

Print ISSN: 0022-2755

# Journal of Mines, Metals and Fuels

Contents available at: www.informaticsjournals.com/index.php/jmmf

# The Effectiveness of Indian Music in Emotion Pattern Recognition under the Framework of Machine Intelligence

## Paromita Das<sup>1\*</sup> and Biswarup Neogi<sup>2</sup>

<sup>1</sup>PhD Scholar, MAKAUT, West Bengal, India. E-mail: rchparomita@gmail.com
<sup>2</sup>Associate Professor, JIS College of Engineering Kalyani, India. biswarupneogi@gmail.com

#### **Abstract**

Experts in music therapy has suggested music as an aid to give the positive state of mind by keeping all sorts of depression and anxiety away. Music helps to bring back the original state of vibration by controlling our emotions [1]. A music is a combined effect of melody, the singer's voice, and linguistics. The singer voice expresses the singer's emotion like glad sorrow, anxiety, peace, tiredness and which in turns control the listener's mental state. Indian music is analyzed and an approach for information retrieval to propose a therapeutic system through detection and identification of Indian music is initiated. Music Information Retrieval is a powerful tool to analyze different characteristics of a music. However, in this approach different traits of music are studied, and categorization has been done which leads to the therapeutic cause.

Keywords: Emotion, Music detection, Perception, Recognition, Signal Processing

# 1.0 Introduction

Music helps to develop a human brain from the prenatal stage of life, exposure to music not only uplift the mother's emotion in pregnancy, but leave a positive effect to unborn<sup>1</sup>. As voice engaged the sentiments of the singer as a consequence, it can easily move the listener's heart and mind. With the advancement of technology computers are implemented to appreciate the emotional part of the multimedia's content for accurate information retrieval<sup>3,4</sup>. Man machine interaction as example automatic recognition of the emotion of the people inside a conference room, a restaurant and the shopping mall is the most appropriate example of Music Emotion Recognition (MER)<sup>5,6</sup>. Music Information Retrieval (MIR) is the tool helps to extract features and these features are based on - categorical method and the dimensional method. The categorical method defines

several stages and dimensional method offers various axes to portray emotion into a plane. These features are well accepted in categorization emotional conditions gradation<sup>2</sup>.

# 2.0 Present State of Art

Most of the previous initiatives deals with different classes of emotion i.e arousal and valance<sup>7,8,9,10,11</sup>. Arousal describes whether the emotion is exciting or calming and valance defines whether it is negative or positive. The pattern recognition technics are employed thereafter to train the classification model. some popular approaches are – AdaBoost. RT (Boost.R), multiple linear regression (MLR) and support vector regression (SVR)<sup>3</sup>. In response to the claim about the diversity of human assessment of musical reaction,<sup>13</sup> maintains acceptance of multi-label classifiers, which allow assigning multiple emotional classes to the comparable song. Figure 1: Thayer's arousal-valence emotion

<sup>\*</sup>Author for correspondence

plane is used to generate AV values in<sup>3</sup>, where the continuous view is developed. MER is created as a regression issue and SVR is used to get AV values. Regression methodology avoids the difficulties inherent in categorical approaches. Additionally, the issue of subjectivity can be avoided because there is more room to explain a song. For instance, in addition to the quadrant in which the song falls, one can further understand the strength of the emotion the song evokes by examining its AV values. Lyrics and the song's title were considered as features by Nadia Lachetar<sup>14</sup>, who then used Naive Bayes and the Ant Colony algorithm to classify emotions. To examine various methods of detection, A. Bram van de Laar described emotion recognition during music study. The concept of emotion assessment of songs

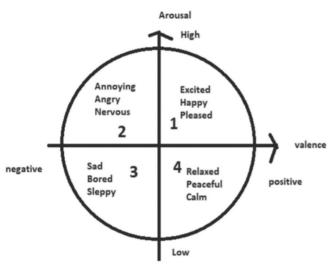


Figure 1: Arousal-valence emotion plane representation by Thayer

was proposed by Adit Jamdar, Karishma Khanna, and Jessica Abraham<sup>16</sup> and focused on auditory elements like danceability, pace, energy, loudness, mode, and lyrics. Studying lyrics involves using natural language processing. the achievement made possible by the K-Nearest Neighbors algorithm.

For easy reference, the emotions/classes are shown in Fig.1, also known as Thayer's arousal valence emotion plane<sup>12</sup>.

# 3.0 Methodological Aspects

The system includes testing and training procedures. Fig.2 details the "music emotion categorization (MEC)" system's detailed design<sup>16</sup>.

The input of the system is classical song, later, it is separated as instrumental and background noise. The process of features selection from the vocal segment for emotion recognition is studied through DSP system toolbox<sup>18</sup>.

# 4.0 Emotion Recognition System Description

The emotion recognition system comprises data collection, data analysis and consequent algorithm to realize singer voice emotion.

#### A. Data Collection

This part contains the collection of data from 50 individuals. These 50 persons are divided into a group of 5 members each depending upon their different characteristics. .5 songs of different types are selected. Each of them is played to each group members. A database is maintained containing feedback from everyone. Similarly, another database on songs is maintained with different traits of a particular song.

# **B.** Methodology

The intricate process of identification and detection has been implemented. Due to Fig.3. The typical process of detection, averaging, thresholding, choosing a wide range of widths, and discovering trivial instants is carried out. The process of identification is further comprising of padding, followed by Discrete Fourier Transform. The thumb rule behind is "exceptions are not examples and examples are not rules".

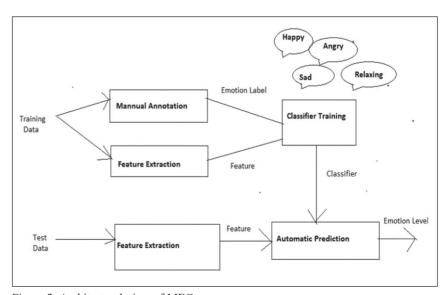


Figure 2: Architectural view of MEC

Table 1: Music Database

Fields	Music ID	Туре	spectrum	Tone	Pitch	Tempo ral Charac teristic
1 <sup>d</sup> Music specimen	Tumi.Amas Ciekechile	Tago re song		Time- 3:55.1 71 Value: 0.0254 61	Time:3:55,821 Value:443.086Hz(A4+1 2c,69)	Time: 3:55.125 Value:127.424bps
2 <sup>nd</sup> Music specimen	Saac.Marii. Cr. Biri.Jairi	<del>Baicul,Sone</del>		Time:3 .575 Value:0 .083501	Time: 3.552 Value: 484,47Hz(84- 33c,71)	Time: 3.555 Value: 75.395 3 bpm
3 <sup>rd</sup> Music specimen	ðasa: Meri- Mehfil- Mein	Ghazal		Time:3 .575 Value:0 .014406	Time:3.552 Value:130.506Hz(C 3-4c,48)	Time: 3.552 Value: 130.653 bpm
4 <sup>th</sup> Music specimen	Shiv <b>Çanday</b> .	Jandas.		Time:3 .575 Value:0 .041360	Time:3.552 Value:110.355Hz(A2 +6c,45)	Time: 3.575 Value: 121.62 8 bpm
5 <sup>th</sup> Music specimen	dbousest utl. bhagavat a	stuti	or property	Time:3 .575 Value:0 .037002 8	Time:3.552 Value:69.349Hz(C#2 +1c,37)	Time: 3.552 Value: 133.74 5 bpm

Table 2: Human Feedback

Fields	Person ID	Age	Sex	Profession	Religion	Response
	001	18	Male	Student	Hindu	liked
	002	45	Female	Housewife	Hindu	liked
Trial 1	003	60	Male	Retired	Hindu	liked
	001	18	Male	Student	Hindu	disliked
	002	45	45 Female Housewife		Hindu	liked
				Retired Army	Hindu	
Trial 2	003	60	Male	Officer		liked
	001	18	Male	Student	Hindu	liked
	002	45	Female	Housewife	Hindu	disliked
				Retired		
Trial 3	003	60	Male	Army	Hindu	disliked
				Officer		



Figure 3: Flow Chart

This research is achieved on various songs. Here every note supports a type of analogous pattern indicated in the Fig.3. The instant we play one note exactly after other, it preliminary produces enough high amplitude but diminishes by the time. Each note duration based on time and domain traits helps to identify and recognize required frequency.

The whole system is assembled into various steps as portrayed in the subsequent subsection.

### C. Proposed Technique

Step 1: In case of non-psychological patient the machine interacts directly with the patient and for others one single contact person from the family to know the symptoms of the disease.

Step 2: Referring the existing database (1) containing the symptoms along with the name of the disease, the conclusion about the disease name is determined. Then, the relevancy is judged by backward reasoning. If some symptoms are unknown to knowledge base those are included for upgradation of database. Theses may be basically realized keeping weights for every single symptom to establish precedence because several diseases may contain identical signs with distinct intensity.

Step 3: After step 2, where disease is revealed then the disease to the treatment database (2) implemented to decide the right kind of music therapy. Several music samples are there from same attributes and traits.

Step 4: The preliminary stage of man machine interaction is already discussed in step 1 and in the second stage, the liking and disliking of the particular individual are taken into consideration. This last stage will be accomplished by using heuristic analysis of already existing predetermined music medicine.

Interpretation of the symbols used to signify production system graphically

1.	Knowledge base
	Modification
2.	A Different set of symptoms, diseases, or music domains
3.	Interface with Patient
4.	Conclusion
5.	Correspondence

The symptoms and disease plotting depicted in the Figure 5 is, clearly built on domain, particular expertise in therapeutic, however, the surgical cases are excluded here. Domain precise

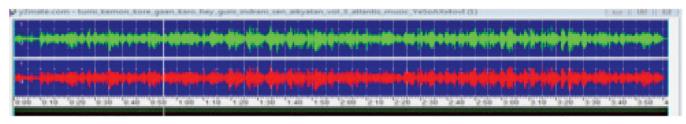


Figure 4: Vocal sample

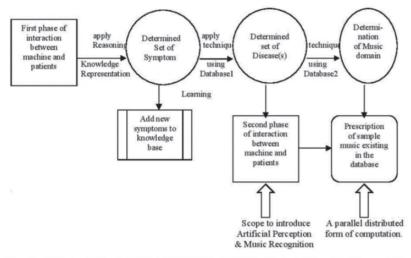


Figure 5: Complete Intelligent System Control flow

knowledge is mandatory to establish a mapping between disease and the music database. Rigorous trial using the learning process enriched both the database.

Here we consider separating the world of music into various categories. Nevertheless, any preliminary database considered as too insignificant as equated to the whole number of music deviations that the actual world holds.

Here is the example of a disease to the music field record for Indian Ragas

The efficient therapy irrespective of race, nationality and religion may be achieve by considering numerous sub fields of music domain. Data mining technic is considered one of the most appreciated processes to find the appropriate sub field in the said domain.

Table 3 Disease to Music Map

Name of the Disease	Appropriate time to Play	The name of the Ragas	
Insomnia	Middle of Night	Bagesri	
Indigestion	Morning	Ahir Bhairab	
Gallstone	Anytime in Spring	Basant-bahar	
lack of Confidence	Late Morning	Asavari	
Rheumatic Arthritis	Early Morning	Bhairavi	
Anorexia	Late Night	Chandrakauns	
High blood pressure	Early Night	Tori or Bhupali	
Anxiety	Early Afternoon	Bhim Palasri	
Heart ailments	Late Night	Chandrakauns	
Low blood pressure	Just after Midnight	Malkauns	
Easing tension	Midnight	Hanswadhani or Kalavati	

Table 4: Mood to Music Map

Specification of Mood	Appropriate time to Play	Ragas
Humid, cool, soothing, and deep mood	Second Prahar of Night	Kafi
heavy, Pleasant, deep, and gloomy mood	Anytime in Monsoon	Megha, Poorya, Dhanasri.
Pleasing, soft, sweet and refreshing feel	Second Prahar of Night	Mishra Mand.

#### D. Detection

#### 1. Averaging

The first step is averaging. The large number of music samples with various fluctuations can effectively reduce by averaging of every 100 samples. The first 100 samples assigned as the 1st while average of the following 100 samples as 2nd sample. Slow decaying signal implies denser averaged signal and rapidly decaying averaged signal indicates envelope of the actual signal.

#### 2. Thresholding

Immediately after averaging thresholding are involved –

- (a) Constant thresholding: This process includes detection of peaks from the averaged signal. Later maximum quantity of peaks are acquired based on single optimum estimate
- (b) Adaptive thresholding: This process includes consideration of certain fixed threshold value, intended for certain consideration it may possibly greater than the highest note value and some cases it is very low that two consecutive note peaks get combined and detected as a single note. The drawback of adaptive thresholding is due to the long silence between two notes leads to very low threshold value and consequently discarded.

### 3. Width Selection and Instant Finding

This step involves the detection of the peaks of the signal obtained afterward thresholding. Among all the threshold values some of them are more than the maximum and some of them are so low that it became very difficult to have two separate peaks. so here implementation of Adaptive thresholding appeared.

The window width is chosen depends on worst-case scenario and contains all zeroes. Implementing our sample, it observed as 57 therefore width considered as 50 for assurance. The worst situation is the variation of speed for a chosen frequency of sampling. If the song is slow, then it is more and if the song is fast, then it is less. Consequently, the smallest width estimated:

Here "fs" is the frequency of s the sampling and "A" defines the complete measures of samples found in the actual signal related to individual sample of the averaged of the signal under consideration.

#### 4. Detection

In this part the decision about the duration of notes played are discussed. The period of the first note played is the time taken between two consecutive instant (preferably the first and the second one), the period of the second notes performed is the time gap between the second and third instance and so on.

#### 5. Padding Zeroes

In case of offered note duration, the Discrete Fourier Transform is obtained after cropping the original signal, the results are not accurate as per our expectation. Here came the need for padding zeros. We can pad zeros of different length and can apply the same used for several portion in the cropped signal and later perform Discrete Fourier Transform. Now from the various outcomes the closet results are chosen considering the point that the zeros will be of the same length and applied on both sides of the picked version. The observed t signal helps to locate the notation.

#### 6. Frequency Detection and Notation Assignment

Once zero padding is done, Discrete Fourier Transform of the subsequent signal happens. Determination of the resultant frequency using equation 1 is the immediate step and the real note is designated using Table 5.

Song Type							
	Time	Peak frequency (Hz)	Bin frequency (Hz)	Peak pitch	Bin pitch	Power in dB	phase
Thumri 1	5:10- 5:11	1536-1553	1539-1550	G6-35e G6-17C	G6-32C G6-20C	-5029	-3.09 3.005
Thumri 2	5:10- 5:11	993.147	979-998	B5+9C	B5-14C— B5-19C	-3524	-2.38— 2.74
Gazal, 1	5:10- 5:11	13867- 13870	13824- 13910	A9-26C	A9-32C— A9-21C	-5443	.853— 3.133
Gazal.2	5:10- 5:11	1722-1728	1701-1722	A6- 38C— A6-32C	A6-38C— A6-32C	-3627	-2.64- 2.97

Table 5: The Audio Features under consideration

# 5.0 Identification of Vocalists

# A Audio Fingerprinting in Identification of Vocalist

Audio fingerprinting is an efficient way for identifying music and vocalist. The different music performance of the same song has different fingerprint. Fingerprinting technics consist of two elementary building blocks i.e., extraction of seed tracks and maintenance of a database of previously extracted fingerprint with a unique identifier coupled with additional metadata of the music considered.

The features included in frequency domain are MFCC<sup>20,21</sup>, Spectral flatness measure or spectral crest factor per frequency band<sup>23</sup>, spectral centroid per sub band<sup>24</sup>, band energies<sup>25</sup>, spectral peaks<sup>26,27</sup>, statistical moments of sub band<sup>28</sup> and modulation frequency<sup>29</sup>. The similarity measures for comparison include correlation measure<sup>30,31</sup>, the Euclidean distance<sup>32</sup> and Manthan distance<sup>33</sup>.

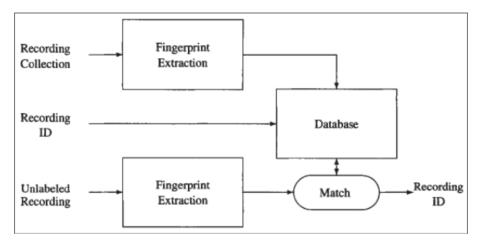


Figure 6: Training and query Phase of Audio Fingerprint

Table 6 Same	Song by	amerent	Singer	(Male	and	remaie)

				Peak	Daale	Bin			
Field	Music ID	Туре	Time	frequenc	Peak Pitch	Frequenc	Bin Pitch	Power(dB)	Phase
ciu	Ke jeno	.,,,,	Time			,	Dill 7 Itell	, one (ub)	1.7430
	go			271.666-		6589.16-			
1st Music	dekeche	Female	12.881-	272.071	C#4-35c-	6718.36	6#8-15c-	-4940	-1.66038- 1.34297
Specimen	amay	Voice	12.961	Hz	c#4-32c	Hz	G#8+19c	'	'
	Ke jeno								
	go			140.928-		10723.5-			-2.37306-2.32999
2nd Music	dekeche	Male	10.106-	141.494H	c#3+29c-	10895.8	E9+29c-	-4846	-2.57500-2.52555
Specimen	amay	Voice	10.129	z	c#3+36c	Hz	F9-44C		

Table 7 Same Song by different Singer (Different Male)

				Peak		Bin			
				frequenc	Peak	Frequenc			
Field	Music ID	Type	Time	у	Pitch	у	Bin Pitch	Power(dB)	Phase
	Jadi			204.407-		6416.89-			
1st Music	kichu	Shyamal	3.477-	207.661	G#3-27c-	6589.16H	G8+40c-	-5347	-2.61054-0.569334
Specimen	amake	Mitra	3.511	Hz	A3-37c	Z	G#8-15c		
	Jadi			186.385-		7364.36-		61 40	3.10252-2.98044
2nd Music	kichu	Srikanta	3.779-	205.428	G3+3C-	7536.62H	A#8-22c-	-6149	'
Specimen	amake	Acharya	3.813 Hz	Hz	G#3-19c	z	A#8+18c		

# 6.0 Conclusion

The present initiative provides a methodology to diagnose a disease of non-surgical category through music therapy. This is a work to identify Indian music and singer and based on the song to disease mapping proposed a music therapy in terms of rule and knowledge based. This research does not deal with music therapy convention, but implementation of digital signal processing and machine intelligence enable the same for treatment. A further extension of the research demands investigation. The work mainly based on Indian classical and no specific DSP tools, simulation process and applications are mentioned. Emotion recognition for specific lyrics and voice are portrayed and consequent application for lifting the original state of vibration of mind is suggested.

# 7.0 References

- P. Das, S. Gupta, B. Neogi, (2020): "Measurement of effect of music on human brain and consequent impact on attentiveness and concentration during reading", Procedia Computer Science, Elsevier Science, ISSN:1877-0509, Vol. 172, pp. 1033-1038, 2020. (https://doi.org/10.1016/j.procs.2020.05.151)
- Kee Moe Han, Theingi Zin, Hla Myo Tun (2016): "Extraction Of Audio Features For Emotion Recognition System Based On Music", *International Journal of Scientific & Technology Research* volume 5, issue 06, june.
- Yi-Hsuan Yang, Yu-Ching Lin, Ya-Fan Su, and Homer H. Chen, "Music Emotion Classification: A Regression Approach", in Multimedia and Expo, IEEE International Conference, 2007, pp.208-211.
- 4. Yi-Hsuan Yang, Ya-Fan Su, Yu-Ching Lin, and Homer H. Chen, (2007): "Music Emotion Recognition: The Role of Individuality." HCM'07, September 28, Augsburg, Bayaria, Germany.
- 5. S. Reddy and J. Mascia, (2006): "Lifetrak: music in tune with your life," Proc. Human-centered multimedia, pp.25–34.
- S. Dornbush, K. Fisher, K. Mckay, A. Prikhodko, and Z. Segall, (2005): "XPOD – A human activity and emotion aware mobile music player," Proc. Int. Conf. Mobile Technology, Applications and Systems.
- E. Schubert, (1999): "Measurement and time series analysis of emotion in music," Ph.D. dissertation, School of Music & Music Education, Univ. New South Wales, Sydney, Australia.
- 8. Y. Feng, Y. Zhuang, and Y. Pan, (2003): "Popular music retrieval by detecting mood," Proc. ACM SIGIR, pp. 375–376.
- 9. D. Yang and W. Lee, (2004): "Disambiguating music

- emotion using software agents," Proc. Int. Conf. Music Information Retrieval, pp. 52–58.
- T. Li and M. Ogihara, (2006): "Content-based music similarity search and emotion detection," Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing, pp.17–21.
- L. Lu, D. Liu, and H.-J. Zhang, (2006): "Automatic mood detection and tracking of music audio signals," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 14, no. 1, pp. 5–18.
- 12. R. E. Thayer, (1989): The Biopsychology of Mood and Arousal, New York, Oxford University Press.
- 13. P. Synak and A, Wieczorkowska, (2005): "Some issues on detecting emotions in music," Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing, Springer, pp. 314–322.
- 14. Nadia Lachetar, Halima Bahi, "Song classification", Computer Science Department.
- 15. Bram van de Laar, (2006): "Emotion detection in music, a survey" in Twente Student Conference on IT vol 1,p.700
- Adit Jamdar, Jessica Abraham, Karishma Khanna and Rahul Dubey, (2015): "emotion analysis of songs based on lyrical and audio features", *International Journal of Artificial Intelligence & Applications* (IJAIA), Vol.6, No. 3, May.
- 17. C. Ho, W.T. Tsai, K.S. Lin, and H. H. Chen, (2013): "Extraction and alignment and evaluation of motion beats for street dance," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May.
- 18. Jay K. Patela, E.S. Gopia, (2015): Musical Notes Identification using Digital Signal Processing, Proceeding of 3rd International Conference on Recent Trends in Computing (ICRTC-2015)
- S. Gupta, (2007): "Music Perception, Cognition and Effect: A case study of Artificial Intelligence illustrating Disease Diagnosis", Proceedings of the 5<sup>th</sup> International Conference on Information Science, Technology and Management (CISTM) 2007, ISBN: 978-0-9772107-8-7, 50, pp. 1-16
- P. Cano, E. Batlle, H. Mayer, and H. Neuschmied, (2002): "Robust Sound Modeling for Song Detection in Broadcast Audio," in Proceedings of the 112th Audio Engineering Society Convention, Preprint 5531, Munich: Audio Engineering Society (AES).
- A. Ramalingam and S. Krishnan, "Gaussian Mixture Modeling of Short-Time Fourier Transform Features for Audio Fingerprinting," *Transactions on Information Forensics and Security*, Vol. 1, no. 4, pp. 457-463, Dec. 2006, ISSN: 1556-6013. DOI: 10.1109/ TIFS.2006.885036.
- 22. O. Hellmuth, E. Allamanche, J. Herre, T. Kastner, M.

- Cremer, and W. Hirsch, (2001): "Advanced Audio Identification Using MPEG-7 Content Description," in Proceedings of the 111th Audio Engineering Society Convention, Preprint 5463, New York: Audio Engineering Society, 2001
- J. S. Seo, M. Jin, S. Lee, D. Jang, S. Lee, and C. D. Yoo, (2005): "Audio Fingerprinting Based on Normalized Spectral Subband Centroids," in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Institute of Electrical and Electronics Engineers (IEEE), 2005, pp. 213-216, ISBN: 0-7803-8874-7. DOI:10.1109/ICASSP.2005.1415684
- 24. A. Kimura, K. Kashino, T. Kurozumi, and H. Murase, (2001): "Very Quick Audio Searching: Introducing Global Pruning to the Time-Series Active Search," in Proceedingsof the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 3, Institute of Electrical and Electronics Engineers (IEEE), pp. 1429-1432, ISBN: 0-7803-7041-4. DOI: 10.1109/ICASSP.2001.941198.
- 25. J. Haitsma and T. Kalker, "Speed-Change Resistant Audio Fingerprinting Using Auto-Correlation," in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 4, *Institute of Electrical and Electronics Engineers* (IEEE), 2003, pp. IV-72831-, ISBN: 0-7803-7663-3. DOI: 10. 1109/ICASSP.2003.1202746.
- C. Yang, "MACS: Music Audio Characteristic Sequence Indexing for Similarity Retrieval," in Proceedings of the Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz: *Institute of Electrical and Electronics Engineers* (IEEE), Oct. 2001, pp. 123-126, ISBN: 0-7803-7126-7. DOI: 10.1109/ASPAA.2001.969558.
- 27. A. Wang, (2003): "An Industrial Strength Audio Search Algorithm," in Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR), Washington.
- 29. S. Kim and C. D. Yoo, (2007): "Boosted Binary Audio Fingerprint Based on Spectral Subband Moments," in

- Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP), vol.1, *Institute of Electrical and Electronics Engineers* (IEEE), Apr. 2007, pp. I-241I-244-, ISBN: 1-4244-0727-3.DOI: 10.1109/ICASSP.2007.366661
- S. Sukittanon and L. E. Atlas, (2002): "Modulation Frequency Features for Audio Fingerprinting,"in Proceedings of the International Conference on Acoustics Speechand Signal Processing (ICASSP), Institute of Electrical and Electronics Engineers (IEEE), pp. II–II, ISBN: 0-7803-0946-4. DOI: 10. 1109/ ICASSP.2002.1006107.
- D. Schonberg and D. Kirovski, "Fingerprinting and Forensic Analysis of Multimedia," in Proceedings of the 12th ACM Multimedia Conference, New York, 2004. DOI: 0.1145/1027527.1027712.
- 32. J. I. Martinez, J. Vitola, A. Sanabria, and C. Pedraza, (2002): "Fast Parallel Audio Fingerprinting Implementation in Reconfigurable Hardware and GPUs," in Proceedings of the Southern Conference on Programmable Logic (SPL), *Institute of Electrical and Electronics Engineers* (IEEE), Apr. 2011, pp. 245-250, ISBN: 978-1-4244-\_8847-6. DOI: 10.1109/SPL.2011.5782656.
- C.J.C. Burges, J.C. Platt, and S. Jana: "Extracting Noise-Robust Features from Audio Data," in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), *Institute of Electrical and Electronics Engineers* (IEEE), ISBN: 0-7803-0946-4. DOI: 10.1109/ICASSP.2002.1005916.
- 34. G. Richly, L. Varga, F. Kovacs, and G. Hosszu, (2000): "Short-Term Sound Stream Characterization for Reliable, Real-Time Occurrence Monitoring of Givi n Sound-Prints," in Proceedings of the 10th Mediterranean Electrotechnical Conference. Information Technology and Electrotechnology for the Mediterranean Countries (CMeleCon), Vol.2, *Institute of Electrical and Electronics Engineers* (IEEE), 2000, pp. 526-528, ISBN: 0-7803-6290-X. DOI: 10.1109/MELCON.2000.879986.