

Research Article

MEL-IRIS: An Online Tool for Audio Analysis and Music Indexing

Dimitrios Margounakis, Dionysios Politis, and Konstantinos Mokos

Computer Science Department, Aristotle University of Thessaloniki, University Campus 541 24, Thessaloniki, Greece

Correspondence should be addressed to Dimitrios Margounakis, dmargounakis@learn-it.gr

Received 1 May 2009; Revised 14 August 2009; Accepted 21 September 2009

Recommended by Peijun Shan

Chroma is an important attribute of music and sound, although it has not yet been adequately defined in literature. As such, it can be used for further analysis of sound, resulting in interesting colorful representations that can be used in many tasks: indexing, classification, and retrieval. Especially in Music Information Retrieval (MIR), the visualization of the chromatic analysis can be used for comparison, pattern recognition, melodic sequence prediction, and color-based searching. MEL-IRIS is the tool which has been developed in order to analyze audio files and characterize music based on chroma. The tool implements specially designed algorithms and a unique way of visualization of the results. The tool is network-oriented and can be installed in audio servers, in order to manipulate large music collections. Several samples from world music have been tested and processed, in order to demonstrate the possible uses of such an analysis.

Copyright © 2009 Dimitrios Margounakis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

This paper presents MEL-IRIS, a tool developed during our research on chroma in music. “MEL-IRIS” derives from the words Melodic Irida. The main task of MEL-IRIS is to chromatically analyse music files (MIDI, WAV, and MP3). It provides a unique method of classification, identification, and visualization of songs and can be applied in large song databases as well as in web applications. MEL-IRIS is designed for processing musical pieces from audio servers, creating a unique chromatic index for each of them, and classifying them according to the chromatic index. Each chromatic index results to a colourful strip, which characterizes a song and can serve as a signature and as a classifier as well.

The initial stages of the research and the fundamental background of chromatic analysis are described in [1]. The first version of MEL-IRIS, as well as early results, has been presented in [2]. These papers are referred where necessary.

MEL-IRIS at present is available in two different versions: the stand-alone PC version 2.0 (which is totally redesigned in order to overcome the weaknesses of the first version) and the network-based version 3.0. The latter is a client-server application and can be found at

<http://nomos.csd.auth.gr:8080/meliris>. This version aims at the organization and manipulation of large music collections on the Internet.

Section A: Theoretical Background

2. Basics

2.1. Chroma-Chromaticism-Chrominance. A thorough discussion about chroma has been described in [1]. What one should keep in mind is that the concept of *music chroma* has not been strictly defined and, though, it is an open problem. Shepard has defined with chroma the note's position within the octave and has created a nonlogarithmic pitch helix, the *chroma circle*, which clearly depicts octave equivalence [3]. This has led to rather complex pitch-space representations in which the chromatic tone scale, the circle of fifths, octave circularity, and other properties are all accounted for. This approach perceives chroma as extension to the concept of tonality. It has been argued that the dimension of tone chroma is irrelevant in melodic perception [4].

Chromaticism in music is the use of notes foreign to the mode or diatonic scale upon which a composition is based,

applied in order to intensify or colour the melodic line or harmonic texture [5, 6]. The piano can only produce discrete frequencies of sound (12 frequencies per octave), and so the chromaticity of the piano is specified only in terms of unrelated to the specific scale notes. Consequently, in this case the concept of chroma coincides with the terminology of Western music.

Our research focuses on musical hearings from all around the world and therefore the western approach of chroma is not sufficient. Nonwestern music recordings (e.g., oriental or Byzantine) define different modes and sounds. Therefore, a general definition of “musical chroma” is needed as “chromatic” is defined as any sound with frequency irrelevant to the discrete frequencies of the scale. In proportion to the distance of the interval, that this sound creates with its “neighbours” (previous and next sound), it can be estimated *how much* chromatic this sound is.

Chrominance is defined as “the difference between a colour and a specified reference colour having a specified chromaticity and an equal luminance.” Our approach uses the intervallic nature of chroma, instead of Shepard’s circle. MEL-IRIS creates the chromatic index of a song, based on the intervals between each note and its “neighbours.” These intervals provide a chroma value (as it is further explained next), which defines the chrominance of each note of the song. The evaluated chroma of the scale of the song (“*Scale Chroma by Nature*”) is considered to be the reference chroma.

2.2. Purpose and Objectives. The purpose of this research is to provide the web of ethnomusicology with an index of background acoustic variability, using chroma as a *musical genus discriminator* [7]. Genuses or species [8] are not perceived as contemporary genres like hip-hop, disco, or jazz; therefore the chromatic index is not a genre discrimination method, like the one proposed by Tzanetakis and Cook [9] but mainly an intervallic systems classifier beyond the Western musical paradigm, introducing metrics for the progress of the concept of chromaticism. So, we would rather classify musical pieces into more or less “chromatic” genres. The approaches and definitions, used for chroma based music characterization, are described in the next four subsections.

Along with the “chromatic” classification, a way of visualizing music in a chromatic manner is proposed. This visualization comes from the meaning of the Greek word chroma (= color). So, if a way of representing chroma were acceptable, this would be in terms of colors. Many attempts to sound modeling with colors have been done, because correspondences between the physical dimensions of sound and colour [10–12] provide a review of auditory-visual associations as these have been investigated in computer music research and related areas.

In our proposal, an effort to correlate music chroma with musical emotion is necessary, since music and colors are associated with emotions. Music patterns have an effect on the listener’s emotions [13], who experiences them as contours of tension and release [14]. The melodic progression creates “chromatic” impression, which in analogy can

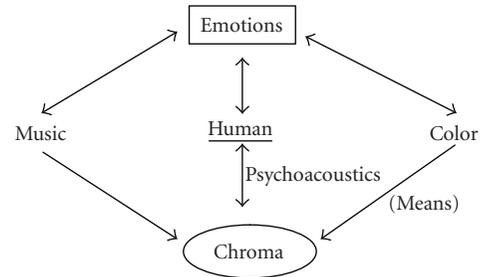


FIGURE 1: Music-Colors-Chroma correlation with human perception.

be depicted as a simultaneous “chromatic” progression. This progression, joint with psychoacoustics theories, results in the proposed chromatic graphs (see Section 11). A graphical representation of the relations between human, music, emotions, colors, and chroma can be seen in Figure 1. Colors are therefore used for the representation of emotions. Although measuring emotions continuously is open to criticism, Schubert has modeled perceived emotion with continuous musical features [15]. Also, Sheirer explored the extraction of psychoacoustic features related to music surface and their use for similarity judgements and high-level semantic descriptions [16]. “Chroma” can be pertained as a continuous musical psychoacoustic feature.

2.3. Definitions. The concept of “chromaticism” is strongly associated with our perception of music. Our research is focused on horizontal chromaticism (melodic line) [5]. In an effort to reach a definition that covers all the discussed aspects of “Chroma” in music we have concluded the following.

- (1) “*Chroma*” is mostly associated with the music intervals. The interval sizes may add an extra chromatic dimension in a specific melodic line. While a specific interval may cause a chromatic perception in a musical piece, in some other may not (Figure 2).
- (2) Each scale has an inherent chromaticity, which we call “**Scale Chroma by nature.**” “Scale Chroma by nature” is also associated with the intervals it contains and their distribution. This kind of chroma is not to be confused with each sound personal chroma. “Scale chroma by nature” provides the listener with the general perceptive impression that the scale gives in a musical piece. For example, the general impression of a major scale in a song is happiness (which does not mean that the same song cannot be momentarily sad), while a minor scale induces in general melancholy. This means that a minor scale has a different “scale chroma by nature” from a major one. We provide a way of measuring each scale’s “chroma by nature.”
- (3) The fundamental basis of the developed metric for “*Scale Chroma by nature*” is the occurrence of intervals other than whole tones (200 cents) and intervals that

are not multiples of whole tones within a scale. It returns to the original meaning of the scales some music uses being “chromatic” (rather than “diatonic” or “enharmonic”), extending this to allow scales to be more or less chromatic.

- (4) A sound in a musical piece is defined as “chromatic,” as far it is produced out of intervals, different from the ones, which the scale that the musical piece is written to contain. This definition also encompasses the occurrence of any sound with frequency irrelevant to the discrete frequencies of the scale in the musical piece. This is the normal definition of the term (stated in Section 2.1) and should not be confused with the definition of “Scale Chroma by nature.” A scale can contain itself chromatic notes and this may result in a greatest “Scale Chroma by nature,” but since they are taken into account in the Scale Chroma Measuring algorithm, they are considered notes of the established scale within the piece. So, if a note of this kind is encountered through the melody of the piece, it is no more chromatic in context, which means that it does not add any extra chroma to the one of the scale and, furthermore, of the musical piece in general (Figure 2).
- (5) In proportion to the distance of the interval, that a sound creates with the precedent one, it can be estimated how much chromatic this sound is. Certain metrics are proposed for this.
- (6) The progression of the melodic line of a musical piece results in a corresponding progression in the chromatic perception of the piece. Therefore, if we associate our proposed metric χ (chromatic index) with the human chromatic perception, we can represent the chromatic progression over time in either a two-dimensional graph or a “chromatic” strip, which is related to the notion of “colour” in music.

The approach that is followed in this research, in order to measure “chroma” in music, according to the previous definitions—explanations is what follows.

- (i) Each scale bears a definitely measurable “Scale Chroma by nature,” which is symbolized by χ° .
- (ii) Each piece can add or eliminate “chroma” due to the following factors:
- use of notes outside the established scale,
 - the melodic line progression,
 - rapidity (this factor affects only the saturation of a color, related to chroma, in the visualization process).

These 3 factors apply on the “chromatic index” χ of a musical piece.

2.4. Related Work. Several research efforts focus on the field of chroma-based music characterization as well as the mapping from music to colors in contemporary literature.

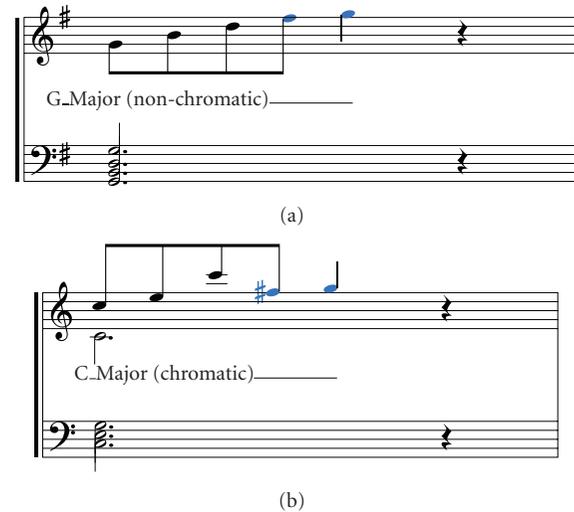


FIGURE 2: (a) The interval F#-G does not add any chroma ($\chi = \chi^\circ$) to the musical piece, which is written in G-Major ($\chi^\circ = 1.2857$). (b) The same interval adds more chroma ($\chi > \chi^\circ$) to the musical piece, which is written in C-Major ($\chi^\circ = 1.2857$).

In this section, we give references to literature and discuss related work in comparison to our research.

A very important issue in MIR research is *audio matching*, meaning the retrieval of similar songs to a given short query audio clip. Kurth and Muler suggest an efficient index-based audio matching procedure [17]. Their algorithm allows for identifying and retrieving musical audio clips irrespective of a specific interpretation or arrangement. Their approach uses advanced techniques for *audio identification* (also known as *audio fingerprinting*) [18]. Musical similarity can be tested by identifying cover songs. Therefore, Bello [19] uses approximate chord sequences from chroma features and hidden Markov Models.

Another interesting task in modern Digital Music Libraries (DMLs) is *audio thumbnailing*, that is the production of short representative samples of selections of music. Spectral features (MFCCs) that have been used in the past for speech processing [20] have been successfully applied by researchers in the area of audio thumbnailing. The problem of thumbnailing is also related to automatic audio segmentation tasks [21, 22]. Bartsch and Wakefield [23] use chroma-based representations for analysis, which includes segmentation and retrieval tasks.

All the aforementioned works extract chroma elements and create relations according to the *harmonic structure* of a music piece. In our work this music structure is not taken into account. Instead, we analyze the *melodic progression* of music by using the salient features. Moreover, the reader should not be confused with chroma features of the previously mentioned works, which use the chroma circle of Shepard (see Section 2.1) [3]. Our research deals with proposed metrics of chromaticism (as defined in Section 2.1) that go beyond the equal-tempered scale of western music and use a wider range of pitches, in order to be applied to the worldwide music, which is a very difficult task. A similar

calculation of chroma feature vectors to [23] is used for “Scale Matching” of the song (see Section 3.1), but the similarities stop there. The calculation of chromatic values χ and the visualization process is a unique approach for chroma-based music characterization. Another point that we should mention here is that identifying cover songs is not a good indicator of performance in our work. This derives from the different use of chroma as distinct pitch estimation in the western equal-tempered scale and our approach that takes into account pitch intervals. This means that another version of the same song may not be recognized as similar to the original, if the artist sings in a totally different idiosyncrasy from the other, which means that a totally different chromatic impression is created. Nevertheless, an effort to demonstrate performance gain on the task of identifying cover songs is evaluated in Section 11.2.2.

In the same context of using melody as an important descriptor of music, Marolt [24] suggests a unique mid-level representation with higher-level semantic features extracted from music signal for melody-based retrieval in audio collections. Marolt estimates salient melodic lines, based on Klapuri’s work [25]. In our work, we use melody (or better the predominant pitches that the listener perceives) as a chromatic descriptor of music.

Several approaches have dealt with the *visualization* of musical content. Smith and Williams [26] describe a method of visualizing music with colors in 3D space. The problem of mapping from music to colors is both theoretically and practically important. Several efforts for mapping music to colors and moods have been done [27]. The mapping and classification in these cases take place using metadata attributes (as mood or genre) that can be obtained from music services. The use of colors in music interfaces has been applied in several encodings. For example, in [28], the authors present a music interface that maps colors with different types of music. In our 2D visualization proposal, colors have a semantic meaning and are produced according to algorithms that calculate the chromatic perception of the song. This allows music analysis in a more intuitionistic way—a chromatic way.

2.5. The Chromatic Analysis Process. The procedure of the chromatic analysis is serial and consists of five steps:

- (1) extraction of melodic sequence (frequencies),
- (2) scale Matching,
- (3) segmentation,
- (4) calculation of chromatic values,
- (5) creation of color strip.

These steps will be further explained in Section B. The reader can also refer to [1, 2] for further discussion on the alphabet (symbols) used and the tasks.

3. The Algorithms

3.1. The Scale Match Algorithm. It is really necessary to know in which scale the musical piece being analyzed is

TABLE 1: Intervals calibration.

<i>Interval</i>	<i>c (cents)</i>
Tone	200
Half-tone	100
Quarter-tone	50
3/2-tone	300
2 tones	400
...	...

written, because its chroma χ is used as a benchmark. The algorithm is described here briefly. The algorithm scans the whole sequence of frequencies, which resulted from melody extraction and writes down how many times each note of the melody was being played on a space of an octave. From the notes most frequently played, it fetches which note spaces predominate on the musical piece. Sliding the values of cents 6 times (one interval at a time) it creates 7 possible modes. If one of them matches perfectly with a mode of the “Scale Bank,” the melody automatically corresponds to that mode (“Scale Bank” is a database, which contains scales and modes, each of them expressed in terms of its individual attributes). “Scale Bank” contains at the moment about 100 scales, which can be classified in categories like: Western, Oriental, Byzantine, and so forth. The database structure supports the enrichment with more “unusual” scales and modes, like African or Thai scales [29].

If there is not an exact match, the closest mode to the interval sequence is considered. Essentially, the scale that yields the minimum error rate (calculated from the absolute values of the differences between the several combinations of spaces) is chosen.

3.2. The Scale Chroma Algorithm. Each scale bears a “Scale chroma by nature” value χ^0 , which is very important for the chromatic analysis. This value is representative of the chromaticity of a scale and is used as the reference at the calculation of the chrominance of each note in the song. As it will be explained later, each value χ corresponds to a real color. So, the predominant color on the visualization of a musical piece is the one that corresponds to its χ value.

An algorithm for finding a value as a metric of chroma in a specific scale is the following.

Algorithm 1. Let a tone correspond to 200 cents, as usual. Table 1 shows the cents of each interval.

- (i) For each interval i in the scale calculate K_i :

$$K_i = 200/c \quad c \leq 200, \quad (1)$$

$$K_i = -0.0002 * c^2 + 0.12 * c - 15 \quad 200 \leq c \leq 400,$$

where c is the interval cents.

- (ii) “Scale Chroma by Nature” is equal to

$$\chi^\circ = \frac{[(\sum_{i=1}^n K_i) + j]}{n}, \quad (2)$$

where n is the number of whole tone steps in the scale (number of notes $- 1$), and j is the amount of the extra accidentals on the scale notation, different from the accidentals at the key signature.

The chroma of C Major Scale is 1.286, while the chroma of C Minor Melodic Scale is 1.571 according to the calculation of the algorithm. Examples of Scale Chroma Calculation can be found in [1].

3.3. The Segmentation Algorithm. There is a great deal of considerations and approaches about the question of melody segmentation. The segmentation of the melody is absolutely necessary in our research, in order to process each segment solely for finding chromatic elements. The ideal way of dealing with a music file would be to process it in a whole, calculating and representing chromatic values for each frequency of the melody, but that would mean a very time-consuming and computational resources—expending process, since an FFT algorithm on a WAV file results in thousands of values, which would need a big storage space in the Hard Disk, along with all the results of the analysis. Segments are considered as complete musical phrases within the context of a song and their subjective analysis is rather wiser, since it is the whole phrase that creates a chromatic impression in human mind, not one note by itself.

Many segmentation suggestions have been used with satisfactory results. The Cambouropoulos-Widmer segmentation algorithm [30] resulted in much a similar way to our arbitrary segmentation, based on a sentient perception of the melody. Therefore, this approach has been preferred for MIDI files. Our standard segmentation method makes use of some possibilities rules from Lerdahl and Jackendoff template [31]. Heuristic rules are also used. These rules apply some constrains on the way the algorithm splits the melodic sequence of a song into segments. Some of them are presented in [2].

4. Drawing Colorful Strips

The initial chroma of a musical piece (χ^0) is the “Scale Chroma by Nature.” According to the sequence of frequencies, which was the output of the first step, each space affects the current χ value (possible increment or reduction), creating this way a continuous chromatic χ value of the musical piece [2].

The amount of χ values is equal to the amount of the notes, which comprise the melody. These χ values produce the final colorful strip. Colorful strips characterize music according to chromatic variations. This chromatic visualization consists of boxes that represent the segments. Each box represents a segment. The length of a box is proportional to the duration of the segment it represents. This results in the real-time lengthways creation of the chromatic strip. The basic color of a segment is the average $\langle \chi \rangle$ of the χ values that correspond to all the notes of the particular segment. As the creation of a box comes near to an end, the basic color changes in order to achieve a smooth transition to the basic color of the next segment.

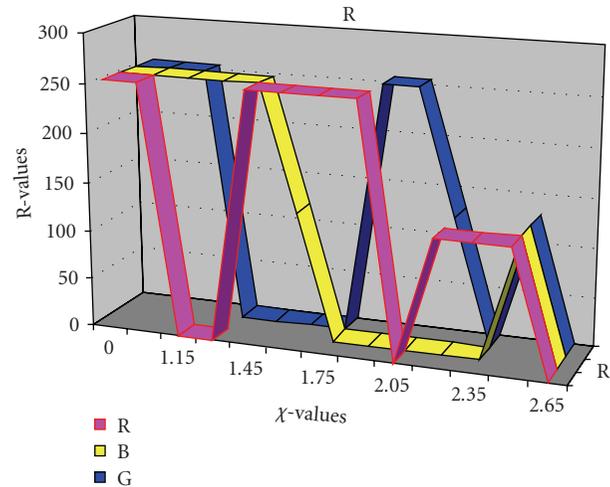


FIGURE 3: R–G–B variables’ variation, based on χ value.

A 12-grade color scale was designed to correspond χ values to colors. Colors are ranged in chromatical order, beginning from white and ending to black [2].

The actual color of each segment is characterized from the combination of the R–G–B variables (Red–Green–Blue). The values of R–G–B for a particular segment are calculated from linear quotations and their graphical representation can be seen in Figure 3.

A musical piece starts with the color, which is defined by the position $(\chi^0, 0)$ of x -axis, as its basic chroma. Since the variable x affects chroma, spot A moves either rightward on x -axis (more chromatic), or leftward on x -axis (less chromatic). Figure 4 shows that if a melody moves higher on y -axis at some color strip, the saturation of the color is increased, while if it moves lower on y -axis, saturation is reduced. This happens because, since rapidity affects the y -axis, the faster a musical piece is, the brighter feeling it provokes. (Imagine a cheerful song in very slow tempo!)

The choice of color on the final representation graph takes place per segment. All values of y in a segment produce the average $\langle y \rangle$, which is the global y -coordinate of the segment. Similarly, all values of χ produce the average $\langle \chi \rangle$ of the segment.

Section B: The MEL-IRIS Tool

In this section we present MEL-IRIS and its capabilities. MEL-IRIS was mainly developed in Borland C++ Builder 6 and uses the MS-SQL database.

5. In General

MEL-IRIS stand-alone PC version is now on version 2.0. The second version of MEL-IRIS is explicitly improved with regard to the previous version. There are improvements both in design and the kernel of the application in terms of musical pieces analysis. Apart from the new algorithms,

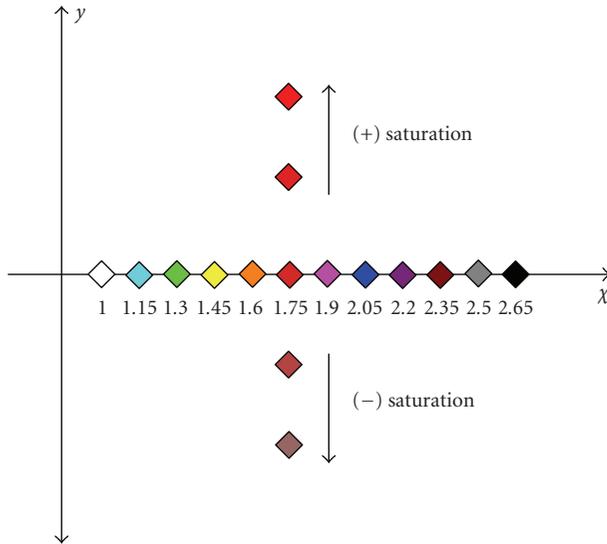


FIGURE 4: A colorful 2-dimensional surface for visualizing the Chromatic Index.

the old algorithms have been redesigned in order to succeed a more effective and correct approach of chrominance in music.

The application's design was lionized in order MEL-IRIS to turn into an easy-to-use program, without remising its substantial proposal in computer music. A simple user-auditor can treat MEL-IRIS as his/her default Media Player, with which he/she can organize and monitor all his/her music collections in an easy and effective way, and also listen to the music while watching this prototypal way of music visualization. On the other hand, the music analyst/composer/researcher can work with MEL-IRIS in his music analyses, since it provides an integrated environment for chromatic analysis and extracts useful statistics from large music collections. *Figure 5 shows the inner structure of the application.*

MEL-IRIS supports directly the following file formats: .wav, .MP3, and MIDI. If the input to the application is an MP3 file, this will automatically be converted to a .wav file. Then, it will be analyzed as a wave file and finally this intermediate file will be erased from disk. The reason for that is that the AUDIO analysis in this version is only invoked in wave files.

MEL-IRIS stores all the data in the "MELIRIS" database, which is structured on Microsoft SQL server.

The diagram in Figure 6 shows analytically the whole chromatic analysis process of a wave file from MEL-IRIS. All the intermediate produced results can be seen in the diagram. The legend shows what each shape stands for. An observation here is that the text files, which are marked as XML, are those that contain analysis data that are finally exported in XML format. All the intermediate files are automatically deleted after the end of the analysis (however their data can be retrieved, since they are stored in the database).

Initially, an FFT (Fast Fourier Transform) is applied on the .wav file. The FFT algorithm is followed by an embedded

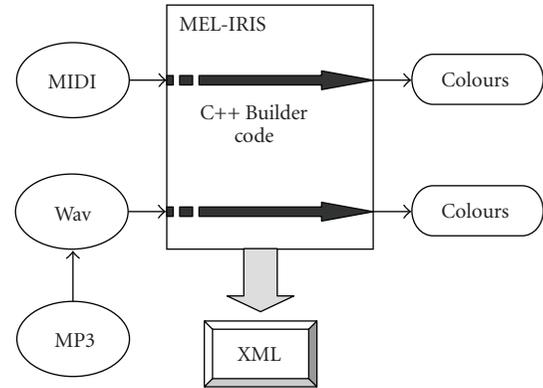


FIGURE 5: MEL-IRIS inner structure.

process for predominant frequencies extraction, so as to arise a table, which (in a satisfactory percentage) contains the fundamental frequencies of the piece's melody. Since pitch is normally defined as the fundamental frequency of a sound, this process executes melody pitch tracking. The resulting values are stored in the files conv1.mel (frequency values) and times1.txt (duration of the frequencies in milliseconds). In these two files are afterwards 3 different algorithms applied: (1) the *Scale Match* algorithm corresponds the melody to a scale and, in extension, determines the "Scale Chroma by Nature" χ^0 , (2) the *Segmentation* algorithm fragments the musical piece and produces the file segments.mel, which determines how many consecutive notes comprise a segment, and (3) the *Mini Conv* algorithm. The latter condenses the files conv1.mel and times1.mel, based on some rules, because of the fact that the initial sampling of the FFT is applied on very short intervals. For example, if we take a sample every 200 milliseconds and the note A4 lasts 2 seconds, the result of the sampling would be ten A4 notes, which is not correct (the default sampling in MEL-IRIS is applied every 20 milliseconds, which is a typical frame length in sound analysis). The algorithm diminishes the melodic sequence by producing two new files: conv.mel and times.mel, as it can be seen in Table 2.

The new conv.mel as well as segments.mel are inputs to the algorithm of *chromatic elements mining* (a series of rules and controls regarding the chromatic variation), which results in the χ sequence (how the chromatic index is altered during the melodic progression over time) in the file x.txt and the avg.mel, which contains the chromatic average values for each segment. Finally, sample.mel contains the appropriate information, which is needed for the visual chromatic diagrams to be produced on screen.

Some discussion is needed here for explaining the reason why the *Mini Conv* algorithm is not directly applied on the initial files. This happens because the *Scale Match* algorithm takes into account the number of the occurrences of each melody's pitch and gives priority to the notes with the most occurrences. If *Mini Conv* had been applied to the example of the previous figure, the *Scale Match* algorithm would calibrate the A4 note as 2, which means a very slight superiority over the other two notes that would bear 1 as the

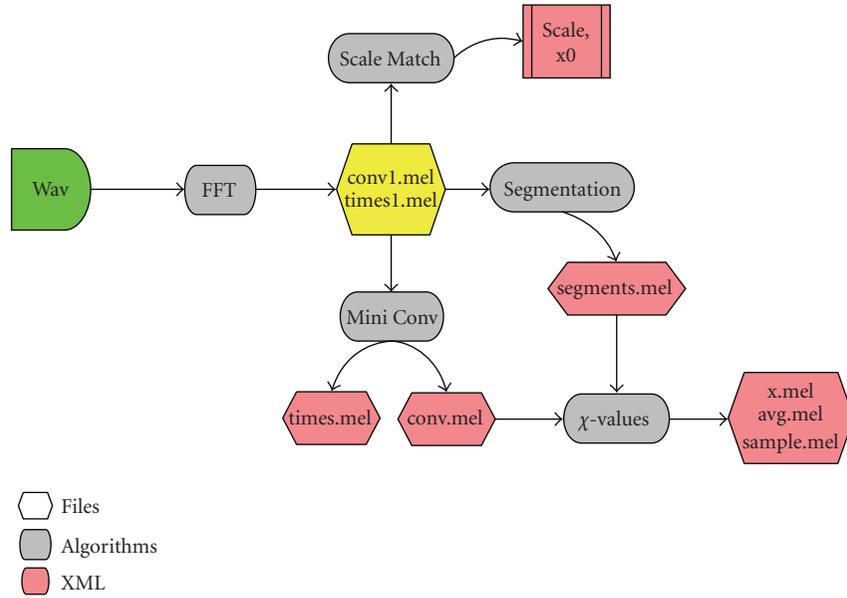


FIGURE 6: Implementation of Chromatic Analysis in MEL-IRIS.

TABLE 2: Transforming melodic elements with *Mini Conv*.

CONV1	TIMES1			CONV		TIMES
	Start Time	End Time		Start Time	End Time	
440	0	200		440	0	2000
440	200	400		356	2000	2600
440	400	600		412	2600	2800
440	600	800		440	2800	3200
440	800	1000				
440	1000	1200				
440	1200	1400	→			
440	1400	1600	MINI			
440	1600	1800	CONV			
440	1800	2000	→			
356	2000	2200				
356	2200	2400				
356	2400	2600				
412	2600	2800				
440	2800	3000				
440	3000	3200				

number of their occurrences. However, this is not the fact. A4 may really have two occurrences on this example, but the sound lasted longer than the other notes. The real weights can be seen in Figure 7.

6. The Stand-Alone Version of MEL-IRIS

This stand-alone version 2.0 of MEL-IRIS provides a completely different interface from the network version, since it serves yet equally as a media player and as a music analyzer as well, with full functionality in contrast to the network

version, where the function of analysis takes place in the server. The following snapshot (Figure 8) shows the graphical interface of the stand-alone version.

As it can be seen in this snapshot, the main screen of the application is oriented towards the visual representation of music and not the numbers—song data. The figure shows five frames of sound representation (namely; spectrogram, waveform, frequency-amplitude graph, volume counter, and the depiction of chromatic analysis) and a list, which contains the pieces of the music collection, on the right side of the window.

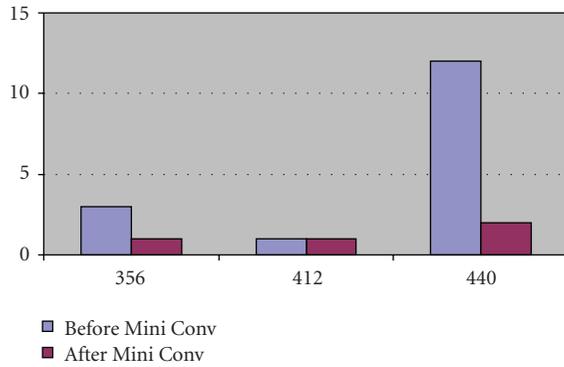


FIGURE 7: The *Mini-Conv* algorithm effect.

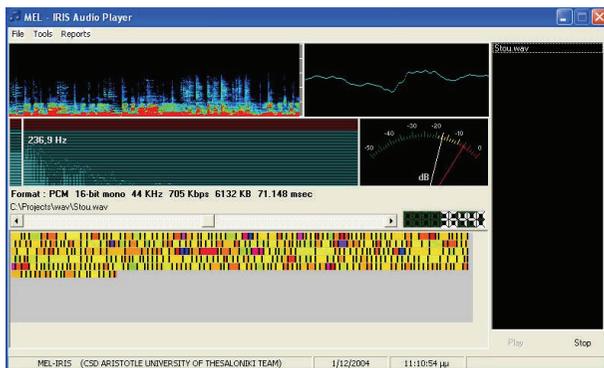


FIGURE 8: The Interface of MEL-IRIS 2.0.

7. Analyzing Single Audio Files or Collections

In MEL-IRIS v.2.0., there is the capability of either opening and analyzing one single file or opening a songs folder for batch processing.

In the case of choosing a single file opening, the following procedures are executed: (a) the database is checked for previous analysis of the same file, (b) if the previous result is negative, a background analysis is initiated (the process is not visible to the user), (c) the results of chromatic analysis are automatically stored into the database, and (d) the musical piece is executed by the player.

The user may also choose a folder, which contains all the songs he/she wants to analyze (e.g., 100 mp3 files, 50 wav files, and 15 MIDI files). The application initiates then the background analysis of all the songs one by one and produces at the end total statistics of the whole collection as well as statistics of each single song (see Section 8).

8. Statistics

As a music analyzer, MEL-IRIS can produce statistics from the analysis that are useful to researchers and musicians. The statistics can be related either to one single file or to full song collections.

General Statistics. The following attributes appear for each song: corresponding scale, χ^0 of the scale, the lowest χ value

that was found in the piece, the highest χ value, the deviations of χ^0 in terms of the lowest and the greatest χ value, and the song's chromatic average.

Chromatic Proportion. It is about a graph that shows in which analogy the 12 basic colors have been detected in a particular piece. It can be viewed in two formats: bar chart or pie chart.

Chromatic Graph. It is a bidimensional graph where x -axis represents the time and y -axis represents the chromatic χ value. It performs the chromatic evolution of the melody in terms of time.

Distribution Graph. It is also a bidimensional graph where x -axis represents χ and y -axis represents the amount of times a χ value appeared in the piece (i.e., the notes that correspond to that χ value). For example, if a musical piece does not change tonality and scale at every turn, its normal distribution graph would be a curve (more like a normal curve) with the highest peak on a value near χ^0 , which is the value of the scale chrominance and is used as a benchmark for this song.

Queries. After the elements of chromatic analysis have been registered in the database, the user is given the option to compose some queries that return statistics (see Section 11).

9. Color Assignment

Our intention from the beginning of the research on musical chroma is the finding of a common counting denominator in a worldwide scale, that is, in musical hearings, scales and modes from the whole world, and not exclusively from western music. The correspondence of music to colors and feelings, however, cannot be absolute. The reason for that is the existence of broad racial differences: various ways of thinking and cultures, for example, in some countries purple is considered as a sepulchral color, while for some other countries the color of mourning is black, and for some others it is white!

One could possibly correlate the chromatic analysis of music with current attempts to discover the meaning of words in a text through statistical analysis [32], that is, the definition of the style of an author in terms of a statistical analysis of word usage. A reader or an audient brings a vast knowledge of other works to help him perceive the current work's structure. Listening to a musical piece triggers reactions unique to each individual, according to his previous music experiences and cultural differentiations. This is the reason why MEL-IRIS is designed to offer flexible variable settings.

MEL-IRIS provides the user the option to recreate the colors-feelings correspondence all by himself, according to his perception. As it is described in theoretical background, there is a default 12-grade series of colors, which ranks 12 basic colors starting with the less chromatic (grade 1—white) and ending at the most chromatic (grade 12—black). The

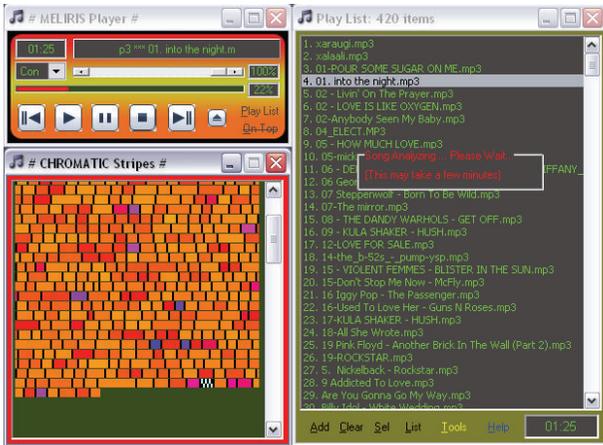


FIGURE 9: The Interface of the client application (MEL-IRIS v.3.0.).

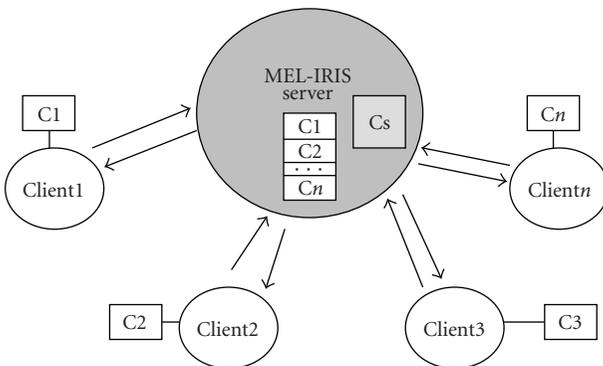


FIGURE 10: MEL-IRIS client/server architecture.

user may not agree to that sequence. On this account, MEL-IRIS is parameterized in a way that it provides the user the freedom to rank the colors anyway he likes. A proper algorithm runs on the background in order to produce the operators, which satisfy the new rank. The results reflect on the chromatic graphs.

Using the “Color Assignment” option, the user faces a simple interface with twelve colorful boxes that represent the basic colors. Any of these can be chosen and moved into one of the blank sequence positions. That way, the new chromatic distribution is created and is ready to be used for chromatic analysis by MEL-IRIS.

10. The Online Application

The online version of MEL-IRIS (<http://nomos.csd.auth.gr:8080/meliris>) is the one that allows the music collection and analysis in audio servers and may be used for audio indexing (Figure 9), classification, search, and retrieval in broadcasting. The client application communicates with the server, where the whole process of chromatic analysis takes place (Figure 10).

Each user has full access to audience and statistics of all the songs that he/she uploads to the server. The client application is easy-to-use and disposes the attributes of the

stable stand-alone PC version, except that the chromatic analysis is not applied locally.

The collection and the statistic analysis of all the songs in the central database of MELIRIS will be soon available on the Internet for all the users. Every user (after registration) occupies his own space on our server in order to upload and analyze his own songs.

10.1. MEL-IRIS Network Architecture. The MEL-IRIS system’s architecture is shown in Figure 10. The main mechanism of statistical analysis and storage exists in a central server, to which several client machines are connected. The server contains the main database of the system as well as numerous stored MP3 files (Cs—Server (Music) Collection). An easy-to-use graphical interface and a smaller (in terms of capacity) local database (C1—Cn) are installed on each client.

Authorized users may take advantage of the several functionalities of MEL-IRIS from the client machines. To begin with, users can listen to the musical pieces that are already stored in the server and/or watch their visualization and examine the results of the chromatic analysis. Clients have access to all the pieces in the music collection of the server. Moreover, they are able to load their own music files either only for listening to (audio player function) or for chromatic analysis. In the second case, the extracted statistics are stored both in the local client database and the central server. This means that if client 1 analyzes the file x.mp3, then client 2 is also able to retrieve the corresponding statistics from his terminal and listen to the piece, since this exists in the server and can be downloaded in his/her personal computer. This is the case where client 1 has uploaded his music for analysis on the server. The gathering of the statistics, which are the result of the analyses of all the clients, aims at a massive data collection for further processing and data mining. These data are accessible from all the users of the system. This means that each user may choose any number of pieces, which have been chromatically analyzed, as a sample for his/her research. Moreover, each user can create locally his/her own profile, so as to interfere on the variable attributes of chromatic perception, for example, the colors-emotions correspondence (see Section 9) and the way of the final results visualization. Finally, each user is able to use the MIR (Music Information Retrieval) functions of the systems, through queries to the database.

11. Experimental Results

11.1. Method. Initially, we tested the analysis tool on the server of Multimedia Lab of the University, which hosts over 2 millions (classified by genre) songs, in MP3 format, from all over the world. 16 students (which we consider as a satisfactory number for our primary results) aging 18–25 years old helped the research with their opinions (10 students of Computer Science, 6 students of Music). They were asked to create their own color sequences (in case they didn’t accept our default sequence) and afterwards run 10 songs of their choice on the server.

TABLE 3: Summarized Statistics of the Sample.

GENRE	Number of songs	Ch. Avg
Classic	75	2,413
Dance	7	1,888
Hip Hop	44	1,890
Metal	13	1,856
Pop	51	2,127
Rock	15	1,869
Greek Old Classics	15	2,517
Greek Traditional	4	2,232
Rebetico	53	2,085
Instrumental	50	2,083
Ecclesiastical	47	2,196
Ethnic	33	2,061
<i>TOTAL</i>	407	2,2

We will present here the results from the chromatic analysis of a sample of 407 MP3 songs that were randomly chosen and initially scattered in 12 genres. We will pick up a few examples, in order to demonstrate some fundamental functions in our chroma-based MIR system using colors. We will (i) examine the common attributes of the songs in each of the categories mentioned before: (ii) apply filtering queries on the sample, and (iii) try to detect similarities (from a psychoacoustic and a musicological perspective) in songs that resulted in congener chromatic graphs.

The full experimental results are available for further exploration at <http://www.csd.auth.gr/~dpolitis/MEL-IRIS>.

11.2. Results

11.2.1. Experimental Results. Table 3 shows summarized statistics for each of the 12 genres of the sample. Although some genres contain only a few songs and are therefore not recommended for general conclusions, we can make some observations on the numbers.

Taking into account the genres that contain over 30 songs, we can observe that the most chromatic genre is classical music. In Figure 11, we can see that there is a great variety of chromaticism in the classical songs of the sample. In contrast, the hip hop genre (the less chromatic from the considered genres) shows no such variation with the most of the tested songs belonging to an orange tint of about $\chi = 1,6$. This is normal, because hip-hop music is more rhythmic and less (or not at all) melodic and creates static little chromatic impression to the audient. Figure 11 also shows the songs distribution of ecclesiastical chants, which is a very chromatic genre. We can note here that it was the only genre, where chromatic averages greater than 3,5 appeared (with an exception of a 3,7 occurrence in classical music).

11.2.2. Queries. Figure 12 shows the results of the query “Find the top-3 songs whose total average is 1,3.” The query

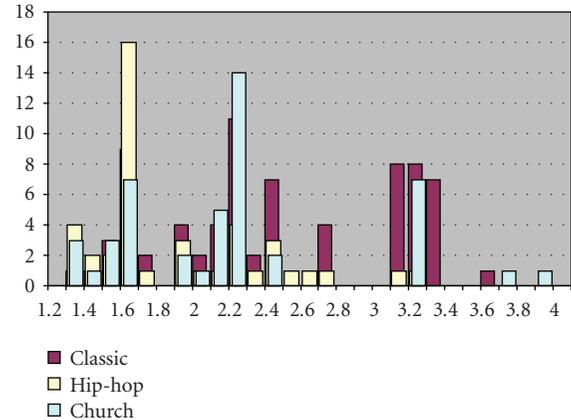


FIGURE 11: Songs distribution from the chromatic analysis.

ROWID	Song Name	Average	Song Destination
1	12-SA TRINCHA _ SMELL OF PARADISE.i	1,29	D:\MP3\METHNI
2	05-ERIC FERNANDEZ _ ADIOS ANDALUC	1,29	D:\MP3\METHNI
3	14-SARAH BRIGHTMAN _ HOW FARE TH	1,29	D:\MP3\METHNI

FIGURE 12: MEL-IRIS queries.

returned the following songs:

- (i) Smell of Paradise (Sa Trincha),
- (ii) Adios Andalucia (Eric Fernandez),
- (iii) How fare this spot (Sarah Brightman).

A second demonstration query on our sample was “Find the 3 most similar songs to the song “How you gonna see me now” (Alice Cooper).” The number one returned song (the best match) was “The trooper” (Iron Maided), which also belongs to Metal songs. Figure 13 presents the chromatic graphs of the two songs, while their colours distribution appears in Figure 14. The similarity is obvious.

Another approach of finding similar music, and not the exact match of a sequence of notes, has been presented by Ghias et al. [33]. The proposed method was to retrieve music based on the relative pitch changes of a sequence of notes. The chromatic approach (presented here) discovers similar colour patterns, which are resulted from the sequence of notes. The chromatic method is more general and allows greater freedom, since two completely different notes sequences may result to a similar colour sequence, and furthermore to similar feelings and mood.

The retrieval method for similar patterns in MEL-IRIS is based on Euclidean distance as a metric, regarding the colors distribution of the analyzed musical pieces. Other difference metrics in color-content-based retrieval systems are histogram intersection and histogram cross distance function [34].

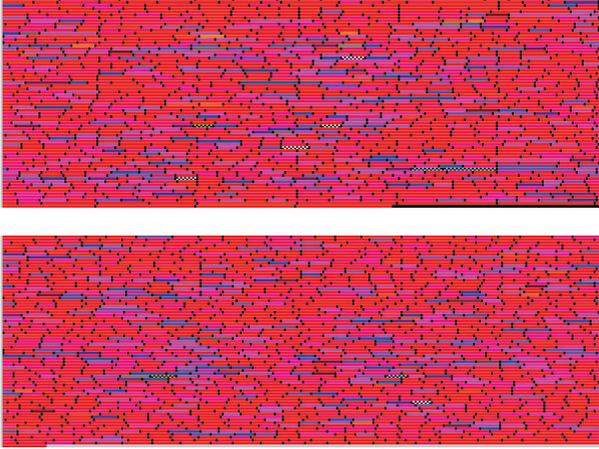


FIGURE 13: Similarities in the same genre: Alice Cooper versus Iron Maiden.

The graphs in Figure 15 show the chromatic progression of χ in all the three metal songs returned from the query. The value range is almost identical (the third song is “Laguna Sunrise”—Black Sabbath).

A crucial query type of the application is the *query by example scenario*. The user may provide a sample of music (a song clip in MIDI or audio format) as input to the system. The system responds by chromatically analyzing the sample (just as any other song). After the analysis, possible results for such a query to be scenario are songs that are similar to the colour ratios of the sample, songs that are closest to the chromatic average of the sample, and also identical patterns to the sample in other songs (especially for short sound clips as samples).

The important fact here is that after the chromatic analysis is finished, the final colour representation is stored as a vector of (χ_i, D, B) elements, where each element represents a segment. The variable χ_i stands for the chromatic index of the segment, (where i is the number of the segment) while D and B stand for Duration and Brightness of the segment, respectively. Such a representation allows the efficient use of known string matching techniques for similarity matters.

In order to evaluate the performance of the system in identifying similar songs, we used the task of identifying *cover songs*. Cover songs are different versions of the same song usually by different artists. For that reason, we enriched our database with 7 cover versions of the song “All by myself” and 6 cover versions of “Somewhere over the rainbow.” The particular songs were chosen in order to compare the results with Marolt’s and Ellis’s works [24, 35]. As in [24], we evaluated retrieval accuracy by querying the database with each of the 13 cover songs and calculating average precision as well as counting the numbers of relevant songs returned in the Top-10 hit list. The formula for average precision (AP) is

$$AP = \frac{1}{R} \sum_{j=1}^H \frac{\text{rel}(j)}{j} \sum_{k=1}^j \text{rel}(k), \quad (3)$$

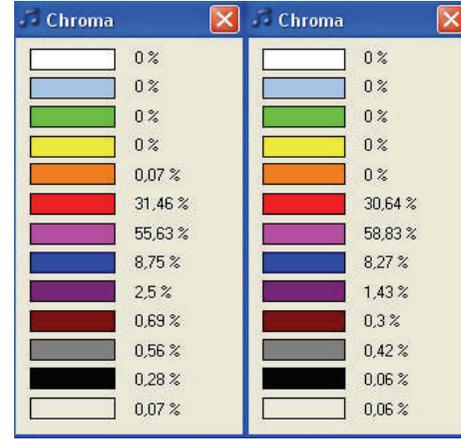


FIGURE 14: Similarities in the percentages of colours.

where H is the number of hits returned, R the number of relevant hits returned, and $\text{rel}(k)$ a binary function returning 1, if the k th entry in the hit list is relevant. The comparison of per-song results can be seen in Table 4.

Although MELIRIS does not intend to discover cover songs, it seems that the results are satisfactory from this experiment. A straight comparison to the other two systems is not applicable (since the methods are completely different and aim at different musical features), but we can see that in the case of “Over the Rainbow” the chromatic interpretation proved to be more accurate.

“All by myself” had smaller differentiation in its different versions with a chromatic average of 2.102 (or royal blue). The strongest feelings appeared in the translated opera version “Solo Otra Vez” by Il Divo. “Over the Rainbow” showed several variances in chroma with an average of 1.913 (or pink). The most chromatic version proved to be that of Criss Alen with emotional changes of convulsion, intensity, and sadness that reached melancholy (darker color grades). Israel Kamakawiwo’s voice caused a depressive result (around purple), while Jason Castro had his own playful interpretation. These differentiations that come from the artists and the music genre of each version cause the chromatic analysis to characterize in a different way each song and therefore difficult to accept them as similar.

Another fact (in order for chromatic indexing to be better understood) is that chromatic perception is that it is not directly related to genres. This means that a melancholic pop song is closer to a melancholic rock song than a cheerful pop one. This is only confirmed by objective user tests. Nevertheless, we also provide some subjective results on the songs sample. Table 5 shows the average number of 5, 10, or 20 closest songs with the same genre or artist as the seed song. We took into account only the genres that contained over 30 songs.

Of course, a user may not be familiar to the concept of chroma and is not able to use any of the query types that were described. Apparently, a more suitable approach would be the use of queries like “which are the 10 most melancholic

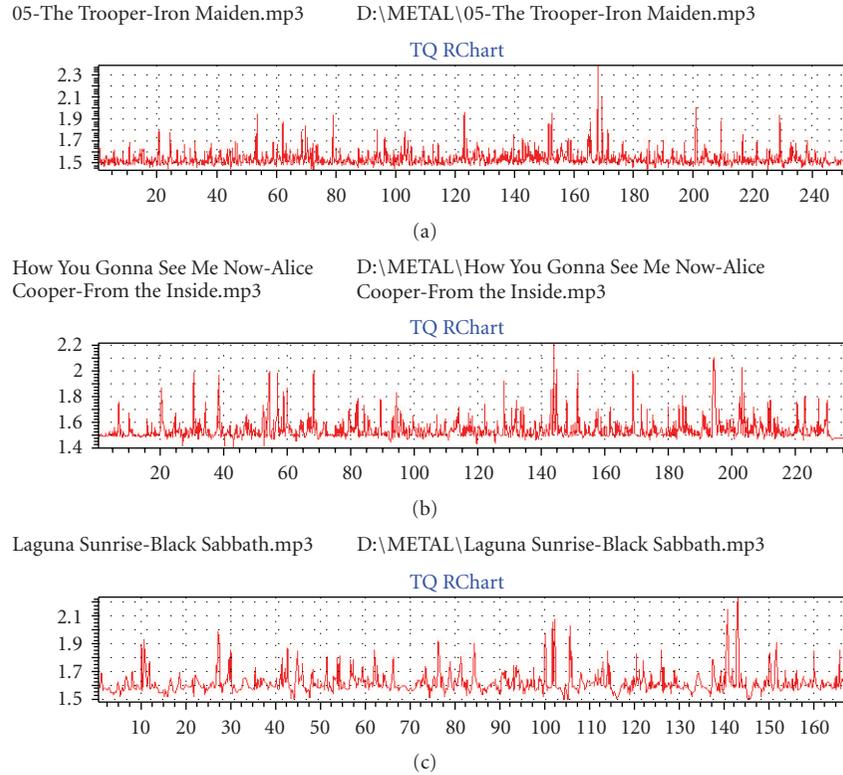


FIGURE 15: Representative chromatic graphs (genre: metal).

TABLE 4: Performance Evaluation.

Song	Our approach		Marolt		Ellis	
	mean AP	top 10	mean AP	top 10	mean AP	top 10
All by myself	0.82	6	1	12	0.53	5
Over the Rainbow	0.55	4	0.15	5	0.13	2

TABLE 5: Average number of closest songs with the same genre or the same artist as the seed song.

Closest to the seed	Avg—Same Genre	Avg—Same Artist
Closest 5	2.12	1.38
Closest 10	3.71	2.67
Closest 20	7.21	4.54

pieces in the database?”. The system supports such a kind of queries by translating each feeling (which is selected from a list) to its chromatic value equivalent and matches that value to the closest chromatic averages of the songs. In this case the user should be aware that by choosing a feeling, he substantially selects a whole cluster of feelings. For example, by demanding “*ten passionate songs*,” he actually retrieves 10 songs that cause the feelings of anger, evil, jealousy, and love, which correspond to the red color. The user is informed by a message for that choice.

In the tested sample the statistical averages state that 78% of the songs in the same genre (e.g., pop, rock) showed the same average chroma on their graph in a percentage of about

59%. For example, 70 songs in our sample belong in the rock-ballad genre. 74% out of them had pink as their average color in a percentage of 60%.

11.2.3. Similarity Patterns. Among the several observations on the chromatic graphs is that several similarities between song’s patterns occurred. The similarities were obvious from the color and the size of the segments. This is absolutely natural and predictable, since a song usually has both identical repetitions (musical phrase repetition—“*physical refrain*”) and slightly varied repetition (musical phrase variation—“*fake refrain*”).

In Figure 16, we can see part of the chromatic graph of “Say something” (James). The checked 3 lines comprise part of the refrain. Watching carefully on the first and third lines, we can observe two exactly similar patterns.

11.3. Music Indexing. The chromatic average of each song that is analyzed with MEL-IRIS may represent a unique way of music indexing with many possible uses. First of all, it ranges the songs into chromatic categories (and furthermore emotional representations). Thus, songs play

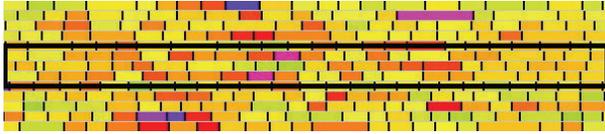


FIGURE 16: Similar patterns on “Say Something” (James).

lists of a category may be automatically created. For instance, asking for a play list of “melancholic” music ($\chi = 2.2$) from our sample database, MEL-IRIS returned:

- (1) “*H parthenos shmeron*” (Petros Gaitanos—Ecclesiastical music),
- (2) “*Ymnos astrapis*” (Vangelis Spanoudanis—Instrumental music),
- (3) “*To alithinon fos*” (Petros Gaitanos—Ecclesiastical music),
- (4) “*Silver Inches*” (ENYA—Ethnic music),
- (5) “*Megalinion fimi mou*” (Petros Gaitanos—Ecclesiastical music).

Moreover, a chroma-based music characterization for songs in a huge music collection is the following: Very Low Chromatic (VLC), Low Chromatic (LC), Medium Chromatic (MC), High Chromatic (HC), and Very High Chromatic (VHC) [2].

Our data collection can be presented in a 2-dimensional space with all the songs as points (Figure 17). The two factors affecting their indexing are their average chromatic index and the deviation of χ during the melodic progress of each song. The greater this deviation is, the more colours, feelings, and chromatic fluctuations the song has. As it is noted in the figure, broad categories are obvious from the points.

These categories are broad and may contain songs that are characterized, for example, as “low chromatic” that are close but are not totally the same in mood movement and patterns. The following graph (Figure 18) shows the chromatic distribution of 3 songs of the same category. We could say that “Let it be” and “When I need you” are much alike in their color distribution, while the Greek hip-hop song “Se alli diastasi” has a greater peak in 1,4 (shine) and lower harmony affect (1,3). MEL-IRIS approaches this way the following MIR tasks: mood, emotion, and style [36].

Finally, the chromatic indices may characterize certain music albums or artists. For example, in our sample, it came to light that 16 songs of Chris de Burgh had all 1, 49 as chromatic average! This comprises a kind of chromatic music identity of this artist and can discriminate him from somebody else, who will perform the same piece (with a different chromatic fluctuation).

The discussed topics are of essence because no manual addition of any metadata is necessary in order to extract these results, since the whole process of audio analysis takes place automatically.

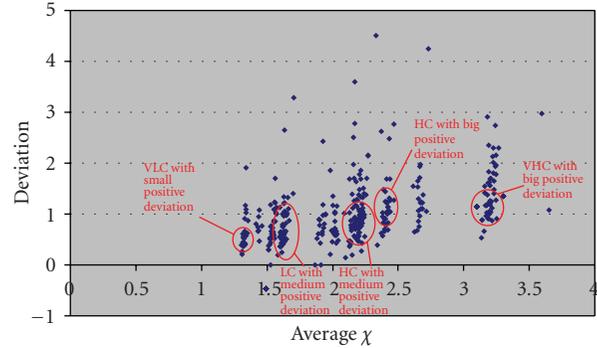


FIGURE 17: The songs of our sample collection on a 2D space create obvious similarity groups.

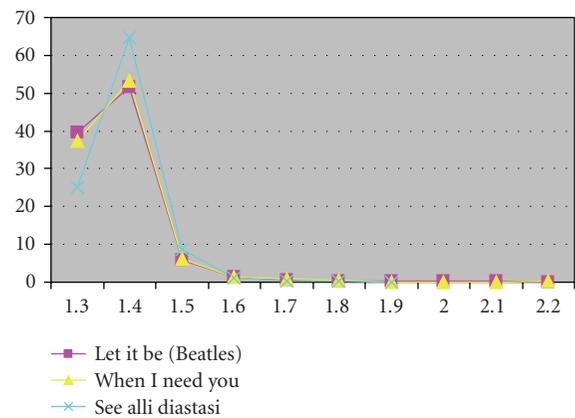


FIGURE 18: Comparison of chromatic distribution between 3 similar songs with the same chromatic average.

12. Future Directions

Very soon, the internet gate of MEL-IRIS will be enriched in order to offer all the advantages of the application to its visitors. By visiting the site, the user will be able to apply all the well-known functions of huge Digital Music Libraries (e.g., playlists, retrieval tasks online, etc.).

Moreover, we will seriously take into account the feedback from the online use of MEL-IRIS and the experimental activities of the Multimedia Lab students in Computer Science Department of the Aristotle University of Thessaloniki, Greece, for the improvement and optimization of the application.

Together with the improvement of our chroma-based music characterization process, one of our goals is also the enrichment of the MEL-IRIS media player with splendid 3d color visualizations and also the design of more complicated queries, which will be useful for academics, researchers, musicologists, and people in music industry.

13. Conclusions

Music Information Retrieval tasks require flexible systems with semantic visualizations, new metrics, and multiple ways of implementation.

MEL-IRIS is designed upon the unique idea of chromaticism, which allows many possible uses in this area: search, classification, similarity patterns finding, indexing, digital signatures, and visualization of sound. An approach on human perception of music leads to emotional music depiction through the sound to image transformation.

Two versions of the software are available at this time: MELIRIS V.2.0 (Stand-Alone PC version) and MELIRIS V.3.0 (Internet Edition). The first one has full functionality on a personal computer, while the second one suits remote access to a central processor-database which stores and organizes huge collections of music, making that way the advantages available to all its users. The Internet edition has been recently developed in order to offer solutions in recent trends of music broadcasting over IP. These solutions include algorithms for audio indexing and retrieval, classification, audio segmentation, as well as semantic metrics, and visualizations for academics, researchers, and musicologists.

Many uses of MEL-IRIS, as well as its basic structure, have been demonstrated in this manuscript and there are still more to come. The development team in Computer Science Department makes continuously improvements on the system, aiming at a fully integrated MIR environment with a great variety of uses.

References

- [1] D. Politis and D. Margounakis, "Determining the chromatic index of music," in *Proceedings of the 3rd International Conference on Web Delivering of Music (WEDELMUSIC '03)*, pp. 95–102, Leeds, UK, September 2003.
- [2] D. Politis, D. Margounakis, and K. Mokos, "Visualizing the chromatic index of music," in *Proceedings of the 4th International Conference on Web Delivering of Music (WEDELMUSIC '04)*, pp. 102–109, Barcelona, Spain, September 2004.
- [3] R. Shepard, "Pitch perception and measurement," in *Music, Cognition and Computerized Sound*, P. Cook, Ed., MIT Press, Cambridge, Mass, USA, 1999.
- [4] W. L. Idson and D. W. Massaro, "A bidimensional model of pitch in the recognition of melodies," *Perception & Psychophysics*, vol. 24, no. 6, pp. 551–565, 1978.
- [5] V. Barski, *Chromaticism*, Harwood, Amsterdam, The Netherlands, 1996.
- [6] A Jacobs, *The New Penguin Dictionary of Music*, Penguin, New York, NY, USA, 1980.
- [7] M. L. West, *Ancient Greek Music*, Oxford University Press, Oxford, UK, 1994.
- [8] A. Wood and I. M. Bowsher, *The Physics of Music*, John Wiley & Sons, New York, NY, USA, 1975.
- [9] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.
- [10] C. Padgham, "The scaling of the timbre of the piping organ," *Acustica*, vol. 60, pp. 189–204, 1986.
- [11] J. L. Caivano, "Colour and sound: physical and psychophysical relations," *Colour Research and Applications*, vol. 19, no. 2, pp. 126–132, 1994.
- [12] K. Giannakis and M. Smith, "Auditory-visual associations for music compositional processes: a survey," in *Proceedings of the International Computer Music Conference (ICMC '00)*, Berlin, Germany, 2000.
- [13] C. L. Krumhansl, "Music: a link between cognition and emotion," *Current Directions in Psychological Science*, vol. 11, no. 2, pp. 45–50, 2002.
- [14] R. Jackendoff and F. Lerdahl, "The capacity for music: what is it, and what's special about it?" *Cognition*, vol. 100, no. 1, pp. 33–72, 2006.
- [15] E. Schubert, "Modeling perceived emotion with continuous musical features," *Music Perception*, vol. 21, no. 4, pp. 561–585, 2004.
- [16] E. D. Scheirer, "Tempo and beat analysis of acoustic musical signals," *Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 588–601, 1998.
- [17] F. Kurth and M. Muler, "Efficient index-based audio matching," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 16, no. 2, pp. 382–395, 2008.
- [18] M. Clausen and F. Kurth, "A unified approach to content-based and fault-tolerant music recognition," *IEEE Transactions on Multimedia*, vol. 6, no. 5, pp. 717–731, 2004.
- [19] J. P. Bello, "Audio-based cover song retrieval using approximate chord sequences: testing shifts, gaps, swaps and beats," in *Proceedings of the International Symposium on Music Information Retrieval (ISMIR '07)*, pp. 239–244, September 2007.
- [20] S. B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 4, pp. 357–366, 1980.
- [21] G. Tzanetakis and P. Cook, "Multifeature audio segmentation for browsing and annotation," in *Proceedings of the IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics (WASPLAA '99)*, pp. 103–106, New Paltz, NY, USA, October 1999.
- [22] D. Kimber and L. Wilcox, "Acoustic segmentation for audio browsers," in *Proceedings of the Interface Conference*, L. Billard and N. I. Fisher, Eds., pp. 295–304, Sydney, Australia, 1996.
- [23] M. A. Bartsch and G. H. Wakefield, "Audio thumbnailing of popular music using chroma-based representations," *IEEE Transactions on Multimedia*, vol. 7, no. 1, pp. 96–104, 2005.
- [24] M. Marolt, "A mid-level representation for melody-based retrieval in audio collections," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1617–1625, 2008.
- [25] A. Klapuri, "Multiple fundamental frequency estimation by summing harmonic amplitudes," in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR '06)*, pp. 216–221, Victoria, Canada, October 2006.
- [26] S. M. Smith and G. N. Williams, "A visualization of music," in *Proceedings of the IEEE Visualization*, pp. 499–503, Phoenix, Ariz, USA, October 1997.
- [27] F. Vignoli, R. van Gulik, and H. van de Wetering, "Mapping music in the palm of your hand, explore and discover your collection," in *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR '04)*, E. Fox and N. Rowe, Eds., pp. 409–414, Barcelona, Spain, October 2004.
- [28] E. Pampalk and M. Goto, "Musicrainbow: a new user interface to discover artists using audio-based similarity and web-based labeling," in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR '06)*, pp. 367–370, Victoria, Canada, October 2006.
- [29] K. Attakitmongcol, R. Chinvejitvanich, and S. Sujitjorn, "Characterization of traditional thai musical scale," in *Proceedings of the 5th WSEAS International Conference on Acoustics and Music: Theory & Applications (AMTA '04)*, Venice, Italy, November 2004.

- [30] E. Cambouropoulos and G. Widmer, "Automatic motivic analysis via melodic clustering," *Journal of New Music Research*, vol. 29, no. 4, pp. 303–317, 2000.
- [31] F. Lerdahl and R. Jackendoff, *A Generative Theory of Tonal Music*, MIT Press, Cambridge, Mass, USA, 1983.
- [32] N. Fakotakis, E. Stamatatos, and G. Kokkinakis, "Automatic text categorization in terms of genre and author," *Computational Linguistics*, vol. 26, no. 4, pp. 471–495, 2000.
- [33] A. Ghias, J. Logan, D. Chamberlin, and B. C. Smith, "Query by humming: musical information retrieval in an audio database," in *Proceedings of the 3rd ACM International Conference on Multimedia*, pp. 231–236, San Francisco, Calif, USA, November 1995.
- [34] J. Wang, W.-J. Yang, and R. Acharya, "Color space quantization for color-content-based query systems," *Multimedia Tools and Applications*, vol. 13, no. 1, pp. 73–91, 2001.
- [35] D. P. W. Ellis and G. E. Poliner, "Identifying 'cover songs' with chroma features and dynamic programming beat tracking," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 4, pp. 1429–1432, Honolulu, Hawaii, USA, April 2007.
- [36] J. H. Lee and J. S. Downie, "Survey of music information needs, uses, and seeking behaviours: preliminary findings," in *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR '04)*, pp. 441–446, Barcelona, Spain, 2004.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

