IRANIAN TRADITIONAL MUSIC DASTGAH CLASSIFICATION

SajjadAbdoli
Computer Department, Islamic Azad University,
Central Tehran Branch, Tehran, Iran
Saj.abdoli@gmail.com

ABSTRACT

In this study, a system for Iranian traditional music Dastgah classification is presented. Persian music is based upon a set of seven major Dastgahs. The Dastgah in Persian music is similar to western musical scales and also Maqams in Turkish and Arabic music. Fuzzy logic type 2 as the basic part of our system has been used for modeling the uncertainty of tuning the scale steps of each Dastgah. The method assumes each performed note as a Fuzzy Set (FS), so each musical piece is a set of FSs. The maximum similarity between this set and theoretical data indicates the desirable Dastgah. In this study, a collection of small-sized dataset for Persian music is also given. The results indicate that the system works accurately on the dataset.

1. INTRODUCTION

Music Information Retrieval (MIR) has grown in many fields but, there is still a significant gap between western and non-western, especially middle-eastern, MIR. As mentioned by Downie et al. [15], it is one of the most important challenges for the second decade of International Society of Music Information Retrieval (ISMIR) to expand its musical horizons to non–western music. To reduce this gap, we develop a system for Iranian traditional musical Dastgah classification.

The Dastgah concept in Persian music is similar to western musical scales and Maqams in Turkish and Arabic music. Middle-eastern music has not been considered in MIR studies largely, however, Gedik et al. [10] constructed a Turkish music Maqam recognition system based on the similarity between pitch histograms; and Heydarian et al. [16] described the Iranian musical Santur instrument and they also implemented an algorithm for the calculation of fundamental frequency.

In this paper, we introduce a Dastgah recognition system based on the similarity between Interval Type 2 Fuzzy Sets (IT2FSs). Fuzzy logic is also used by Bosteels et al. [17] for defining dynamic playlist generation heuristics. Sanghoon et al. [18] also used fuzzy logic in a music emotion recognition system. Leon et al. [19] also modeled musical notes by fuzzy logic to integrate music tuning theory and practice.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.
© 2011 International Society for Music Information Retrieval

After feature extraction, the proposed system assumes each performed note as an IT2FS, so each musical piece is a set of IT2FSs.The maximum similarity between this set and theoretical Dastgah prototypes, which are also sets of IT2FSs, indicates the desirable Dastgah. Gedik et al. [10] used the songs of the dataset to construct the patterns, whereas in this study, the system makes no assumption about the data except that different Dastgahs have different pitch intervals. Figure 1 shows the schematic diagram of the system. We also show that the system can recognize the Dastgah of the songs of the proposed dataset with overall accuracy of 85%.

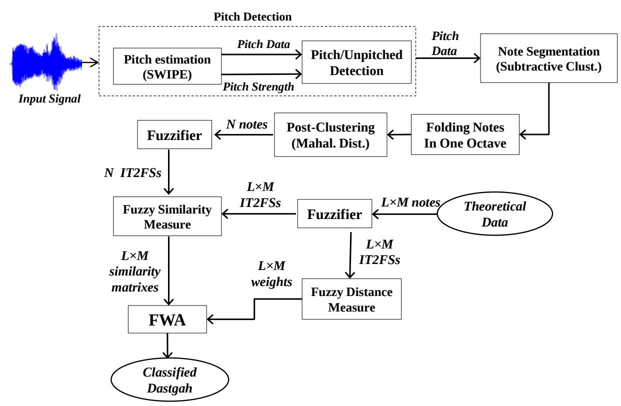

Figure 1. Dastgah classification system.

2. IRANIAN TRADITIONAL MUSIC

Persian music is a very old eastern music and has had outstanding impacts on other eastern musical cultures like Central Asia, Northern Africa, Southern Europe and also the countries around the Persian Gulf.

Iranian traditional music intervals consist of 24 equal Quartertones per each octave. This division first suggested by Vaziri [1]. He called half-sharp quartertone Sori and half-flat quartertone Koron. In practice, Sori and Koron are not exactly half-sharp or half-flat and can reside anywhere between two semitones.

Persian music is based on a set of seven major Dastgahs: Shur, Segah, Chahargah, Homayun, Mahur, Nava and Rast-panjgah. The Dastgah in Persian music is similar to the western musical scales (major and minor) and also Maqams in Turkish and Arabic music. Like western musical scales, Dastgah represents a specific pattern of the pitch ratios of successive notes. Each Dastgah consists of some partial melodies, called Gushe, which are created according to Dastgah patterns; however, some of them are not compatible to those patterns; therefore, their tuning might be different since they are used for moving from one Dastgah to another one (modulation) or for making the performance more pleasant, like Salmak Gushe in Shur Dastgah.

The arrangement of Gushes in each Dastgah during the performance is known as Radif which is presented by the masters of Persian music; such as Mahmud karimi's Radif for vocal or Mirza-abdollah's Radif for fret instruments.

For representing each Dastgah, we prefer the cent scale to tempered western intervals (note, half note, etc.). As it is mentioned, Sori and Korons can be resided anywhere between two half notes. Better results will be obtained if the cent scale is used rather than dividing the octave into equal divisions (12, 24 etc.). The scale steps of each Dastgah according to Karimi's Radif and Farhat [2] is shown in Table 1. Dastgahs like Mahur and Rast-panjgah, and also Nava and Shur have the same tuning.

Table 1. The scale steps for each Dastgah of Persian music.

Dastgah	Tuning Cents
1.Chahargah	(134,397,497,634,888,994,1200)
2.Homayun	(100,398,502,715,800,990,1200)
3.Mahur&Rst.	(208,397,497,702,891,994,1200)
4.Segah	(198,352,495,707,826,1013,1200)
5.Shur&Nava	(149,300,500,702,783,985,1200)

3. PITCH DETECTION

The proposed model for Iranian traditional music Dastgah recognition must be applicable on new and old songs. The majority of available old songs are converted to digital form from tape, so the white noise is an inseparable part of them, and we need a system to discriminate pitch form unpitched signals.

In order to do this, SWIPE′ algorithm [3] is used which can estimate the pitch and its strength at (discrete) time n as the spectral similarity between the signal (in the proximity of n) and a sawtooth waveform with missing non-prime harmonics and same (estimated) pitch as the signal. The pitch vector is refined and classified to pitch/unpitched clusters using the method was presented by Camachao [4]. It tracks the pitch strength trace of the signal and searches for clusters of pitch and unpitched sound according to the local maximization of the distance between the centroids.

The result of using SWIPE′ is shown in Figure 2. The pitches are retrieved from the vocal of Mahmud Karimi in Shur Dastgah. The system estimates the pitch of the signal at each 45 millisecond. The bold black circles are the pitch cluster centers which will be described in Section 4.1. Persian music is a center oriented music, as it shown in Figure 2 the vocalist starts with the Shahed (tonic) note, here about 180 Hz, and circulates around it during the performance and again backs to it.

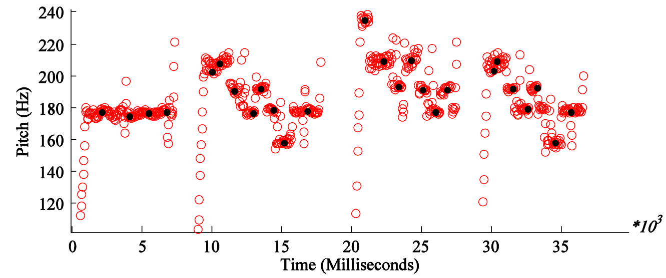

Figure 2. The pitches of vocal of Karimi in Shur mode. Circles are the pitches and the bold black circles are the pitch cluster centers.

4. PREPROCESSING

4.1 Note Segmentation

First of all our system needs to recognize which musical notes are used during the performance; moreover, it is needed to eliminate the wrong estimated pitches. A special situation may occur when we use vocal as our raw data. As it is shown in Figure 2 at the beginning of each note, it takes some milliseconds that the vocalist achieves the desirable frequency of voice and also at the end of each note we have some irrelevant points. To omit the redundant points, we need to use a clustering method to discriminate the notes form irrelevant data. Subtractive Clustering [5] is used.

This algorithm uses the data points, in time-frequency scale, as candidates for the centers of the clusters. Also, the number of clusters is not needed to be predefined. Since each point of data (X_i) is a candidate of clusters centers, a function for measuring the density in X_i is defined as

$$D_i = \sum_{j=1}^{n} exp\left(-\frac{\|X_i - X_j\|^2}{{r_a/2}^2}\right), \quad (1)$$

Where r_a is a positive constant representing a neighborhood radius, thus a data point with many neighboring data points will have a high potential value. After computing the potential value of every data point, we select the data point with the highest potential value as the cluster center. Let X_{c_1} to be the location of the first cluster center, then the potential of each data point (X_i) will be revised as

$$D_i = D_i - D_{c_1} exp\left(-\frac{\|X_i - X_{c_1}\|^2}{{r_b/2}^2}\right), \quad (2)$$

Where D_i and D_{c_1} is the potential value of X_i and the first cluster center, respectively and r_b is a positive constant which defines a neighborhood that has measurable reductions in density measure (typically $r_b = 1.5 r_a$). Thus we subtract the amount of potential value of each data point as a function of its distance from the first cluster center. After revising the density function, the next cluster center is selected as the point having the greatest density value. This process continues until $D_k < \varepsilon D_{c1}$ at the kth iteration, where ε is a small fraction. An algorithm is presented by Chiu [5] for finding the suitable amount of ε. Figure 2 shows the extracted pitch cluster centers. Note that, each pitch cluster center has two features (Time and Pitch). However, the pitch feature of each cluster center will be used in the next steps.

4.2 Folding Notes

It is convenient to fold all the extracted notes in one octave because the process of classification will be easier if we deal with one octave. The distance between A3 to A4 (220 Hz to 440 Hz) is selected. We fold the note f_i in the proposed octave by

$$FL(f_i) = \begin{cases} \dfrac{f_i}{2^{ceil\left(\log_2 \frac{f_i}{440}\right)}}, & f_i > 440 \\ f_i * 2^{ceil\left(\log_2 \frac{220}{f_i}\right)}, & f_i < 220 \\ f_i\,, & otherwise \end{cases} \quad (3)$$

After that, all the notes will be translated into cents with respect to 220 Hz. In order of brevity, it is not included here.

4.3 Post-Clustering

After folding notes in one octave, Mahalanobis distance [6] is applied to recognize which point on the reference octave corresponds to each musical note. Little et al. [7] also used this method for note segmentation of a query by humming system.

We find the distance between adjacent frames in the sequence using the Mahalanobis distance measure, Shown in Eq. (4). Given a frame p_i, we assume a new note has begun wherever the distance between two adjacent frames p_i and p_{i+1} exceeds a threshold, T

$$\sqrt[2]{(p_i - p_{i+1})M^{-1}(p_i - p_{i+1})} > \text{T} \rightarrow \text{new note} \quad (4)$$

Where the matrix M is a covariance matrix, which calculated from the variance within a rectangular window around the frame p_i as

$$M(p,p) = \frac{1}{2\tau}\sum_{k=i-\tau}^{i+\tau}(p_k - \overline{p})(p_k - \overline{p}), \quad (5)$$

Where τ is the size of a window surrounding the current frame and the average for p, $\overline{p}$ are calculated over this window.

The amount of T is set according to the quarter notes of Persian music, about 0.22, and a small window size for calculating the matrix $M(\tau = 4$ frames) is used. The result of this process is shown in Figure 3 which the performed notes of Hoseyni Gusheh in Shur mode based on Karimi's vocal are classified. The green thick lines and dashed red lines are the beginning and the end of each note, respectively.

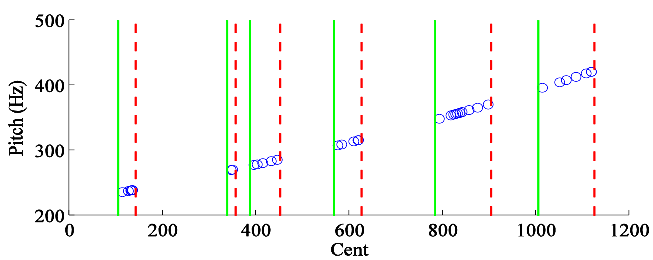

Figure 3. Clustering notes within one octave.

5. FUZZY LOGIC TYPE 2 AS DASTGAH CLASSIFIER

5.1 Interval Type 2 Fuzzy Sets

Type-2 fuzzy logic is an extension of type-1 fuzzy logic that first was introduced by Zadeh [8]. It can describe the uncertainty associated with our data when it is vague or incomplete, effectively. A special kind of type-2 fuzzy set, IT2FS, is used as the basic element of the classifier. IT2FSs include a secondary membership function to model the uncertainty of exact (crisp) type-1 fuzzy sets.[1]

An IT2FS in the universal set X, denoted as $\tilde{A}$, can be expressed as

$$\tilde{A} = \int_{x\in X}\mu_{\tilde{A}}(x)/x = \int_{x\in X}\left[\int_{x\in J_x} f_x(u)/u\right]/\,xJ_x \subseteq [0,1], \quad (6)$$

Where $f_x(u)$ is the secondary membership function and J_x is the primary membership of x which is the domain of the secondary membership function [9]. Figure 4 shows this region. The shaded region bounded by an upper and lower membership function is called the footprint of uncertainty (FOU). The FOU of $\tilde{A}$ can be expressed by the union of all the primary memberships as

$$FOU(\tilde{A}) = \cup_{\forall x\in X} J_x = \{(x,u): u \in J_x \subseteq [0,1]\}, \quad (7)$$

The upper membership function (UMF) and lower membership function (LMF) of $\tilde{A}$ are two type-1 Fuzzy Membership functions that bound the FOU. The UMF denoted by $\bar{\mu}_{\tilde{A}}(x)$ is associated with the upper bound of FOU, and the LMF denoted by $\underline{\mu}_{\tilde{A}}(x)$ is associated with the lower bound of FOU. They can be represented as

$$\bar{\mu}_{\tilde{A}}(x) \equiv \overline{FOU(\tilde{A})}\,\forall x \in X, \quad (8)$$

$$\underline{\mu}_{\tilde{A}}(x) \equiv \underline{FOU(\tilde{A})}\,\forall x \in X. \quad (9)$$

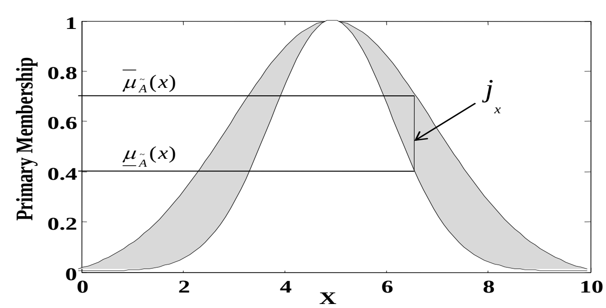

Figure 4. An Interval Type 2 Fuzzy set.

5.2 Fuzzifiers

5.2.1 Theoretical Data and the Data From Signal

We must manage the uncertainty associated with both each performed note and each note of the theoretical data. First; we must define a boundary for each note. We find it

[1]The membership value for ordinary fuzzy sets is a crisp number in [0,1].

convenient to use a region of about 67 cents for each note. Gedik et al. [10] also used this region for the widths of Gaussians of theoretical patterns for Turkish Maqams.

The mean of each segment, which are received from the post-clustering phase, is considered as a reference. Then, the upper bound ψ_k and the lower bound ϕ_k of kth frame in 67 cent scale are computed as

$$\psi_k = min((\sigma_k + 33.96), 1200), \qquad (10)$$

$$\phi_k = max\big((\sigma_k - 33.96), 1\big), \qquad (11)$$

$$\sigma_k = \frac{U_k - L_k}{2} + L_k, \qquad (12)$$

Where U_k and L_k are the beginning and the end of the kth segment, respectively.

5.2.2 Fuzzifing Upper and Lower Bounds

The upper and lower bounds of each note must be fuzzified in a [0,1] scale with a membership function. Considering one octave, there is a non-linier relation between the cent degree and frequency of each note that can be expressed as

$$f(x) = B * 2^{\frac{x-1200}{1200}}, \qquad (13)$$

Where x is the degree of cent of any note and B is the frequency of the final note of the proposed octave (e.g. 440 Hz). If we assign the membership value zero and one to the first and the last note, respectively Eq. (13) is rewritten as

$$f(x) = \frac{(A*2)*2^{\frac{x-1200}{1200}}}{A} - 1, \qquad (14)$$

Where A is the frequency of the first note of the proposed octave (e.g. 220 Hz). After simplification, Eq. (14) can be rewritten as

$$\overline{f}(\psi) = 2^{\frac{\psi}{1200}} - 1, \qquad (15)$$

$$\underline{f}(\phi) = 2^{\frac{\phi}{1200}} - 1. \qquad (16)$$

Where ψ and ϕ are the upper and lower bounds of any note, respectively. Both Eq. (15) and Eq. (16) can be considered as suitable type-1 fuzzy membership functions for fuzzifing musical notes. We call them Musical Fuzzy Membership Functions (MFMF).

5.2.3 Creating Footprint of Uncertainty

Two Gaussians are used for creating FOUs. Kreinovich et al. [11] also prove that Gaussian membership functions are the best choice for representing uncertainty in measurement. The constructed Gaussians are also mapped on MFMF to obtain more similarity degree between overlapped IT2FSs.

The UMF and LMF of the FOU for a note with a domain from ϕ to ψ are constructed as

$$\bar{\mu}_{\tilde{A}}(x,\psi) = \int_X e^{\frac{-(x-c)^2}{2\sigma_1^2}} dx * \overline{f}(\psi), \qquad (17)$$

$$\underline{\mu}_{\tilde{A}}(x,\phi) = \int_X e^{\frac{-(x-c)^2}{2\sigma_2^2}} dx * \underline{f}(\phi). \qquad (18)$$

Where $X = [\phi, \psi]$, c is the center of the $[\phi, \psi]$ boundary and ${\sigma_1}^2$and ${\sigma_2}^2$ are the standard deviations and $\overline{f}(\psi)$ and $\underline{f}(\phi)$ are the fuzzification functions for fuzzifing the upper and lower bounds of each note with MFMF, respectively. The pattern of Shur and Nava scale is shown in Figure 5.

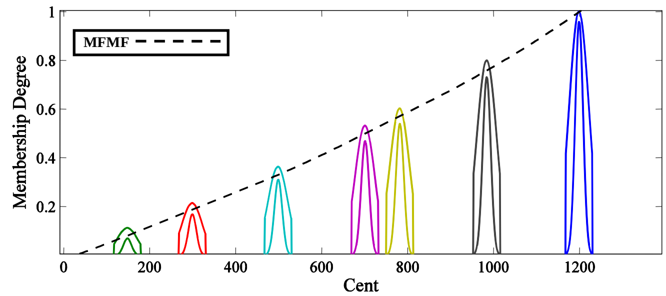

Figure 5. Shur Dastgah prototype that consists of seven IT2FSs which are mapped on MFMF (dashed line).

5.3 Fuzzy Similarity Measure

A suitable Fuzzy Similarity Measure (FSM) is used for computing the degree of similarity between prototypes and unknown patterns.

Basically, a robust FSM must satisfies four properties such as reflexivity, symmetry, transitivity and overlapping [13]. There are only six methods for computing the similarity between IT2FSs. Wu et al. [13] evaluated the six methods. Wu et al. [13] defined a new FSM, called Jaccard similarity measure (JSM), which satisfies the mentioned properties. It is also the fastest algorithm among the other FSMs [9]. It is used for our classifier and it can be defined as

$$\tilde{S}(\tilde{A},\tilde{B}) = \frac{\int_X min(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x))dx + \int_X min(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))dx}{\int_X max(\bar{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{B}}(x))dx + \int_X max(\underline{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{B}}(x))dx}, \qquad (19)$$

Where X is the domain of the data (here 1 to 1200).

5.4 Fuzzy Distance Measure

The distance between two IT2FSs are computed as

$$\tilde{D}(\tilde{A},\tilde{B}) = 1 - \tilde{S}(\tilde{A},\tilde{B}), \qquad (20)$$

Where $\tilde{S}(\tilde{A},\tilde{B})$ can be any FSM for IT2FSs [14].

The average distance between ith note (IT2FS) of any Dastgah prototype and the other notes from different Dastgahs is assigned as a weight to the ith note. This assignment helps to establish more discrimination between Dastgahs. It also indicates the degree of the uniqueness of each specific note. A constant weight (0.10) is assigned to the seventh and common note of each Dastgah. The assigned weight to each note is shown in Table 2.

5.5 Fuzzy Weighted Average

Fuzzy Weighted Average (FWA) is computed by Eq. (21). Mendel et al. [9] discussed about five different situations of the variables of Eq. (21) which make its computation different.

$$y = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}, \tag{21}$$

Where x_i and w_i are two crisp numbers, so Eq. (21) can be computed as simple as ordinary weighted average.

Table 2. The assigned weight to each step of Dastgah scales.

Dastgah	Weight Of Each Scale Step
1.Chahargah	(0.78,0.45,0.07,1.00,0.76,0.20,0.10)
2.Homayun	(0.90,0.45,0.15,0.47,0.74,0.23,0.10)
3.Mahur&Rst.	(0.80,0.45,0.07,0.36,0.77,0.20,0.10)
4.Segah	(0.21,0.16,0.89,0.62,0.17,0.51,0.10)
5.Shur&Nava	(0.77,0.96,0.10,0.36,0.80,0.30,0.10)

5.6 Dastgah Classification

Assume that $n \in \{1,2,\ldots,N\}$ IT2FSs are extracted from the input signal and also $m \in \{1,2,\ldots,\mathrm{M}\}$ IT2FSs for each Dastgah prototype is proposed. We also have $l \in \{1,2,\ldots,L\}$ Dastgahs. Assume that $Z^l_{m\times n} = \tilde{S}(\widetilde{M},\widetilde{N})$ is a similarity matrix between m IT2FSs of lth Dastgah prototype and n extracted IT2FSs from input signal where $\tilde{S}(\widetilde{M},\widetilde{N})$ can be any fuzzy similarity measure for $\widetilde{M}$ and $\widetilde{N}$. Let $\bar{Z}^l_m = max_n(Z^l_{m\times n})$ to be the maximum amount of each row of matrix $Z^l_{m\times n}$, then we may write the process of classifying or assigning, the unknown pattern to the Dastgah prototypes as

$$L^* = argmax_l(FWA_m(\bar{Z}^l_m, W^l_m)), \tag{22}$$

Where W^l_m is the assigned weight to each note (IT2FS) of each the Dastgah prototype.

6. RESULTS

6.1 Dataset

Lack of reliable dataset for Persian music was one of our main problems, so for evaluating the system a dataset for Iranian traditional music is collected. The dataset consists of 210 tracks from different Dastgah types. The Dastgah types and the number of recordings from each Dastgah type are as follows: 89-Shur & Nava, 30-Segah, 41-Mahur & Rast-panjgah, 26-Homayun and 24-Chahargah.

The collection was mainly based on vocal, and some monophonic musical pieces from some popular traditional instruments such as Santur, Tar, Setar and Kamancheh. The vocals were from three prominent Iranian vocalists such as Mahmud Karimi, 69 tracks, Abdullah Davami, 57 tracks, Muhammad Reza Shajarian, 20 tracks and also some other well trained vocalist. For a better evaluation, we also used 21 tracks from Arabian Maqams.[2]

6.2 Pattern Similarity

The Persian musical scales are so similar to each other and it is a considerable obstacle for Dastgah detection. Table 3 shows the degree of similarity between our Dastgah prototypes based on JSM for IT2FSs. The Chaharga, Mahur and Rast-panjgah modes have the maximum similarity degree, about 73%, while Chahargah and Segah modes have the minimum similarity degree, about 43%.

Table 3. The similarity degree between Dastgah prototypes.

Pattern Sim.%	A	B	C	D	E
A.Chahargah	100	59.22	73.30	43.07	49.28
B.Homayun	59.22	100	63.36	50.73	59.16
C.Mahur&Rst.	73.30	63.36	100	60.25	56.19
D.Segah	43.07	50.73	60.25	100	50.38
E.Shur&Nava	49.28	59.16	56.19	50.38	100

6.3 Evaluation

For system evaluation, both original and segmented songs of the dataset are used. We segment each song of our dataset to several portions with arbitrary lengths. By evaluating the system with the song segments, it is found that about one minute of any song is necessary and sufficient for Dastgah detection, so we can use only one minute of a given song to make the process of Dastgah detection faster.

The Dastgah recognition system can recognize the modes with overall accuracy of 85%. It is evaluated by computing the parameters such as Recall, Precision, Accuracy, F-measure and Matthews Correlation Coefficient (MCC). Table 4 shows the performance of the classifier according to above measures. The MCC is computed as

$$MCC = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}, \tag{23}$$

Where *TP*: True Positive, *TN*: True Negative, *FP*: False Positive and *FN*: False Negative.

The MCC is used as a measure of the quality of binary (two-class) classifications. It balances true and false positives and negatives. It can be used even if the classes are of very different sizes, like our dataset which the number of songs varies for each Dastgah. The MCC is also the best way for describing the confusion matrix. The confusion matrix of the classifier is presented in Table 5.

Table 4. The results of the evaluation of the classifier.

Dastgah	Recal	Precision	Acc.	F.mes.	MCC
Mahur&Rst.	90.24	90.24	96.19	90.24	0.87
Shur&Nava	85.39	98.70	93.33	91.56	0.86
Segah	83.33	83.33	95.23	83.33	0.80
Homayun	80.76	75.00	94.31	77.77	0.74
Chahargah	87.50	61.76	92.38	72.41	0.69

Moreover Receiver Operating Characteristic (ROC) space is shown in Figure 6. The ROC space is a graphical plot of the recall, or true positive rate (benefits), versus false positive rate (costs). The best possible prediction method would yield a point in the upper left corner or coordinate (0,100) and the

[2]Segah Maqam (Dastgah) is a common mode in Iranian and Arabian music. Ajam Maqam in Arabian music is also so similar to Iranian Chahargah scale.

worst prediction method is a point in the lower right corner or coordinate (100,0) of the ROC space, respectively.

Table 5. Confusion matrix.

Dastgah	A	B	C	D	E
A.Chahargah	0	1	1	1	0
B.Homayun	4	0	0	1	0
C.Mahur&Rst.	2	2	0	0	0
D.Segah	2	1	1	0	1
E.Shur&Nava	5	3	2	3	0

According to ROC space, confusion matrix and the mentioned measures, Dastgah recognition system is successful for the Dastgah types Mahur, Rast-Panjgah, Shur, Nava and Segah but, it is not very successful for the Dastgah types Homayun and Chahargah. The system can work precisely on small-sized dataset however, the dataset is needed to be expanded for a better evaluation of the system.

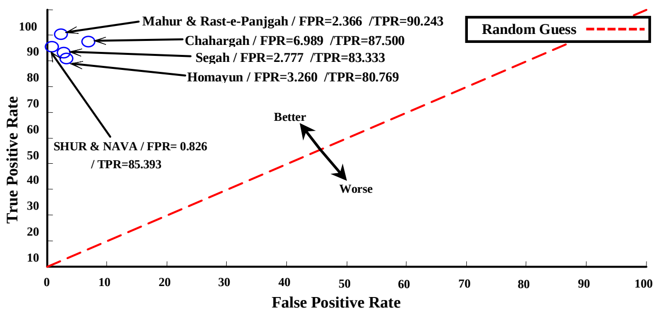

Figure 6.ROC space.

7. CONCLUSION

We presented a method for Iranian traditional music Dastgah classification. The method works by assuming each piece of music as a set of IT2FSs and recognize the Dastgah of the song by finding the maximum similarity between IT2FSs of the song and Dastgah prototypes. The system makes no assumption about the data except that the Dastgahs have different scale steps. The method was shown to work on small-sized dataset accurately.

Using Gaussian shaped FOUs; the Dastgah recognition system only supports intrauncertainty, which is the uncertainty a musician has about the scale steps. It is a candidate of future work to collect data from several musicians about tuning the scale steps of Dastgahs. Then use the method of liu et al. [12] for constructing the FOUs. After that, the system can support interuncertainty, which is the uncertainty that a group of musicians have about the scale steps [9]. Moreover, in the future work, the system must be equipped with a Gushe (or melody) recognition system.

8. REFERENCES

[1] Vaziri, A. N.: *Dastur-e Tar*, Tehran,1913.

[2] Farhat, H.: *The Dastgah Concept in Persian Music*, Cambridge University Press, 1990.

[3] Camacho A., Harris, J. G.: "A sawtooth waveform inspired pitch estimator for speech and music," *Journal of the Acoustical Society of America*, vol. 124, pp. 1638-1652, 2008.

[4] Camacho A.: "Detection of Pitched/Unpitched Sound Using Pitch Strength Clustering," *In Proceedings of ISMIR*, pp. 533-537, 2008.

[5] Chiu, S.:"Fuzzy Model Identification Based on Cluster Estimation," *Journal of Intelligent & Fuzzy Systems*, Vol. 2, No. 3, pp.267-278, 1994.

[6] Mahalanobis P. C.: "On the generalized distance in statistics," *Proceedings of the National Institute of Science of India*, pp. 49-55, 1936.

[7] Little D., Raffensperger D., Pardo B.: "A Query by Humming System that Learns from Experience," *In Proceedings of ISMIR*, pp. 335-338, 2007.

[8] Zadeh L.A.: "The Concept of a Linguistic Variable and its Application to Approximate Reasoning I," *Information Sciences*, Vol.8, No.3, pp.199-249, 1975.

[9] Mendel J. M., Wu D.: *Perpetual Computing*, John Wiley & IEEE Press, 2010.

[10] Gedik A., Bozkurt B.: "Pitch-frequency histogram-based music information retrieval for Turkish music," *Signal Processing*, Vol.90, No.4, pp.1049-1063, 2010.

[11] Kreinovich V., Quintana C., Reznik L.: "Gaussian membership functions are most adequate in representing uncertainty in measurements," *In Proceedings of NAFIPS'92*, pp.618-24, 1999.

[12] Liu F., Mendel J. M.: "Encoding Words into Interval Type-2 Fuzzy Sets Using an Interval Approach," *IEEE Trans. on Fuzzy Systems*, vol. 16, pp. 1503-1521, 2008.

[13] Wu D., Mendel J. M.: "A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets," *Information Sciences*, vol. 179, pp. 1169-1192, 2009.

[14] Zhang H., Zhang W.: "Hybrid monotonic inclusion measure and its use in measuring similarity and distance between fuzzy sets," *Fuzzy Sets and Systems*, Vol.160, No.1, pp.107-118, 2009,

[15] Dowine J. S., Byrd D., Crawford T.: "Ten years of ISMIR: reflections on challenges and opportunities," *In Proceedings of ISMIR*, pp. 13-18, 2009.

[16] Heydarian P., Reiss J. D.:"The Persian music and the Santur instrument," *In Proceedings of ISMIR*, pp. 524-527, 2005.

[17] Bosteels K., Pampalk E., Kerre E.: "Evaluating and analysing dynamic playlist generation heuristic using radio logs and fuzzy set theory," *In Proceedings of ISMIR*, 2009.

[18] Sanghoon J., et al.: "A fuzzy inference-based music emotion recognition system," *In Proceedings of VIE*, pp.673-677, 2008.

[19] Leon T., Liern V.: "Fuzzy logic helps to integrate music theory and practice," *In Proceedings of FUZZ-IEEE*, pp.1-5, 2010.