Detecting Diagonal Activity to Quantify Harmonic Structure Preservation with Cochlear Implant Mapping

Sherif A. Omran sherif.omran@gmx.de
Mathematics and Natural science / Institute of Neuroinformatics / University of Zurich, 8057
Experimental Audiology Department / University Hospital of Zurich, 8091
Zurich, Switzerland

Abstract

Matrix multiplication is widely utilized in signal and image processing. In numerous cases, it may be considered faster than conventional algorithms. Images and sounds may be presented in a multi-dimensional matrix form. An application under study is detecting diagonal activities in matrices to quantify the amount of harmonic structure preservation of musical tones using different algorithms that may be employed in cochlear implant devices. In this paper, a new matrix called "Omran matrix" has been proposed. When it is post multiplied with another matrix, the first row of the output represents indices of fully active detected diagonals in its upper triangle. A preprocessing matrix manipulation might be mandatory. The results show that the proposed matrix is powerful in this application and illustrated higher grade of harmonic structure preservation of one algorithm used for cochlear implants with respect to other algorithms.

Keywords: Robotics, Cochlear Implant, Diagonal Detecting, Lines Detecting, Matrix, Harmonic Structure, Music.

1. INTRODUCTION

Cochlear implants (CIs) are devices aimed at restoring speech perception. The Nucleus cochlear implant uses a standard frequency to channel mapping algorithm that distorts harmonic structure of overtones in musical tones especially at low frequencies. Two new mapping algorithms based on semitone (Smt) mapping were developed and tested [1, 2]. These algorithms in addition to the standard mapping algorithm were used to process synthetic musical sounds and resynthesize them using a noise band vocoder acoustic model [3]. This paper aims at quantifying the degree these algorithms can reach in preserving the harmonic structure of overtones of musical tones. Sound signals are describable in three dimensions (intensity, frequency and time) and music is a complex sound that consists of tones. In general, tones have harmonically related partials [4]. Detecting active diagonals in matrices may be considered one way to detect harmonic structures in successive tones. The index of a detected diagonal in a frequency time matrix can represent the ratio between different overtones in such a case. Pattern recognition of objects inside matrices may be time consuming, but using the proposed Omran matrix (O'mat), matrix manipulation may be an easier approach. This paper is organized as follows:- Section 2 introduces a hypothesis and section 3 describes O'mat in details in addition to some examples in different special cases. Later, the algorithm is summarized in simple steps followed by a study using special acoustic signals, representing synthetic musical notes processed with the different algorithms; standard and semitone (Smt-MF and Smt-LF) mappings [1]. The output is preprocessed and is multiplied with O'mat. Section 4 is a proposed application for quantifying harmonic structure preservation using different music processing algorithms for CIs. Section 5 presents the results, where the following sections are discussion and conclusion.

2. HYPOTHESES

Preserving harmonic structure of music while being processed with a non-linear device, algorithm or a telecommunication system is expected to ameliorate music naturalness. Smt mapping which is one of the algorithms under study is expected to increase preservation of harmonic structure of overtones in CI devices because of its underlying nature.

3. METHOD

Definition: O'mat is a vertical rectangular or square matrix O[i,j] described by Equation 1. A matrix A is singular if the system Ax = 0 has non trifle solutions [5]. The vertical rectangle form of O'mat is invertible unless using the Moore-Penrose pseudo inverse. The first square B of the Moore-Penrose pseudo inverse of O'mat can be singular [6], where B matrix has dimensions equals to the rank of O'mat.

Matrix O[i, j] = 0, $i=m, j=n$ for $i \neq j$ and $i \geq j$ Equation 1
= m-n+1, $i=m-n+1, j=c$ where $1 \leq c \leq m-1$ and $i \geq j$

Each diagonal in the matrix is given a representative incrementing index number (see Figure 1).The first diagonal starts with index =1 and increments up to the main diagonal of the matrix. For appropriate recognition, each diagonal has to be fully active otherwise shifting the matrix one row upward and one column left or a matrix manipulation may be mandatory in a preprocessing phase. Lower triangles cannot be directly detected unless with matrix shifting or transposition.

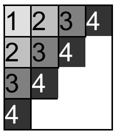


FIGURE 1: Diagonal indices of a matrix are shown in the same color. Each diagonal in the upper triangle of the matrix has a unique index number.

O'mat is a newly proposed matrix with zeros every where except the main diagonal of the lowest square as defined by Equation 1 and illustrated in Figure 2. All the values along the main diagonal increment by one and have their maximum which equals to the number of rows (i) at the lowest right corner. Figure 2 shows an example for a 4x4 (square) and 4x3 (vertical rectangle) O'mat. It illustrates two O'mats with 4 rows (i=4) (square and rectangular shapes).

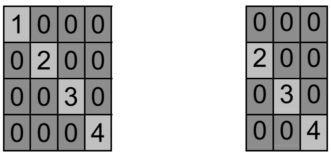


FIGURE 2: An example of a square (left) and a vertical rectangular (right) shape of O'mat with i=4. O'mat can not be horizontal by definition.

3.1 Proof

Assume an image B (4x4) with 2 diagonals active and the corresponding O'mat (4x4) matrix O

B=

0	a	b	0
c	d	0	0
e	0	0	0
0	0	0	0

O=

1	0	0	0
0	2	0	0
0	0	3	0
0	0	0	4

Result = B x O

=

0	2a	3b	0
1c	2d	0	0
1e	0	0	0
0	0	0	0

R(i,j) = B(i,j).O(j,j)

Assuming that the B image is binary (black=0 and white=1) and B(n,m) = 1 are elements in an active diagonal.

R(n,m)= 1. O(m,m) Equation 2

Equation 2 shows that the elements in the resulting matrix (R) that represent the active diagonals will not be dependent on the image, however the output will represent O'mat values in a decrementing arrangement. The following section has numerous examples in many special cases to illustrate the idea.

To understand the concept, let's start with a simple example for a non-singular O'mat and a binary image matrix image with four pixels. The lowest right element at i = m and j = n should be divided by the number of rows (i) in O'mat prior to being multiplied with O'mat. This is illustrated in the following examples:-

3.2 Example 1

Image x O'mat = Output

0	1	1	0
1	1	0	0
1	0	0	0
0	0	0	¼

x

1	0	0	0
0	2	0	0
0	0	3	0
0	0	0	4

=

0	2	3	0
1	2	0	0
1	0	0	0
0	0	0	0

EXAMPLE 1: To detect diagonal activities in an image using a non singular O'mat.

| Determinant (Output) | = 3! x 2! = 6

$$Weight = \prod_{1}^{n} output(1,n)! \text{ where } , output(1,n) \neq 0$$ Equation 3

= 3! x 2! = 6

The weight value is calculated by multiplying all non zero elements in the first row of the output matrix as described by equation (3). In this example, the determinant of the output equals to its weight.

3.3 Example 2

Image x O'mat = Output

0	1	1	0
1	1	0	0
1	0	0	¼

x

0	0	0
2	0	0
0	3	0
0	0	4

=

2	3	0
2	0	0
0	0	1

EXAMPLE 2: To detect diagonal activities in an image using O'mat, whose first square of its pseudo inverse is singular while the output is a non-singular square matrix.

$$Weight = \prod_{1}^{n} output(1,n)! \text{ where } , output(1,n) \neq 0$$

= 3! x 2! = 6

Both examples illustrated that the weight value is a factorial multiplication of the diagonal indices in the images that were fully active (see Figure 1).

3.4 Example 3

The following examples illustrate the case when the image is singular.

Image x O'mat = Output

1	1	0	0
1	0	0	0
0	0	0	0
0	0	0	¼

x

1	0	0	0
0	2	0	0
0	0	3	0
0	0	0	4

=

1	2	0	0
1	0	0	0
0	0	0	0
0	0	0	1

EXAMPLE 3: To detect diagonal activities in an image represented by a singular square matrix. One drawback of O'mat is inability to detect activity in diagonal with index =1.

| Determinant (Image) | = 0

$$Weight = \prod_{1}^{n} output(1,n)! \text{ where } , output(1,n) \neq 0$$

= 2! x 1! = 2

3.5 Example 4

Image x O'mat = Output

0	1	1	0
1	1	0	0
1	0	0	¼

x

0	0	0
2	0	0
0	3	0
0	0	4

=

2	3	0
2	0	0
0	0	1

EXAMPLE 4: To detect diagonal activities in an image represented by a rectangular matrix having diagonal with index numbers 2 and 3 fully active using a rectangular O'mat.

$$Weight = \prod_{1}^{n} output(1,n)! \text{ where } , output(1,n) \neq 0$$

= 2! x 3! = 6

The weight value may be very complex especially in large matrices; it would practically exceed double precision length. To simplify the formula, an estimated weight function is proposed as in Equation 4:

$$\text{Estimated Weight} = \sum_{1}^{n} 2^{output(1,n)} \quad \text{Equation 4}$$

Steps to detect a diagonal activity:-

1 - Construct a zeros vertical rectangular or square matrix $_{(i,j)}$, where i≥j
2 - Fill the diagonal of the lowest square with decrementing numbers starting at element (i=n, j=n)
3 - Change the last (i,j) element in the image matrix into unity divided by n
4 - Post multiply O'mat with an image matrix that has a diagonal activity.
5- Check output diagonal values being decreasing by 1.
6 - Multiply the factorial of all non-zero elements in the first row of the output by each other to get the weight.
7- The indexes of the active diagonals exist in the first row of the output matrix.
8- Check the activity of the first elements (1,j) of the input image manually.

The matrix could be built with a simple Matlab script as shown in Figure 3.

```
function x=omran_matrix(i,j)
% Input : Required matrix dimension (i,j)
% Output: Omran Matrix
if nargin < 2,
    x=[];
    disp('please give matrix dimensions');
    return
end
if i<j,
    x=[];
    disp('Error: dimensions are not correct. Matrix can be vertical or square i>=j');
    return
end

x=zeros(i,j);
counter=1;
for q=i-j+1:i
    x(i,counter)=q;
    counter=counter+1;
end

end
```

FIGURE 3: A function to construct O'mat written in Matlab from Mathworks.

4. APPLICATION

Music is a complex sound consisting of tones which in turn, consists of fundamentals and overtones that generally have harmonic structure [7]. The majority of musical instruments have fundamental frequencies below 1 kHz [4]. Music is a sound that can be organized mainly in three elements (melody, rhythm and harmony) [8]. Melody is one of the important aspects of music [9], and is described as a sequence of individual tones that are perceived as a single entity [10]. Preserving the harmonic structure of individual tones is assumed to be necessary for preserving the melody perception.

One way to improve melody representation in CIs would be to ensure that the fundamental frequencies of adjacent tones on the musical scale are assigned to separate electrodes [1]. Such an approach involves mapping fundamental frequencies of musical tones to electrodes based on a semitone scale. The idea was initially investigated in a study by Kasturi and Loizou [11], using the 12 electrode Clarion CII (Advance Bionics) implant with a limited range of semitone frequencies. They concluded that semitone spacing improved melody recognition with CI recipients. Additionally, music could be enhanced by increasing the frequency representation. This may be possible using virtual channels (VCs) formed by stimulating two adjacent electrodes simultaneously with the same current level. Busby and Plant [12] reported that VCs invoked the perception of an intermediate pitch. VCs on an array of 22 electrodes would yield a total number of 43 channels, which would cover three and a half octaves with Smt mapping with one-semitone intervals between the characteristic frequencies of successive channels. A comparison between Std and Smt mapping with normal hearing and CI recipients showed a perceptual improvement in melody recognition [2]. In this paper, the amount of harmonic structure preservation is calculated quantitatively and relatively utilizing a special constructed sound signal that was processed with Std, Smt-MF and Smt-LF mappings and then re-synthesized with a noise band vocoder acoustic model for CI devices [3].

A sequence of 36 complex harmonic tones is prepared; each tone has the fundamental frequencies of piano note for a period of 100 msec with 4 overtones. The semitone mapping is claimed to preserve the harmonic structure of overtones as shown in Figure 4 and Figure 5 (22 and 43 channels respectively) where the ratio between the overtones with Smt mapping is almost constant; unlike with the Std mapping where there is a distortion of the harmonic structure at low frequencies. A Matlab program was developed for preprocessing before applying the O'mat algorithm. Harmonic structure preservation is demonstrated by a linear frequency to channel relationship, as can be seen for Smt-MF and SMT-LF for each of the partials. With the Std mapping, the linearity is seen only for the higher frequencies. At lower frequencies, the partials can not even be resolved. With Smt-MF components below 440 Hz are filtered out and with Smt-LF the high frequency partials greater than 1.51 kHz are filtered out (as indicated by arrows) [1].

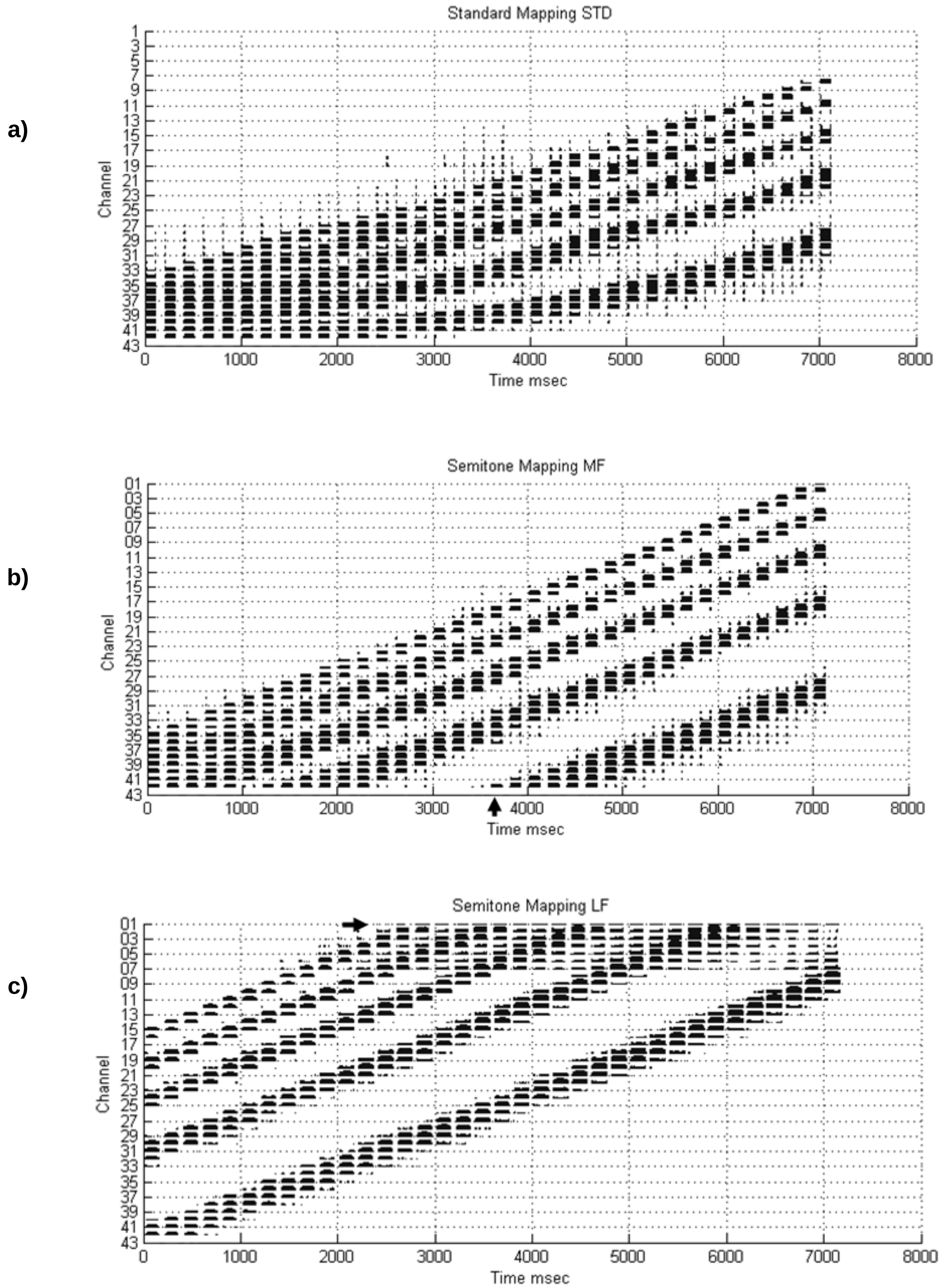
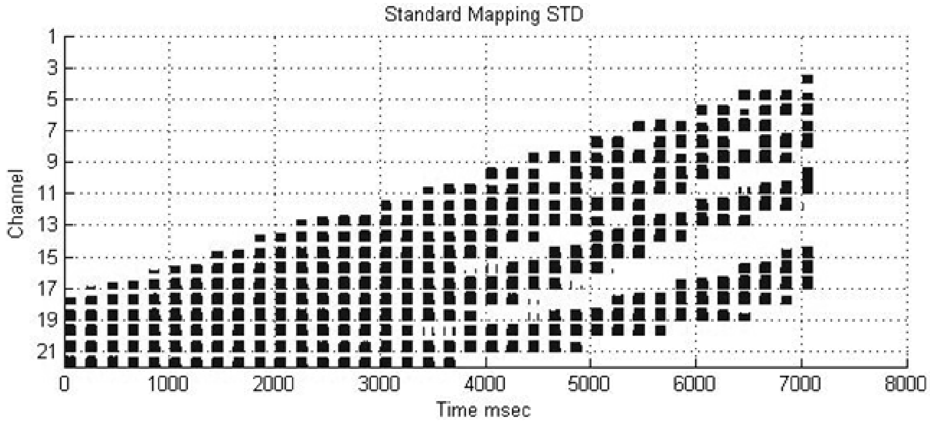
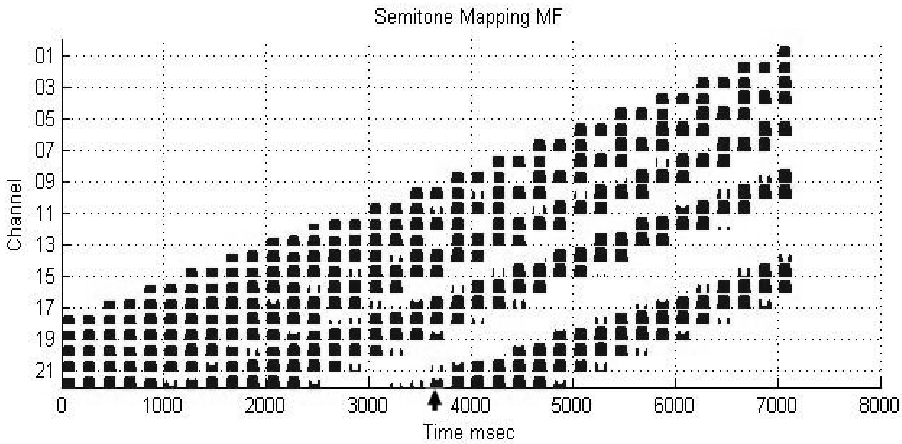


FIGURE 4: CTM outputs from the AMO for the Std (left), Smt-MF (middle) and Smt-LF (right) mapping using 43 channels.

a)



b)



c)

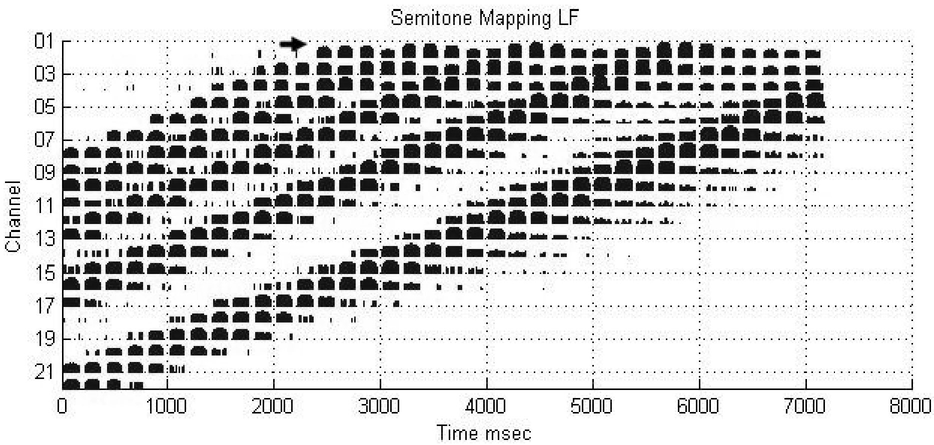


FIGURE 5: CTM outputs from the AMO for the Std (left), Smt-MF (middle) and Smt-LF (right) mapping using 22 channels.

5. RESULTS

Signals were preprocessed with an acoustic model [3] representing different algorithms for CIs using Std and Smt mappings with 43 and 22 channels [1, 2], The results are demonstrated in Figure 6 as a relative ratio of harmonic structure preservation with respect to Std mapping.

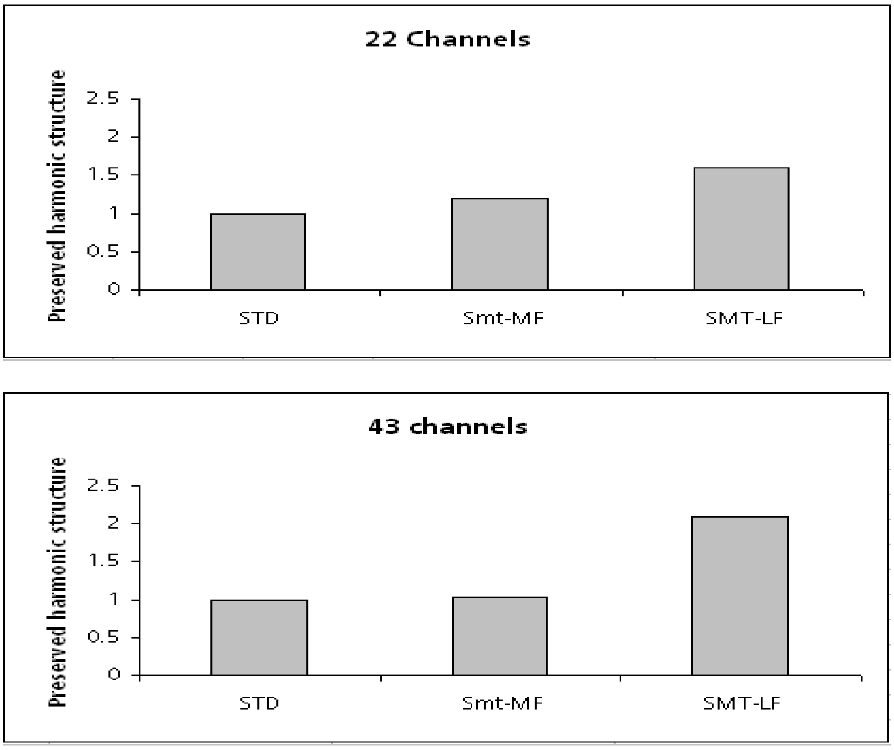


FIGURE 6: Preserved amount of harmonic structure with respect to standard mapping for 22 channels (up) and 43 channels (down).

Figure 6 shows that using 22 and 43 channels the preservation of harmonic structures increase with Smt-MF and increases more with Smt-LF than the Std mapping. In 43 channels the increase between Smt-MF and Std mapping is relatively smaller than of 22 channels, while Smt-LF with 43 channels provides higher preservation of harmonic structure than with 22 channels. Although Smt-LF shows the highest preservation, but it transposes sounds to higher frequencies [1] and patients described it as higher in pitch than similar sounds with Smt-MF [2].

6. DISCUSSION

Music has many descriptions; Heinrich Hueschen (1961) described it among the artistic disciplines and said “it is the one whose material consists of tones. Of the raw materials available in nature, only a small proportion is actually used in music. The finite number of tones selected for musical use from the infinity available in nature is organized into specific tone systems through defined rational processes” [13]. Hans Heinrich Eggebrecht (1967) provided a characterizing definition and said “music is – in the area in which the concept is relevant, western culture – the artistic formation of those sounds that represent the world and the spirit in the form of a voice of nature and emotion in the realm of hearing, concretely conceived, and which achieves significance as an art, become both meaningful and meaning-creating material through reflected and ordered cognition and theory. For the basic element of music, the tone is on one hand the bearer of meaning, while on the other hand it is the vehicle of meaning as the beneficiary of the tonal order. These lend to the unit of music, tone, its specifically cultural forms, meanings and conceptions and at the same time, as a natural phenomenon, it remains accountable to the laws of nature” [14]. Till now there is neither a complete definition of music nor a unit for its underlying aspects. Rather, various descriptions of the concept that appear in the literature and emphasize a particular aspect of the total phenomenon [9, 13].

Music consists of tones; where a tone is a sound played or sung with a particular pitch and duration. It is one of the important aspects of music. Melody is an agreeable succession of tones that represents a link between the individual musical tone and tonal music as a whole [10]. The sequence of tones in a melody forms its melodic line. Within each melodic line a sequence of oriented intervals defined by neighboring tones can be traced. This sequence demonstrates the movement of the melodic line [15]. Melodies are often described as being made up of phrases. They may also be perceived in short musical ideas (melodic, harmonic, rhythmic or any combination of these three) called motifs [9]. A motif is the smallest meaningful unit, a perceivable or salient recurring fragment or succession of notes that may be used to construct parts of a melody [16]. Long melodies that reappear in a music piece are called themes.

Another two important aspects of music are rhythm and meter. A series of events (whether musical notes or blows of a hammer) is commonly characterized as ‘rhythmic’ if some or all of those events occur at regular time intervals. Rhythm is a major aspect of music, dance, and most poetry; it involves the pattern of both regular and irregular durations that are present in music. Meter involves the perception and anticipation of rhythm patterns [17]. Meter is a structured attending to time which allows listeners to have precise expectations as to when subsequent musical events are going to occur. It plays a crucial role in the determination of relative durations. Not all durations are perceived the same, as there are a number of psycho-physical limits on the ability to perceive durations and durational succession. Divenyi (1974) demonstrated that two onsets must be separated by at least 2 ms in order to be distinctly perceived, and that at least 15-20 ms are required to determine which onset came first and 100 ms seems to be the threshold for reliable judgments of the length [18]. In rhythm, any periodic transient signal that marks it; is called a beat. Beat may be described as a sudden change in energy with respect to the preceding history. Much music is characterized by the sequence of stressed and unstressed beats [7]. The perception of a beat is not only necessary for a sense of ‘connectedness’ among successive events; it may also be necessary for a sense of motion [19]. A sense of beat, while necessary, is not sufficient to engender a sense of meter. Timbre is a characteristic feature of musical instruments. Helmholtz was the first person to describe the timbre as a property of the spectral components of the sound [20]. Helmholtz view of timbre was that the perceptual cues came from the Fourier series coefficients. Unlike pitch or rhythm, timbre is developed from many physical features and can not be measured with a scale [7]. Timbre is defined physically by the temporal envelope (particularly the onset) and the spectral shape of the acoustic sound [21]. The characteristics of timbre do not come particularly from the ratio of the strengths of partials, but from the formants.

Harmony is another important aspect of music. The term harmony is derived from the Greek word ‘harmonia’ and means ‘fit or join together’. The term ‘harmony’ refers to the combination of notes simultaneously to produce chords and successively to produce chord progressions [9]. Chords are groups of notes built on major or minor triads and sounded simultaneously. Sound keeps pleasant to the ear if a musical relationship exists between the overtones of a chord. From the engineering point of view, music can be described as a series of complex acoustic sounds composed of tones. Such tones have a harmonic structure of overtones [4]. Also, harmony can be produced from inharmonic spectrum [7]. An inharmonic spectrum is formed from aperiodic tones; such as frequencies with the “Goldener Schnitt” ratio 1.61803 [22] where the signals in the time domain do not overlap except at t=0.

Acoustic sound in general is described in three domains (loudness, frequency and time). Music has many important aspects and some of them are summarized in Figure 7.

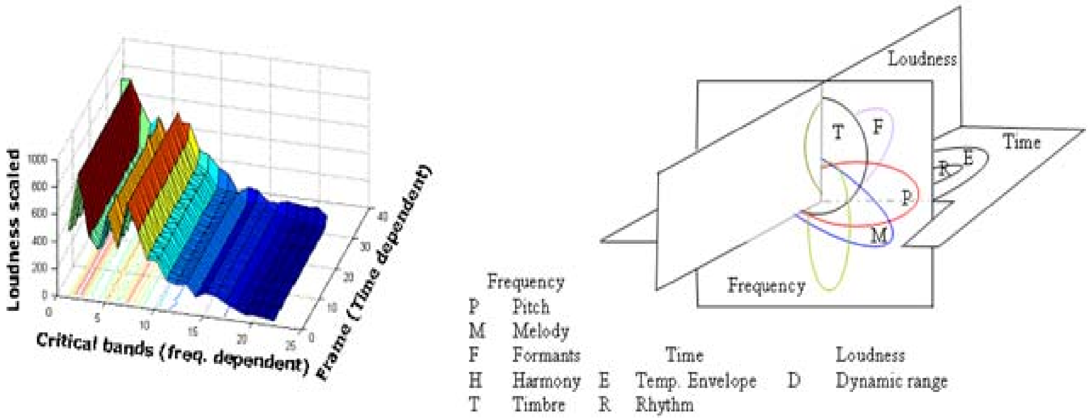


FIGURE 7: Represents a spectro-contrast gram (loudness in Sones scaled along z-axis, critical bands on the x-axis and different frames along y-axis) of a complex clarinet tone in nature. To the right side, different aspects of music are presented categorized according to the three dimensions of sound (frequency, time and loudness).

Up to now, there is no analytical measure of many musical aspects such as the harmony or how appreciated could musical sounds be compared to other processing algorithms. But psychoacoustic tests carried out with either normal hearing subjects through a simulator or CI recipients had proved an enhancement with Smt-MF mapping [2]. However, the proposed method adds a new analytical approach to address this problem. Simulations in Figures (4-5) show clearly that the proposed Semitone mapping which is based on a logarithm semitone scale can enhance representation of musical partials with respect to the current Std mapping algorithm.

On another hand, the proposed method could be used as well in other applications such as robots cameras, mainly in detecting road lines using an index maximization simple matrix rotation and multiplication operations of segregated image parts, which is expected to be effective than currently used methods like Sobel edge detection.

7. CONCLUSION

The O'mat proposed in this paper, proved it can be used to detect diagonal activities in matrices or images, with a detection limitation of the first element (index=1). However, this can be simply compensated with one line of a conditional code. The O'mat algorithm was applied to relatively quantify the amount of harmonic structure preservation with Smt-MF and Smt-LF mappings with respect to the Std mapping for CI devices. Results show that 43 channels provided higher quantity of preservation than 22 channels with Smt-LF mapping. It shows also that the Smt-LF highly ameliorates harmonic structure preservation of overtones in musical notes with respect to Smt-MF mapping which is higher as well than Std mapping.

8. ACKNOWLEDGMENT

This work was partially supported with a scholarship from the University of Zurich. Many thanks to my professors Norbert Dillier and Rodney Douglas.

9. REFERENCES

1. S. Omran, W.K. Lai, N. Dillier. *"Semitone Frequency Maps to Improve Music representation for Nucleus Cochlear Implants"*. Audio speech and music processing journal, submitted, 2011

2. S. Omran, W.K. Lai, N. Dillier. *"Pitch ranking, Melody contour and instrument recognition tests using two semitone frequency maps for Nucleus Cochlear Implants"*. Audio speech and music processing journal, 2010

3. J. Laneau, M. Moonen, J. Wouters. *“Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants”*. J Acoust Soc Am, 119(1):491-506, 2006

4. J. Pierce. *“The science of musical sound”*, Scientific American Books, New York (1983)

5. F. Gantmacher. *“The Theory of Matrices”*, American Mathematical Society, Rhode Island (1990)

6. A. Neubauer. *“Irreguläre Abtastung: Signaltheorie und Signalverarbeitung (Irregular sampling: signal theory and signal processing)”*, Springer Inc, Berlin (2003)

7. M. Neukom. *“Signale, Systeme und Klangsynthesse: Grundlagen der Computermusik (Singals, Systems and Sound synthesis: basics of computer music”*, Peter Lang Inc, Bern (2005)

8. K. Wyatt, C. Schroeder. *“Harmony and Theory: a comprehensive source for all musicians”*, Musicians Institute Press (1998)

9. S. Sadie, G. Grove. *“The New Grove Dictionary of Music and Musicians”*, Grove, London (1995)

10. E. Terhardt. *“Akustische Kommunikation (acoustic communication)”*, Springer-Verlag, Berlin (1998)

11. K. Kasturi, P. Loizou. *“Effect of filter spacing on melody recognition: acoustic and electric hearing”*. J Acoust Soc Am, 122(2):29-34, 2007

12. P.A. Busby, K.L. Plant. *“Dual electrode stimulation using the nucleus CI24RE cochlear implant: electrode impedance and pitch ranking studies”*. Ear Hear, 26(5): 504-515, 2005
13. F. Blume. *“Die Musik in Geschichte und Gegenwart (Music in history and present)”*, Finscher Ludwig (1961)

14. B. Riemann. *“Musik Lexikon (Music lexicon)”*, Direct media Publishing GmbH, Berlin (2000)

15. J. Hofman-Jablan. *“Antisymmetry and colored symmetry of musical work”*. Culture & Science, 6(2): 249-251,1995

16. R. Scruton. *“The Aesthetics of Music”*, Clarendon Press, Oxford (1997)

17. G. Houle. *“Meter in Music, Performance, Perception, and Notation”*, Indiana University Press, pp. 1600-1800 (1987)

18. P.L. Divenyi, I.J. Hirsh. *“Identification of temporal order in three-tone sequences”*. J Acoust Soc Am, 56(1): 144-191, 1974

19. A.S. Bregman. *“Auditory Scene Analysis: The Perceptual Organization of sound”*, The MIT Press, Cambridge, Massachusetts (1990)

20. H.L.F. Helmholtz. *“On the Sensations of Tone as a Physiological Basis for the Theory of Music”*, New York Dover (1954)

21. S. Handel. *“Timbre perception and auditory object formation”*, Academic Press, San Diego (CA). pp. 425-461 (1995)

22. H. Walser. *“Der Goldene Schnitt (The Golden Schnitt)”*. Teubner, Stuttgart (1993)