International Journal of Engineering & Technology, 7 (3.29) (2018) 538-542

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

When music and emotions meet machine

Soumya Kota [1] *, M. Krishna Prasanna [2], D. Madhusudhan Reddy [2]

[1] *IV B. Tech, Student Department of IT, BVRIT Narsapur, Medak*
[2] *Assistant Professor Department of IT, BVRIT Narsapur, Medak*
**Corresponding author E-mail: soumyakota1810@gmail.com*

Abstract

Perception about music varies person to person. But, what most of them love listening to, is music. The discussion that's exciting is, emotion and cognition go hand in hand in music.Genres like classical, jazz, rock, blues, country, no matter what, people listen, feel and enjoy the essence of it.Surprisingly we can relate music and emotions. Every piece of music is created to convey certain emotions and that is why emotion has been considered a powerful way for accessing music information. Large collection of music is being created and accessed today. To make the user's choice easier, the (MIR) music informational retrieval community has made an effort to train a machine (like computer) to detect the emotion of a music signal.Making a machine identify the emotion of music leads to better interaction between the machine and the user. It is possible to play the music according to the user's mood by detecting the psychological or facial cues. Also, this can be used to sort user's music collection according to their mood with an adjective described. Intensity and quality of musical emotions change overtime.Therefore, various discussions and debates led to research and proposal of data model training and result set according to the user's emotion.

Keywords: *Music Cognition; Emotions; Mir (Music Information Retrieval) System; Smart Places.*

1. Introduction

Music is an art that everybody enjoys. Probably people sang or made sounds before they could even speak, which makes music a part of human evolution. Music is linked strongly to social contact, motivation and stress buster ability. Only certain number of people can play music, but most of them can atleast hum a tune. Music is also an experience that generates reflexes to the cognitive capabilities. It is connected with many basic human needs which leads to the result of thousands of years of neurobiological development. Music, though it appears to be similar to features of language, is more rooted in the primitive brain structures that are involved in motivation, reward and emotion. Music and emotion [5] is a branch of music psychology which seeks to understand the relationship between music and human effect. This topic leads to numerous areas of study, including characteristics of listener, nature of emotional reactions to music, components of a certain musical performance or creation. With the evolution of large collection of music and various preferences of listeners, Music Information Retrieval (MIR) [1] community develops various algorithms and increased rapidly over the last decade. MIR is used for retrieving information from music. It uses data source, feature representation, statistics and machine learning. This music can be in symbolic format (e.g., a MIDI file), in audio format (e.g. an mp3 file), or in vector format (e.g., a scanned score). MIR research takes advantage of technologies and knowledge derived from signal processing, machine learning, music cognition, database management, human-computer interaction, music archiving or sociology of music. As the data is large, managing music pieces based on album name, artist name, song title, and genre are no longer sufficient.Music can be a type of perceptual illusion. The brain imposes structure and order on a sequence of sounds that, in effect, creates an entirely new system of meaning. The appreciation of music is tied to the ability to predict what willoccur next in the song. But this structurehas toinvolve some level of the unexpected, or it becomes emotionally devoid.Skilled composers manipulate the emotion within a song by knowing what their audience's expectations are, and controlling when those expectations will be met. This successful manipulation is what that excites the song. Professor Daniel Levitin, a neuroscientist and composer, unpacks the mystery of the emotion in music by explaining how the brain's emotional [4], language and memory centres are connected during the processing of music – providing what is essentially a synesthetic experience. The extent of this connection varies among individuals, which is how certain musicians tend to create pieces of music which are brimming with emotional quality, and others simply cannot.

Music is like a binding factor in our social culture.It is a feature with and about us.How music came into being is, like most other features in our evolution [6], hard to pinpoint. Evolutionary evidence over a wide range of cultural groups reveals diversity of song and instrument, yet gaps and speculative considerations remain: some cultures sing a lot, some sing less, but most do sing and perhaps Neanderthals sang more than Sapiens (Mithen, 2006). Music is typically something shared something social; we may sing in the shower or on a solitary walk (Whitehead, 1938/1967; Rousseau, 1966), but music is most of the time social, communicative, expressive, and oriented toward others.Music cuts across diverse cognitive capabilities and resources, including numeracy, language, and spatial perception. In the same way, music intersects with cultural boundaries, facilitating our "social self" by linking our shared experiences and intentions. Therefore, MIR (Music Information Retrieval) is a combination of Machine Learning, SignalProcessing, ComputerScience, Psychology, Musicology, Human Computer interaction as shown in Fig 1.

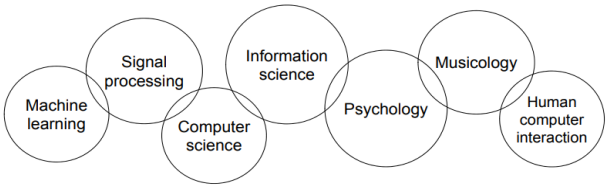

Fig. 1: The Components, Which Make MIR Interdisciplinary.

2. Related works

Stephen Davies and Jerrold Levinson,being the most influential philosophers in the aesthetic of music, Davies termed his concept as the expressiveness of emotions in music appearance emotionalism, which holds that music expresses emotion without feeling it. Objects convey emotion because their structures contain certain features that depict emotional expression. Music's temporally unfolding dynamic structure and configurations of human behaviour associated with the expression of emotion, count most for music's expressiveness. Emotions can be noted by an observer through listener's gait,posture,attitude and gestures. Appearance emotionalism claims many listeners' perceiving associations constitute music's expressiveness. Expressiveness of music is response-dependent i.e it is realized in the listener's judgement.PhilosopherJenefer Robinson argues that the process of emotional elicitation begins with an automatic, immediate response that initiates motor and autonomic activity and prepares us for possible action causing cognition that might enable listeners to name the emotion they felt.Robinson also argues that music is a series of simultaneous processes. Therefore, it is an ideal medium for mirroring cognitive aspects of emotion as musical themes desiring resolution. These simultaneous musical processes can reinforce or conflict with each other and thus express the ay one emotion morphs into overtime.

Mood/Emotion Detection in Music can evoke a wide variety of emotional responses which is one of the reasons it is heavily utilized during movies. Even though culture and education can significantly affect the emotional response of listeners to music there has been work attempting to automatically detection mood and emotion in music. As this is information that listeners frequently used to discuss music and not readily available in most cases there has been considerable interest in performing it automatically. Even though it is occasionally cast as a single-label classification problem it is more appropriately considered a multi-label problem or in some cases a regression in a continuous "emotion" space. Researchers gave experts of fast tempo, slow tempo, major tone music, minor tone music to some participants and these structures were chosen because they are known to convey happiness and sadness. Participants related their own emotions with levels of happiness or sadness after listening to the specified structures. Researchers have also found that familiarity of a piece of music increases the emotions experienced by the listener. Musicmay not only bring out new emotions, but connect listeners with other emotional sources as it serves as a powerful cue to recall emotional memories back into our awareness. Therefore, bringing a strong connection between emotion and music within memory makes it a lot easier to recall when the other prompts it.

3. Role of technology in music

Technology has always played a vital role in music creation, distribution, and listening. Throughout history the making of musical instruments has followed the technological trends of the day. Additionally, more conceptual technological developments such as the use of music notation to preserve and transmit music, or the invention of recording technology have also had profound implications on how music has been created and distributed. In the ancient times, for humans to communicate, the only means of transmission and preservation of musical information was direct contact between a performer and a listener during a live performance. The first technological development that changed this was the gradual invention of musical notation. Although several forms of music notation exist, Western music notations are widely used in the world today. Thetwo historical developments in music technology: musical notation and audio recording, are particularly relevant to the field of MIR. These digital representations were originally engineered to simply store digitally the corresponding analog (for example vinyl) and physical (for example typeset paper) information but eventually were utilized to develop completely new ways of interacting with music using computers. It is fascinating to know how the old debates contrasting the individual model of purchasing recordings with the broadcasting model of radio are being replayed in the digital domain with digital music stores that sell individual tracks such as iTunes and subscription services that provided unlimited access to streaming tracks such as Spotify.

3.1. MIR (music information retrieval)

MIR is interdisciplinary by nature. Music Information Retrieval (MIR), the ideas behind music information and retrieval deals with the analysis and retrieval of music in digital form. It reflects the tremendous recent growth of music-related data digitally available and the consequent need to search within it to retrieve music and musical information efficiently and effectively. Arguably the birth of Music Information Retrieval (MIR) as a separate research area happened in Plymouth, Massachusetts in 2000 at the first conference (then symposium) on Music Information Retrieval (ISMIR). Although work in MIR had appeared before that time in various venues, it wasn't identified as such and the results were scattered in different communities and publication venues. During that symposium a group of computer scientists, electrical engineers, information scientists, musicologists, psychologists, and librarians met for the first time in a common space, exchanged ideas and found out what types of MIR research were happening at the time. MIR research until then was published in completely separate and different conferences and communities. The seed funding for the symposium and some predatory workshops that preceded it, was provided by the National Science Foundation (NSF) through a grant program targeting Digital Libraries [9]. The original organizers Don Byrd, J. Stephen Downie and Tim Crawford had backgrounds connected to symbolic processing and digital libraries but fortunately the symposium attracted some attendees that were working with audio. The resulting cross-fertilization was instrumental in establishing MIR as a research area with a strong sense of interdisciplinarity. When MIR practitioners convey what is Music Information Retrieval to a more general audience they frequently use two metaphors that capture how the majority, but not all, of MIR research could potentially be used. The first metaphor is the so called "grand celestial jukebox in the sky" and refers to the potential availability of all recorded music to any computer user. The goal of MIR research is to develop better ways of browsing, choosing, listening and interacting with this vast universe of music and associated information. To a large extent online music stores, such as the iTunes store, are well on their way of providing access to large collections of music from around the world. A complimentary metaphor could be termed the "absolute personalized radio" in which the listening habits and preferences of a particular user are analyzed or directly specified and computer algorithms provide a personalized sequence of music pieces that is customized to the particular user. As was the case with the traditional jukebox and radio for the first metaphor the emphasis is on choice by the user versus choice by some expert entity for the second metaphor.

3.2. Dimensions of MIR research

There are multiple different ways in which an MIR research can be organized into certain areas. They are 1) specificity which refers to the semantic focus of the task, 2) data sources which refers to the various representations (symbolic, meta-data, context, and

lyrics) utilized to achieve a particular task, and 3) stages which refers to the common processing stages in a MIR systems such as representation, analysis, and interaction. Specificity one can view many MIR tasks and systems though a retrieval paradigm in which a query is somehow specified by the user and a set of hopefully relevant results to that query is returned by the computer. This paradigm is very familiar from text search engines such as Google. MIR queries can take many forms such as an actual audio recording, a set of keywords describing the music, or a recording of a user singing a melody. Specificity refers to the range of retrieved results that are considered relevant for a particular query.In Emotion recognition returned results should all have a particular emotion or evokes a particular mood possibly spanning multiple genres and styles. The most wide circle of specificity is exemplified by simply shuffling in which any music recording in a database is considered relevant to a particular query. Although shuffling is trivial it is a common way of interacting with large audio collections.

3.3. Data sources and information

Another way of organising MIR tasks and systems is through the data sources and information they make use of. An important source of information is the music itself or the musical content as it is frequently termed as. This can be the digital audio recording or some symbolic representation of the music which in most cases can be viewed as a type of a musical score. In addition there is a lot of additional information that can be utilized for retrieval and analysis purposes that is not part of the actual music. This information that is about the music rather than the actual music is termed the context. Frequently it consists of text for example lyrics, blogs, web pages, and social media posts. Music information retrieval systems can combine multiple source of information. For example, a genre classification system might rely on both audio feature extraction from the recording, and text feature extraction from the lyrics.The audio of the music is taken as input.The content of the music is then analysed and then the content is retrieved. (Fig 2)

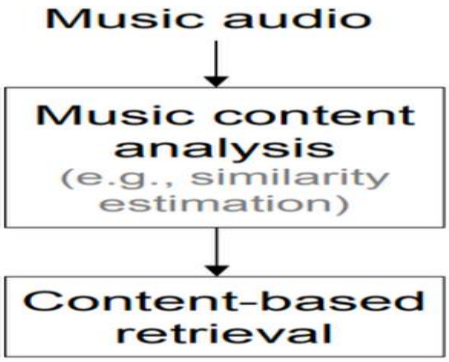

Fig. 2: Sequence of Music Processing.

3.4. Stages

Another way of organizing MIR algorithms [8] and systems is more procedural and through stages that are involved in most such systems. Parallels to these stages can be found in how humans perceive, understand, and interact with music.

- Representation

Hearing Audio signals are stored (in their basic uncompressed form) as a time series of numbers corresponding to the amplitude of the signal over time. Although this representation is adequate for transmission and reproduction of arbitrary waveforms, it is not particularly useful for analyzing and understanding audio signals. The way we perceive and understand audio signals as humans is based on and constrained by our auditory system. It is well known that the early stages of the human auditory system (HAS), to a first approximation, decompose incoming sound waves into different frequency bands. In a similar fashion, in MIR, time-frequency analysis techniques are frequently used for representing audio signals. The representation stage of the MIR pipeline refers to any algorithms and systems that take as input simple time-domain audio signals and create a more compact, information-bearing representations. The most relevant academic research area to this stage is Digital Signal Processing (DSP).

- Analysis - Understanding

Once a good representation is extracted from the audio signal various types of automatic analysis can be performed. These include similarity retrieval, classification, clustering, segmentation, and thumbnailing. These higherlevel types of analysis typically involve aspects of memory and learning both for humans and machines. Therefore Machine Learning algorithms and techniques are important for this stage.

- Interaction - Acting

When the signal is represented and analyzed by the system, the user must be presented with the necessary information and act according to it. The algorithms and systems of this stage are influence by ideas from Human Computer Interaction and deal with how information about audio signals and collections can be presented to the user and what types of controls for handling this information are provided.

3.5. Classification and clustering

Classification and Clustering Classification refers to the process of assigning one or more textual labels in order to characterize a piece of music. Humans have an innate drive to group and categorize things including music pieces. In classification tasks the goal is given a piece of music to perform this grouping and categorization automatically. Typically this is achieved by automatically analyzing a collection of music that has been manually annotated with the corresponding classification information [9]. The analysis results are used to "train" models [10] (computer algorithms) that given the analysis results (referred to as audio features) for a new unlabelled music track are able "predict" the classification label with reasonable accuracy. This is referred to as "supervised learning" in the field of machine learning/data mining. At the opposite end of the spectrum is "unsupervised learning" or "clustering" in which the objects of interest (music pieces in our case) are automatically grouped together into "clusters" such that pieces that are similar fall in the same cluster. There are several interesting variants and extensions of the classic problems of the classification and clustering. In semi-supervised learning the learning algorithms utilizes both labelled data as in standard classification as well as unlabeled data as in clustering. The canonical output of classification algorithms is a single label from a finite known set of classification labels. In multi-label classification each object to be classified is associated with multiple labels both when training and predicting.

When characterizing music several possible such groupings have been used historically as means of organizing music and computer algorithms that attempt to perform the categorization automatically have been developed. Although the most common input to MIR classification and clustering systems is audio signals there has also been work that utilizes symbolic representations as well as metadata (such as tags and lyrics) and context information (peer-to-peer downloads, purchase history).

4. Proposal method:-MER (music emotion recognition)

Music Emotion Recognition [2] [7] (MER) is a branch of MIR that is used for the identification of emotions/moods in musical pieces. In the academia, more and more multimedia systems that involve emotion analysis of music signals have been developed such as MusicSense, Mood Cloud. The various emotions of a human being are considered for categorising them.(Fig 3.)

This is mainly used in the fields of cinema, advertising, health, sports, and stress management.

Two main approaches are used for emotion retrieval:

- -Categorical models where discrete number of emotions (adjectives) are used.
- -Dimensional models that are for continuous values of emotions depending on axes.

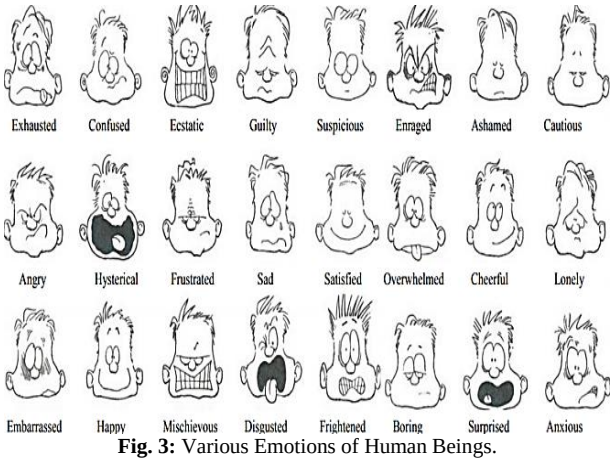

Fig. 3: Various Emotions of Human Beings.

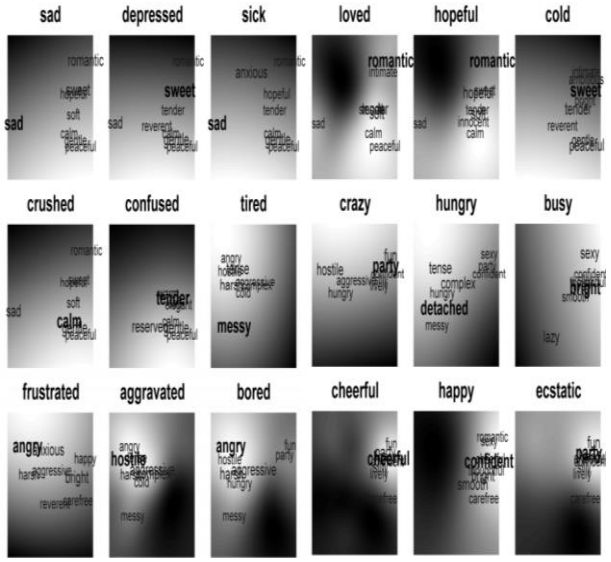

Fig. 4: A) Shows How Various Emotions are Categorized.

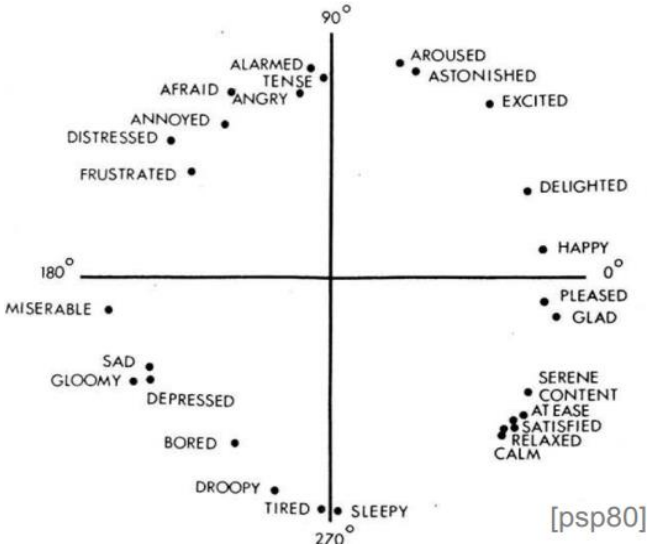

Fig. 4: B) Emotions Are Categorized Into Various Axis.

The above figures (Fig 4a and 4b) show us how different emotions are categorised into groups and into various axis.

4.1. Overall approach of classical pattern recognition approach

1) Selection of a mood model
2) Acquisition of training and test data
3) Feature extraction and selection
4) Selection of representation model
5) Model training
6) Model validation (Fig 5)

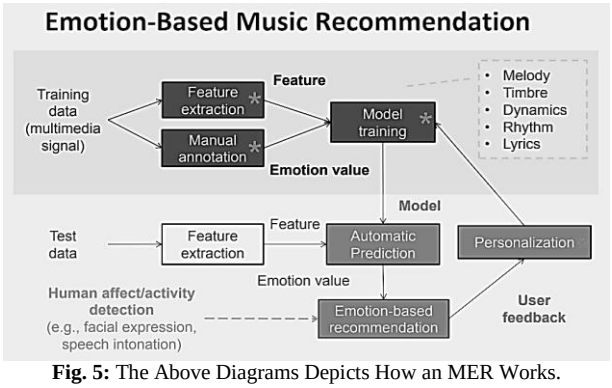

Fig. 5: The Above Diagrams Depicts How an MER Works.

Various applications can be made using MIR techniques and the below figure (Fig 6) shows how an application is made using the concepts of MIR and MER where the emotion matches at the activation level and predicts the songs to the user.

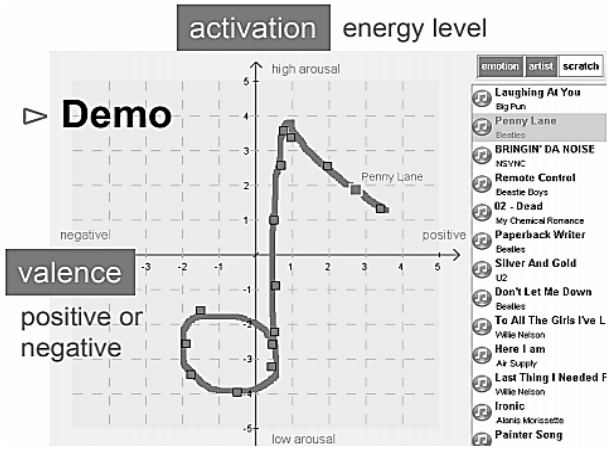

Fig. 6: Shows an Application Made Using MIR System

4.2. Music retrieval in emotion space has the following advantages

- No need of human labelling
- Scalable
- Easy to update and edit
- Emotion based music recommendation
- It is content based and easy to use.

4.3. Smart places

Music Emotion Recognition [3] is being used for smart spaces for example a residence,café, conference room,restaurant where they can play the background music based on the people inside the room which will help the people enrich their mood.

4.4. Applications of MER and MIR

This MIR and MER are basically used for:

1) Recognising the emotion of user by EEG for music recommendation.
2) Music therapy to cure patients
3) Incidental music selection

4) Karaoke System

5. Conclusion

MIR is an interdisciplinary field of research that involves many researchers to develop MIR and MER tools for emotion recognition.In this research the evolution of MIR has been discussed. The link between music and emotion [5] that involves music theory has been brought to contrast. Then the role of technology has been discussed.Later the MIR system has been discussed briefly including the data sources and classification stages.Finally, the applications of MIR has been quoted.Researchers feel this as a wide scope of research because it involves music, emotion,cognition,technology and many other aspects as well which helps them to think in a broader perspective.

6. Future works

As this field needs to be viewed from a broader perspective the development and scope of research still lies in the following aspects:
User studies - research should be done in ways to model the user along with his cultural and environment to create accurate user models.
Multimodality- Multi-modal features should be understood and the features that already exist should be made available to be integrated to build personalized music retrieval systems.
Multidisciplinary- Even though music information retrieval is multidisciplinary, we should still need to understand cross-disciplinary of knowledge and related methodologies.
Music through senses-MIR systems should also work for different sensory modalities like visual, tactile to provide extreme experience both in music listening and creation.
Musical companions- Music has an adverse potential in different factors of human life like education, therapy, cognitive development and social well being where MIR systems are used for connecting music to our personal memories and emotions. Therefore, MIR will help in developing and improving overall well-being of mankind.

References

[1] Markus Schedl, Emilia Gómez, JuliánUrbano "Music Information Retrieval: Recent Developments and Applications"-Volume 8, Issues 2-3 of Foundations and trends in information retrieval, ISSN 1554-0669, Now Publishers, 2014.
[2] Yi-HsuanYang,Homer H. Chen- Machine Recognition of Music Emotion: A Review, ACM Transactions on Intelligent Systems and Technology (TIST), Volume 3 Issue 3, May 2012.
[3] Yi-Hsuan Yang, Homer H. Chen, Music Emotion Recognition, CRC Press 2011.
[4] MaliniMohana" Music & How it impacts your brain,Emotions from Psych Central, in 2016.
[5] William Forde Thompson, "Music and Emotion: Psychological Considerations", Oxford University Press in 2011.
[6] Jay Schulkin and Greta B. Raglan "The evolution of music and human social capability", Front Neurosci, 2014; 8:292. Published online 2014 Sep 17.
[7] Yi-Hsuan Yang ; Yu-Ching Lin ; Ya-Fan Su ; Homer H. Chen- A Regression Approach to Music Emotion Recognition- IEEE Transactions on Audio, Speech, and Language Processing,IEEE 2008.
[8] Rainer Typke, FransWiering, Remco C. Veltkamp, "A SURVEY OF MUSIC INFORMATION RETRIEVAL SYSTEMS", ISMIR in 2005.
[9] N. Scaringella G. Zoia D. Mlynek, Automatic genre classification of music content: a survey, IEEE Signal Processing Magazine (Volume: 23, Issue: two, March 2006).
[10] Rudi Cilibrasi Paul Vitányi and Ronald de Wolf Algorithmic Clustering of Music Based on String CompressionComputer Music JournalVolume 28 | Issue 4 | winter 2004.