed above. From the staff notation of the scale, the intervals can be deduced and vice versa. For example, the scale for maqam *Hüseyni* is specified with basic intervals: KSTTKST which corresponds to the following intervals in Hc with respect to the tonic: [8 13 22 31 39 44 53]. Figure 11 presents the scale for makam Hüseyni.

In this notation system, whole tones are represented as T (Tanini – 9 Hc) and half tones are represented as B (Bakiye – 4 Hc). Starting from the tonic *dügah* (A), the first interval is a whole tone lowered by a 1 Hc sized bemol. The first interval is then computed to be of size 8 Hc corresponding to K (Büyük Mücennep). The second interval is half tone plus 1 Hc, again due to the 1 Hc bemol of the first note, resulting in an interval size of 5 Hc, S (Küçük Mücennep).

It is often useful to display theoretical intervals together with the estimated f0 values to have an insight about frequency variations. In Figure 12 we present such an example.
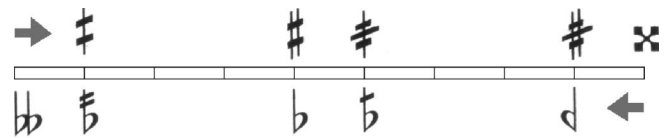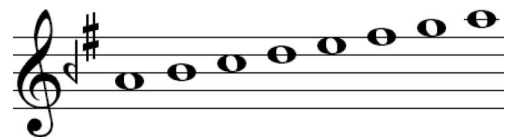

Fig. 10. Accidentals used in the AEU system.

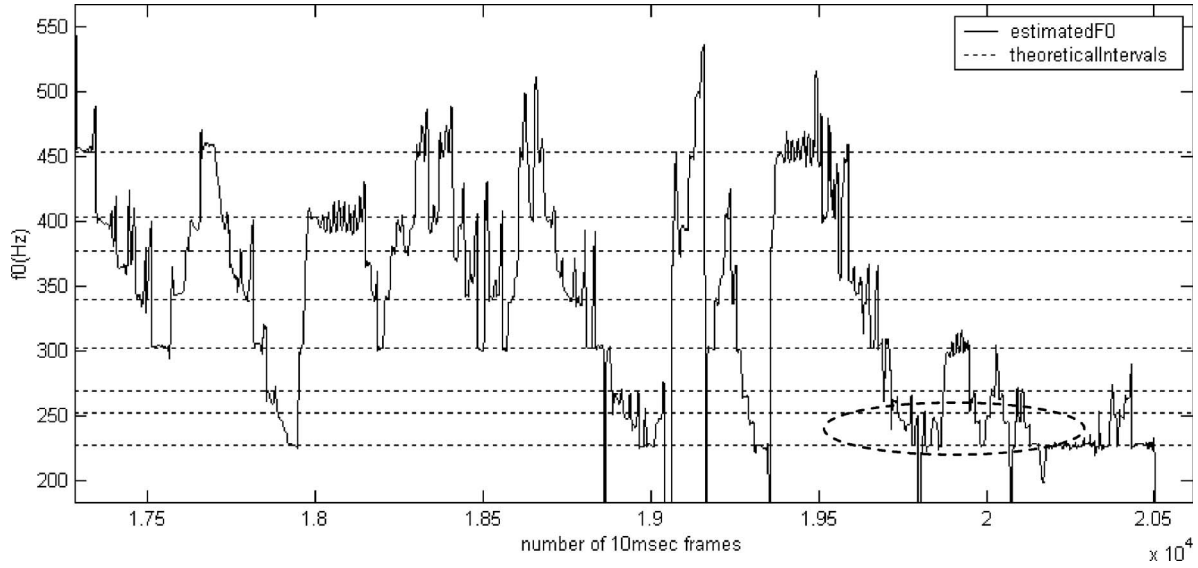Fig. 11. The scale for maqam Hüseyni according to the AEU system.

Fig. 12. An example pitch track for makam Hüseyni (Vol. 4, Track 4, *Hüseyni Taksim*, *end part*).

Table 1. Some maqam scale intervals (with dügah as tonic) in AEU.

Maqam	Intervals in Hc with respect to the tonic										
Hicaz	0	5			17	22	31	35		44		53
Rast	0		9		17	22	31		40		48	53
Segah	0	5		14		22	31	36			49	53
Kürdili Hicazkar	0	4		13		22	31	35		44		53
Hüzzam	0	5		14	19		31	36			49	53
Nihavend	0		9	13		22	31	35		44		53
Hüseyni	0		8	13		22	31		39	44		53
Uşşak	0		8	13		22	31	35		44		53
Saba	0		8	13	18		31	35		44		53

For maqam Hüseyni, there is a theory–application mismatch for the first interval (marked on the figure). A similar characteristic is observed throughout a recording. For this reason, matches and mismatches between theory and application can be observed on pitch histograms as Figure 6(a) (the centre of the second peak differs slightly for the theoretical template and the actual histogram). Therefore it is very convenient to study theory–application mismatches using pitch histograms. Actually, many studies base their claims on observations of pitch histograms (Akkoç, 2002; Zeren, 2003; Karaosmanoğlu & Akkoç, 2003; Karaosmanoğlu, 2004).

From these representations, one can deduce a list of intervals with respect to the tonic as presented in Table 1 for some commonly used maqams.

For automatic processing, using such tables as descriptors of maqam scales is very convenient. We utilize such tables in text format as complementary files in our implementations.

Appendix B: Audio material

For tests and plots, we have used 67 recordings as audio material from "Tanburi Cemil Bey: 1–5" Crossroads CD 4264, remastered from the Orfeon 10498 original 78 rpm record (1910–1914). From a total of 73 tracks, six tracks were excluded due to various reasons:

– the rotational speed variation in recording mechanism causing pitch shifts within a recording;
– piano being used as accompanying instrument;
– f0 estimation problems due to too high background noise.

The remaining material used for analysis contains performances in the following maqams:

Maqams with dügah as tonic (27 recordings):
Uşşak (6), Hüseyni (4), Saba (3), Gülizar (2), Tahir Buselik (2), İsfahan (2), Neva (1), Tahir (1), Bayati (1),

Kürdi (1), Şehnaz (1), Hicaz (1), Nişaburek (1), Muhayyer (1).

Maqams with rast as tonic (16 recordings):
Rast (2), Pesendide (1), Neveser (1), Nihavend (2), Mahur (3), Kürdili Hicazkar (2), Hicazkar (3), Suzinak (2).

Maqams with segah as tonic (5 recordings):
Segah (3), Müstear (1), Hüzzam (1).

Maqams with ırak as tonic (8 recordings):
Irak (1), Bestenigar (3), Eviç (2), Ferahnak (2).

Maqams with yegah as tonic (8 recordings):
Yegah (3), Ferahfeza (3), Şedaraban (2).

Maqams with hüseyni aşiran as tonic (2 recordings):
Nuhuft (1), Suzidil (1).

Maqams with acem aşiran as tonic (1 recording):
Şevkefza (1).

On these recordings Tanburi Cemil Bey plays kemence (an unfretted instrument), tanbur (movable fretted instrument), yaylı tanbur (tanbur played with a bow), lute and violin. It is acknowledged in various texts that Tanburi Cemil Bey used fewer number of frets (43 for two octave range whereas today's tanburs have more than 50 frets) and adjusted fret locations time to time before starting a taksim on a specific makam.

A few pre-processing operations were performed on recordings. Most of the recordings contain speech at the beginning as announcement of the content of recording. These portions of the recordings are silenced out. When necessary and possible, periodic noise is suppressed via frequency selective filters.