Knowledge Management On The Semantic Web: A Comparison of Neuro-Fuzzy and Multi-Layer Perceptron Methods For Automatic Music Tagging

Sefki Kolozali, Mathieu Barthet, and Mark Sandler

Centre for Digital Music
Queen Mary University of London
{sefki.kolozali,mathieu.barthet,mark.sandler}@eecs.qmul.ac.uk

Abstract. This paper presents the preliminary analyses towards the development of a formal method for generating autonomous, dynamic ontology systems in the context of web-based audio signals applications. In the music domain, social tags have become important components of database management, recommender systems, and song similarity engines. In this study, we map the audio similarity features from the Isophone database [25] to social tags collected from the Last.fm online music streaming service, by using neuro-fuzzy (NF) and multi-layer perceptron (MLP) neural networks. The algorithms were tested on a large-scale dataset (Isophone) including more than 40 000 songs from 10 different musical genres. The classification experiments were conducted for a large number of tags (32) related to genre, instrumentation, mood, geographic location, and time-period. The neuro-fuzzy approach increased the overall F-measure by 25 percentage points in comparison with the traditional MLP approach. This highlights the interest of neuro-fuzzy systems which have been rarely used in music information retrieval so far, whereas they have been interestingly applied to classification tasks in other domains such as image retrieval and affective computing.

1 Introduction

In the last decade, there has been extensive research on the development and use of the semantic web to make the web data interpretable, usable and accessible across a wide variety of domains. The key idea of this effort is to provide web content with conceptual background which is referred to as ontologies. For this purpose, the data models, such as ontology web language (OWL) and resource description format (RDF) have received considerable attention from researchers and the industrial sectors. Many research groups built ontologies manually to represent different types of data (e.g. music data, social data) within the formation of the semantic web [1]. Some examples of ontologies in the music domain are the music ontology[1] (MO) and the music performance ontology, grounded in the MO [22].

[1] http://musicontology.com/

9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)
19-22 June 2012, Queen Mary University of London

The main disadvantage of the current ontology engineering process is that it cannot operate independently from human supervision. There is a growing interest for automated learning systems which can handle knowledge acquisition and also build ontologies from fast growing and large datasets [3], since current ontologies have an inflexible structure, and are incapable of handling these problems.

Social tags represent a potential high-volume source of descriptive metadata for music. Tags are useful text-based labels that encode semantic information about the music content (e.g. genres, instrumentations, geographic origins, emotions). In the music domain, popular web systems such as Last.fm[2] provide possibility for users to tag with free text labels an item of interest. Such metadata can either be used to train audio content-based classification systems for semantic annotation and retrieval, or likewise, automatic ontology generation. There has been recently a significant amount of research on content-based music similarity and tagging systems. Both fields use content-based descriptors extracted from audio signals. The Isophone dataset [25] provides an excellent opportunity to undertake reproducible research on large-scale music collection with readily-available mel-frequency cepstral coefficient (MFCC) features that can be jointly used with other datasets.

In this paper, we propose an audio tagging system based on neuro-fuzzy (NF) neural networks in comparison with the more traditional multi-layer perceptron (MLP) algorithm. The system was tested using the Isophone database in conjunction with Last.fm social tags. The use of neuro-fuzzy systems is driven here for further linking it with fuzzy spatial reasoning as an ontology generation solution. Hence we are motivated here by the comparison of the performance of NF networks relatively to another classifier, rather than by the obtention of state-of-the-art classification accuracies. Neuro-fuzzy systems have only been scarcely used in MIR (e.g. [29]) whereas they have shown to be powerful in other domains, such as image retrieval [23] and affective computing [10].

The remainder of this paper is organized as follows; in the next section, previous works related to automatic ontology generation are described. Section 3 explains the automatic tagging system and algorithms used in this work. Section 4 presents the experiments and results. Finally, in the last section, the paper concludes on the importance of the current research problem, and presents the next steps in our research.

2 Related Work

Although there are many ways of collecting experimental data for music information retrieval (MIR) research, the main challenges are the sparsity of the data, and the bias introduced by erroneous annotations. Besides, the cognitive processes underlying the representation and categorization of music are not yet fully understood, and it is often difficult to know what makes a tag accurate and what kinds of inaccuracies are tolerable [12, 9].

[2] www.last.fm

Last.fm is a popular online streaming service and social network which provides metadata assigned to songs or artists by users through an application programming interface (API). Social network users usually prefer to use the most frequent tags rather than by entering new tags in the system. Therefore, the obtained metadata may suffer from a popularity bias.

The most used classification systems for audio tagging are standard binary classifiers such as support vector machines (SVMs) and AdaBoost [26]. As supervised techniques, these classifiers rely on a training and a testing stage. Thereby, the classifier is engaged in predicting the musical tags of a testing dataset. Gaussian mixture model (GMM) is another well known technique that has been widely used in music tag prediction. The approach has shown to provide good semantic annotations for an acoustically diverse set of songs and retrieved relevant songs given a text-based query in [27]. In many studies, a time series of mel-frequency cepstral coefficient (MFCC) vectors are used as a music feature representation. MFCCs are a general purpose measure of the smoothed spectrum of an audio signal which primarily represent the timbral aspects of the sound. Although MFCCs are based on a simple auditory model and are common in the music and speech recognition world [5, 2]. The multi-layer perceptron (MLP) is one of the most commonly used neural networks. It can be used for classification problems, model construction, series forecasting and discrete control. For the forecasting problems, a backpropagation (BP) algorithm is normally used to train the MLP Neural Network [20, 19]. Since the MPL is very common in many research fields, and that neuro-fuzzy neural networks are based on the same learning framework, we have used this algorithm in our experiments, for comparison.

Parallel to this, there are ontologies in use today focusing on cases such as the classification of musical instruments [15]. For such sets of data, the primary organizational structure often involves spatial relationships; for example, object A connects to object B, object B is part of object A, object C is externally connected object B, object C is part of object A. One formalization of spatial relationships for the purpose of qualitative reasoning in ontological models is provided by Coalter and Leopold, in [4]. Fuzzy spatial reasoning is based on spatial relationships that provides a framework for modeling spatial relations in the fuzzy-set theory [24, 17, 6].

3 Audio Tagging System

The general architecture of the proposed audio tagging system is shown in Figure 1 and presented in the sections below.

3.1 Data Acquisition

For the data acquisition, two large databases were used: *i*) the Isophone database[3], [25] and *ii*) the Last.fm database. The Isophone database is based on the SoundBite plugin [16], which is available as iTunes and Songbird[4] plugins. The Sound-

[3] http://www.isophonics.net/
[4] `http://getsongbird.com/`

Bite plugin extracts features (MFCCs) from the entire user audio collection and stores them for further similarity calculations. The extracted features are also uploaded to a central server and expand dynamically the Isophone database.

The Isophone database uses MusicBrainz[5] identifiers as a source for unique identifiers. MusicBrainz is a comprehensive public community music metadata service. It can be used to identify songs or CDs, and provides valuable data about tracks, albums, artists and other related information. In order to associate the Isophone database to the MusicBrainz dataset, the GNAT[6] application is used, which implements a variant of the automated inter linking algorithm. In the metadata (tags) filtering process, MusicBrainz IDs of the tracks included in the Isophone database are matched against those of the Last.fm database by using Last.fm's AP. The collected tags were sorted out by their frequency of appearance within the Isophone database.

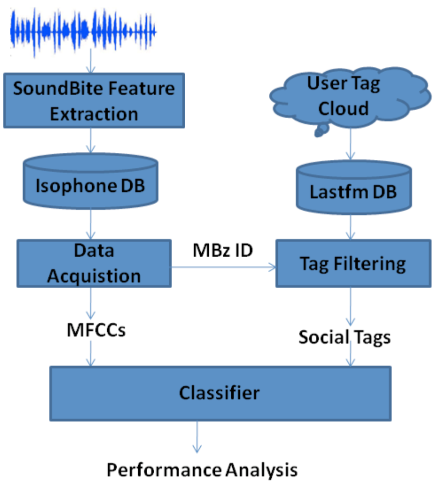

Fig. 1. Audio Tagging System

3.2 Classifiers

The classification is performed by using multi-layer perceptron and neuro-fuzzy systems which are supervised methods. Our goal is to associate an audio signal with various labels from a priori defined tag sets.

Multi-Layer Perceptron Neural Networks have been used in many different areas to solve pattern recognition problems. The multi-layer perceptron

[5] http://musicbrainz.org/
[6] http://www.sourceforge.net/projects/motools/

(MLP)[21] is one of the most common Neural Networks in use. It consists of two main computational stages: a feed-forward network and a backpropagation network. In the forward pass, input vectors are applied to the input nodes of the network, and at each node (neuron), the weighted sum of the input is computed. In the final stage of the forward pass, the set of outputs is produced as the actual output of the network. During the backward pass, the actual output of the network is subtracted from a desired output to produce an error signal, and the network weights are adjusted to move to the desired response according to the errors that are propagated backwards through the network. Fig. 2 shows the architecture of the Multi-Layer Perceptron used for deriving music tagging outputs from MFCCs.

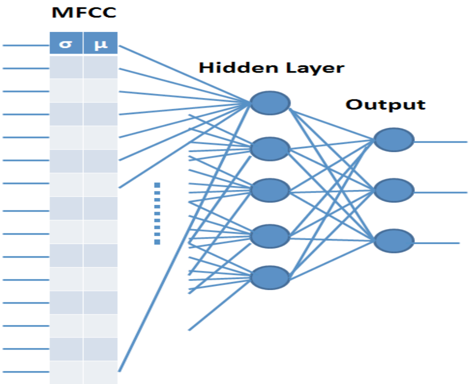

Fig. 2. Multi-Layer Perceptron for Music Tagging. σ and μ represent the variance and mean of the MFCCs time series, respectively

Neuro-Fuzzy Neuro-fuzzy (NF) systems [11] are a combination of neural networks and fuzzy logic [14] that merge the learning ability of neural networks and the reasoning ability of fuzzy logic. Automatic linguistic rule extraction is a typical application of neuro-fuzzy when there is little or no prior knowledge about the process. Figure 3 shows the architecture of a Neuro-Fuzzy network with two inputs and one output.

Considering the fuzzy sets of MFCC coefficients, the following linguistic rule set illustrates a simple fuzzy reasoning process. The MFCC coefficients are defined as the input variables, denoted $x_{1,1}, x_{1,2}, ...x_{i,j}$, where i and j refer to the rules and fuzzy sets, respectively. The rules can be expresses as follows:

$$\text{Rule } 1: \overbrace{If\ x_{1,1} \text{ is } M_{1,1} \text{ and } x_{1,j} \text{ is } M_{1,j}}^{antecedent}\ \overbrace{\text{then } y_1 \text{ is } y_d}^{consequent}$$

$$\vdots$$

$$\text{Rule } i: If\ x_{1,1} \text{ is } M_{i,1} \text{ and } x_{i,j} \text{ is } M_{i,j} \quad \text{then } y_i \text{ is } y_d$$

where M represents the fuzzy sets for the MFCC coefficients and y_d is the desired output provided based on music tags. In the fuzzification process, we used triangular symmetric membership functions. By acting on the parameters of the triangular membership functions, denoted a_{ij} and b_{ij}, it is possible to generate different types of functions (e.g. low, medium, high). Corresponding parameters of the membership function is defined below in Eq.1. Once the rules are determined, the inputs are fuzzified to obtain a membership degree, $\mu_{i,j}$, for each membership function of fuzzy sets, as follows:

$$\mu_{i,j} = \begin{cases} 1 - \dfrac{2 \mid x_j - a_{i,j} \mid}{b_j} & , a_{i,j} - \dfrac{b_{i,j}}{2} < x_j < a_{i,j} + \dfrac{b_{i,j}}{2} \\ 0 & , \text{otherwise} \end{cases} \tag{1}$$

Next, each satisfied fuzzy set's membership degree is used as an input to the fuzzy reasoning process which performs T-norm product operation. The consequent of a fuzzy rule assigns the entire rule to the output fuzzy set which is represented by a membership function that is chosen to indicate the related music tag. In the next layer the firing strengths of each rule are normalised. The normalised consequent fuzzy sets encompass many outputs, so it must be resolved into a single output value by a defuzzification method. In the defuzzification stage, the fuzzy sets which represent the outputs of each rule are combined into a single fuzzy set and distill a single output value from the set. The centre of gravity method which is one of the most popular defuzzification method has been used in the proposed approach to resolve the aggregated fuzzy set.

There are three types of parameters to be adapted in the learning stage which determine the parameter vector z:

$$z = (a_{11}, ..., a_{ij}, b_{11}, ..., b_{ij}, w_1, ..., w_i) \tag{2}$$

where a_{ij}, b_{ij} are the MFCC membership functions and w_i is the weight parameter that is used to tune the membership functions. The learning stage of the neuro-fuzzy approach uses neural nets learning system by optimising a criterion function (V) given by:

$$\nabla_z V = \left[\frac{\partial V}{\partial z_1}, ..., \frac{\partial V}{\partial z_i} \right] \tag{3}$$

where $-\nabla_z V$ is the gradient of V with respect to z. In order to tune the fuzzy set parameters, the weights and membership function's parameters need to be adjusted so as to minimize the error. Eq. (4) shows how to apply the method of stochastic approximation on the criterion loss function to identify the parameters of the system. It is an iterative procedure given by:

$$z(t+1) = z(t) - \eta \nabla_z V[z(t)] \tag{4}$$

where z is the vector parameters to adapt and η is the predefined learning rate constant which specifies the computation speed of the learning task.

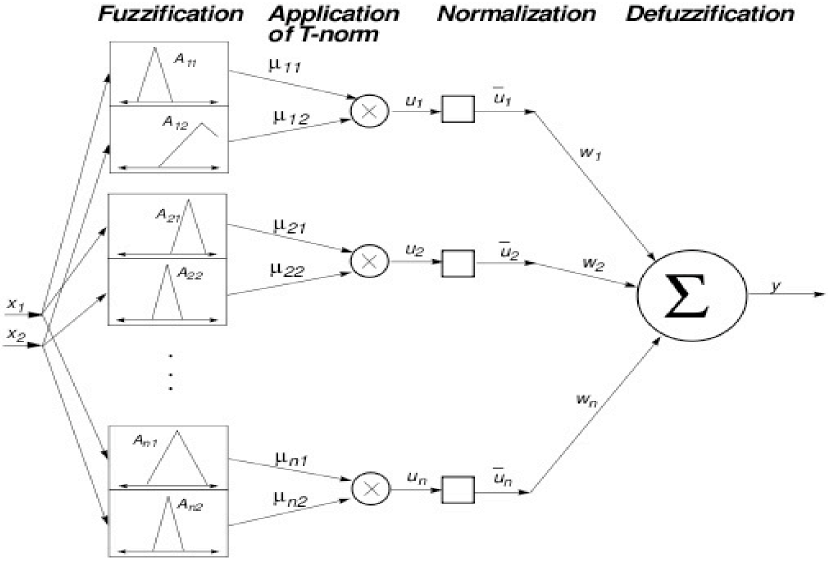

Fig. 3. Neuro-fuzzy system architecture (based on [7])

4 Experiments

Both of the neuro-fuzzy (NF) system and the multi-layer perceptron (MLP) neural network are based on the same network topologies and they were designed with multi-network system.

4.1 Dataset

The experimental dataset is a merge of Last.fm social tags for the Isophone database. In the experiments, 41 962 songs have been used out of 152 410 songs of the Isophone database. For each track we collected tags related to the five following categories: genre, mood, instrumentation, locale, and time-period. By summing up the subclasses associated with these tag categories, 32 tag subclasses were considered in total (e.g. pop, chillout, guitar, american, 90s). For each given tag, 50% of the associated tracks were used for training, and 50% were used for testing. The repartition of tracks according to the various types of tags is given in Table 1. For each track, an audio feature vector of 40 values representing the mean and variances of 20 MFCCs is computed, as in [25].

Genre	Data %	Instrumentation	Data %	Mood	Data %	Locale	%	Time-Period	Data %
Pop	38.52	Electronic	11.51	Dance	7.75	American	20.69	00s	14.67
Alter. Rock	26.45	Acoustic	11.48	Relax	6.14	French	1.92	90s	20.91
Classic Rock	25.70	Guitar	9.20	Fun	4.81	Swedish	1.10	80s	15.22
Electronica	12.18	Piano	10.66	Melancholic	17.40			70s	14.55
Punk	13.92	Vocal	10.14	Party	13.46			60s	10.20
Hard Rock	13.70			Romantic	14.32				
Jazz	13.74			Atmospheric	7.77				
Blues	12.70								
Ambient	9.41								
Trip Hop	5.35								
Soul	10.30								
Metal	11.00								
Total	**88.13**		**36.87**		**51.13**		**23.65**		**57.89**

Table 1. Repartition of tracks in the experimental data set according to genre, instrumentation, mood, locale, and time-period

4.2 Analysis parameters

The number of iterations in the neuro-fuzzy and MLP algorithms were identified according to the lowest point on the mean square error curves obtained in the training stage. The best learning rate ($\eta = 0.6$) was determined empirically. For each tag, the structure of the MLP consisted of 40 input nodes, 20 hidden nodes, and 1 output node. In calculating the hidden and output units of the MLP the *tanh* function was used as the activation function. In the neuro-fuzzy system each network was created with the 40 inputs and 1 output rule set. Three membership functions have been used for each fuzzy set (low, medium, and high). Both algorithms comprised 32 different networks in total.

4.3 Results

In order to evaluate the performance of these algorithms, standard evaluation metrics (precision [P], recall [R], F-measure [F]) have been used [18].

The results are shown in Table 2. On overall, the neuro-fuzzy system achieved an F-measure of 46% in the identification of a large number of music tags (32). The multi-layer perceptron's overall F-measure was 21% that is lower by 25% points in comparison with that of the NF method. The better results obtained for the labels "vocal", "melancholic", "metal", "classic rock", and "60s". The labels "party", "atmospheric", "romantic", "fun" obtained the lowest performance in this experiment. This is probably due to the fact that other factors than timbre (as modeled by the MFCCs) are involved to characterise these genres and emotion-eliciting situations (e.g. rhythm for party music is deemed to be very important). The results indicated that neuro-fuzzy systems performed much better than the multi-layer perceptron on large-scale experiments.

		P		R		F	
		NF	MLP	NF	MLP	NF	MLP
Genre	Pop	**0.66**	0.57	**0.52**	0.46	**0.58**	0.51
	Alter. Rock	**0.65**	0.55	**0.51**	0.32	**0.57**	0.41
	Classic Rock	**0.70**	0.58	**0.54**	0.32	**0.61**	0.41
	Electronica	**0.64**	0.57	**0.41**	0.22	**0.50**	0.31
	Punk	**0.62**	**0.62**	**0.35**	0.29	**0.45**	0.39
	Hard Rock	**0.68**	0.54	**0.48**	0.20	**0.56**	0.29
	Jazz	0.67	**0.73**	**0.41**	0.34	**0.51**	0.46
	Blues	**0.62**	0.45	**0.34**	0.08	**0.44**	0.14
	Ambient	**0.62**	0.49	**0.29**	0.19	**0.40**	0.27
	Trip Hop	**0.67**	0.40	**0.36**	0.04	**0.47**	0.06
	Soul	**0.64**	0.45	**0.36**	0.13	**0.46**	0.21
	Metal	**0.73**	0.61	**0.57**	0.31	**0.64**	0.41
	Average	**0.65**	**0.54**	**0.42**	**0.24**	**0.51**	0.32
Instrumentation	Electronic	**0.64**	0.36	**0.44**	0.07	**0.52**	0.11
	Acoustic	**0.53**	0.46	**0.23**	0.10	**0.32**	0.17
	Guitar	**0.54**	0.32	**0.24**	0.06	**0.33**	0.11
	Piano	**0.56**	0.55	**0.20**	0.02	**0.29**	0.04
	Vocal	**1.00**	0.43	**1.00**	0.04	**1.00**	0.07
	Average	**0.65**	0.42	**0.42**	0.05	**0.49**	0.10
Mood	Dance	**0.53**	0.31	**0.20**	0.04	**0.30**	0.07
	Relax	**0.51**	0.39	**0.14**	0.03	**0.22**	0.05
	Fun	0.31	**0.36**	**0.07**	0.01	**0.12**	0.02
	Melancholic	**1.00**	0.64	**1.00**	0.32	**1.00**	0.42
	Party	0.21	**0.53**	0.02	**0.18**	0.04	**0.27**
	Romantic	0.34	**0.44**	**0.05**	0.03	**0.08**	0.06
	Atmospheric	0.37	**0.45**	0.07	**0.11**	0.11	**0.17**
	Average	**0.46**	0.44	**0.22**	0.10	**0.26**	0.15
Locale	American	**0.58**	0.42	**0.36**	0.06	**0.44**	0.10
	French	**0.67**	0.15	**0.40**	0.04	**0.50**	0.06
	Swedish	**0.64**	0.26	**0.47**	0.09	**0.54**	0.13
	Average	**0.63**	0.27	**0.41**	0.06	**0.49**	0.09
Time-Period	00s	**0.56**	0.45	**0.30**	0.11	**0.39**	0.18
	90s	**0.63**	0.44	**0.45**	0.11	**0.52**	0.17
	80s	**0.65**	0.52	**0.43**	0.14	**0.52**	0.23
	70s	**0.63**	0.50	**0.45**	0.10	**0.53**	0.17
	60s	**0.72**	0.56	**0.56**	0.12	**0.63**	0.20
	Average	**0.63**	0.49	**0.43**	0.11	**0.51**	0.19
Overall		**0.61**	0.47	**0.38**	0.14	**0.45**	0.20

Table 2. Performance of the neuro-fuzzy (NF) system and multi-layer perceptron (MPL) network in the classification of five music tag classes: genre, instrumentation, mood, locale, and time-period

5 Discussion

Reasonably good performance were obtained for the neuro-fuzzy system in the case of genre, time period, and location, considering the large number of classes (32) in these experiments. However the results were poor for the mood and instrumentation labels showing the need to refine the features and/or classification framework. Research on music emotion recognition has shown that the regression approach applied to arousal/valence values outperformed the classification approach applied to categorical labels [13]. Research on polyphonic musical instrument recognition is still in its early days [8], and it is not surprising to obtain low recognition accuracy for the instrumentation since the MFCCs only capture the timbre of the music at a “macro” level (globally). It should also be noted that label inaccuracies in the social data may have affected the results for both classifiers. However as previously mentioned the main goal of the study was to compare the relative performance of the NF and MLP methods with regards to the promising application of NF systems in automatic ontology generation.

Our study provides a framework for future studies to assess systems using the Isophone dataset. Although no means are offered for automatically extracting and proposing axioms to ontology engineering in this study, future work will investigate the identifications of the relationships between different conceptual entities as in [4]. As an example of the future use of ontologies on music annotation systems, it is also worth to mention a recent study proposed by Wang et al.[28] in which an ontology-based semantic reasoning is used to bridge content-based information with web-based resources. The authors pointed out that the proposed ontology-based system outperformed content-based methods and significantly enhanced the mood prediction accuracy.

6 Conclusion

Our research is motivated by the fact that, current ontology designs have inflexible structure and have not been used with any automated learning system which leads to a danger to fossilise the current knowledge representation by static ontologies. Preliminary analyses were conducted with a neuro-fuzzy (NF) system and a multi-layer perceptron (MLP) neural network in a music-tag annotation task. The results showed that NF outperformed MLP by 25% points in F-measure, which indicated that fuzzy systems are promising classifiers for audio content-based ontology construction. In our future work, our study will continue towards the automatic ontology generation by using fuzzy spatial reasoning systems.

References

1. G. Antoniou and F. can Harmelen. *Semantic Web Premier 2nd Edition.* Massachusetts Institute of Technology, 2008.

2. J.-J. Aucouturier and F. Pachet. Music similarity measures: Whats the use? *Proceedings of the 3rd International Symposium on Music Information Retrieval*, pages 157–163, 2002.
3. D. S. C. Catton. The use of named graphs to enable ontology evolution. *W3C Workshop on the Semantic Web for Life Sciences*, 2004.
4. A. Coalter and J. L. Leopold. Automated ontology generation using spatial reasoning. *KSEM'10 Proceedings of the 4th international conference on Knowledge science, engineering and management.*, pages 482–493, 2010.
5. D. P. W. Ellis. Classifying music audio with timbral and chroma features. *The International Society of Music Information Retrieval*, 2007.
6. A. Esterline, G. Dozier, and A. Homaifar. Fuzzy spatial reasoning. *Fuzzy spatial reasoning. In:Proc. of the 1997 Int. Fuzzy Systems Association Conference*, 1997.
7. J. Godjevac. Comparative study of fuzzy control, neural network control and neuro-fuzzy control. *Technical Report*, 1995.
8. P. Herrera-Boyer, A. Klapuri, and M. Davy. Automatic classification of pitched musical instrument sounds. In A. Klapuri and M. Davy, editors, *Signal Processing Methods for Music Transcription*, pages 163–200. Springer, 2006.
9. M. D. Hoffman, D. M. Blei, and P. R. Cook. Easy as cba: A simple probabilistic model for tagging music. In *Proceedings of the 10th International Society for Music Information Retrieval Conference*. International Society for Music Information Retrieval, October 2009.
10. S. V. Ioannou, A. T. Raouzaiou, V. A. Tzouvaras, T. P. Mailis, K. C. Karpouzis, and S. D. Kollias. Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, 18:423–435, 2005.
11. E. M. Jyh-Shing Roger Jang, Chuen-Tsai Sun. *Neuro-Fuzzy and Soft Computing.* Prentice-Hall, Inc, 1997.
12. J. H. Kim, B. Tomasik, and D. Turnbull. Using artist similarity to propagate semantic using artist similarity to propagate semantic information. In *10th International Society for Music Information Retrieval Conference*, 2009.
13. Y. E. Kim, E. M. Schmidt, R. Migneco, and B. G. Morton. Music emotion recognition: a state of the art review. *11th International Society for Music Information Retrieval (ISMIR) Conference*, pages 255–266, 2010.
14. G. J. Klir and B. Yuan. *Fuzzy Sets and Fuzzy Logic Theory and Applications.* Prentice Hall PTR, 1995.
15. S. Kolozali, G. Fazekas, M. Barthet, and M. Sandler. Knowledge representation issues in musical instrument ontology design. In *12th International Society for Music Information Retrieval Conference (ISMIR)*, Miami, Florida (USA), 2011.
16. M. Levy and M. Sandler. Lightweight measures for timbral similarity of musical audio. In *Proc. of the 1st ACM workshop on Audio and Music Computing multimedia*, New York, 2006.
17. C. Liu, D. Liu, and S. Wang. Fuzzy geopspatial ontology model and its application in geospatial semantic retrieval. *Journal of Convergence Information Technology*, 6, 2011.
18. C. D. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval.* Cambridge University Press, 2008.
19. S. Mirjalili and A. Sadiq. Magnetic optimization algorithm for training multi layer perceptron. *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*, pages 42 – 46, 2011.
20. R. Neumayer. Musical genre classification using a multi layer perceptron. *Proceedings of the 5th Workshop on Data Analysis (WDA'04)*, pages 51–66, 2004.

21. S. Rajashekaran and G. Vijayalksmi. *Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications*. Springer, 2004.
22. V. Sebastien, D. Semastien, and N. CONRUYT. An ontology for musical performances analysis. In *2010 Fifth International Conference on Internet and Web Applications and Services*, pages 538–543, Barcelona, May 2010.
23. N. Singhai and S. K. Shandilya. A survey on: Content based image retrieval systems. *International Journal of Computer Applications*, 4(2):22–26, 2010.
24. U. Straccia. Towards spatial reasoning in fuzzy description logics. *The annual IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 512–517, 2009.
25. D. Tidhar, G. Fazekas, S. Kolozali, and M. Sandler. Publishing music similarity features on the semantic web. *Proceedings of the 10th International Society for Music Information Retrieval Conference*, 2009.
26. K. Trohidis, G. Tsoumakas, G. Kalliris, and I. Vlahavas. Multi-label classification of music into emotions. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)*, 2008.
27. D. Turnbull, L. Barrington, D. Torres, and G. Lanckriet. Semantic annotation and retrieval semantic annotation and retrieval of music and sound effects. In *IEEE Transactions on Audio, Speech, and Language Processing*, volume 16, February 2008.
28. J. Wang, X. Anguera, X. Chen, and D. Yang. Enriching music mood annotation by semantic association reasoning. *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1445–1450, 2010.
29. T. Weyde and K. Dalinghaus. A neuro-fuzzy system for sequence alignment on two levels. *Mathware & softcomputing*, 11:197–210, 2004.