Neural perceptron-based Training and Classification of Acoustic Signal

*Yoon-ho Kim

*Div. of Computer Eng. , Mokwon University

e-mail:yhkim@mokwon.ac.kr

ABSTRACT

The MPEG/audio standard results from three years of co-work by an international committe of high-fidelity audio compression experts in the Moving Picture Experts Group (MPEG/audio). The MPEG standard is rigid only where necessary to ensure interoperability. In this paper, a new approach of training and classification of acoustic signal is addressed. This is some what a fields of application aspects rather than technonical problems such as MPEG/codec, MIDI. In preprocessing, acoustic signal is transformmed using DWT so as to extract a feature parameters of sound such as loudness, pitch, bandwidth and harmonicity. these accoustic parameters are exploited to the input vector of neural perceptron. Experimental results showed that proposed approach can be used for tunning the dissonance chord.

1. Introduction

The physical properties of the acoustic environment are encoded in the sound field and nust be decoded by an immersive audio system in order to accurately simulate the original environment. The influence of the local acoustic environment is reflected in the perception of spatial attributes such as direction and distance, as well as that of room spaciousness and source size. Furthermore, such acoustic features as loudness, pitch, and bandwidth are directly related to the musician or composer whether who is majored or not in acoustic signal.

In the past, musics were composed by some invisible rules, but, now a days, the popular music composition rules are almost unrestrained. The development of computer music helps non-musicians to compose musics. An important stage in evolution of structurred representation for musician as well as composer has been the emergence of MIDI(Musical Instrument Digital Interface) as an industry standard for musical score exchange. Although MIDI is a kind of protocol, it was originally conceived as a metod for communication between musical instrument and computers which control them. By using MIDI, musician can design musical copositions using powerful, user-friendly tools that interact with their home equipment.

That is why it does not make us feel constrained if we listen to a music which is composed by a non-musician. In other words, non-musicians practice some counter points and harmonics everyday without any trainings. Even though non-musicians, they do not make some unnatural chord progresses and bizarre melodies because they know it is strange.

In general, almost of us think that musician can only compose a music, but non-musician can compose a music and melodies by using the auto chord composition program and perform the music with the band-in-a-box which is a auto performance program [3].

2. Generation of harmonic chord for music composition

Music composition divides into two categories: the melody and the chord composition [2]. The first mean is to compose a music with various chords, and the melody composition means compose by using h notations. In here, the chord means the accord, for instance, C

chord is the complex sound of do, mi and sol.G chord is the complex sound of sol, si and re.The chord progress rule is not defined but we probably know about what is crude chord progress and what is harmonic chord progress. For instance, to use F chord after C chord is very harmonic chord progress, but to use Bb chord after C chord should not make harmonic chord progress. Because of having differet opinion of harmonic chord individually, in this paper, we should choose lots of musics which is chosen the top 10 music during 5 year, and then statistical analyzed and arranged the harmonic chord progress. Table. 1 listed in harmonic chord pattern with 100 musics [3], and Fig. 1 showed the quad-tree representation for assigning the teaching vector.

Table 1. Harmonic chord pattern

Start Chord	harmonic chord progress				Teaching vector
C	G	F	Em	Am	0000
D	G	Em	Gm	C	0001
E	Am	D	A	F	0010
F	Fm	G	C	Bm	0011
G	C	Dm	Em	F	0100
A	Dm	F	E	D	0101
B	Gm	Em	Cm	G	0110
Cm	Fm	G	Gm	D	0111
Dm	F	Am	G	C	1000
Em	Am	A	C	F	1001
Fm	C	G	Em	Gm	1010
Gm	Am	Em	C	D	1011
Am	Dm	Em	D	F	1100
Bm	Gm	E	Cm	G	1101

Table. 1 illustrates a four possible chord pattern (chords are arranged by high probability to low probability) after each chords. In the popular music, the pattern is not always the same, but generally, the one of most widely used chord progress is consist of three parts: Prelude, interlude and postlude.

3. Extraction of acoustic features using DWT

Continuous wavelet transform (CWT) of a signal $\chi(t)$ can be expressed as

$$W_\psi x(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-b}{a}\right) dt. \quad (1)$$

Let us denote scale a = 1/2s and the

translation b = k/2s where s and k belong to the integer set Z. The CWT of $\chi(t)$ is a number at (k/2s,1/2s) on the time-scale plane. It represents the correlation between $\chi(t)$ and Ψ (t) at that time-scale point. We call this the discrete wavelet transform (DWT). Namely, the multi-resolution analysis is a set of function spaces $(V_j)_{j \in Z}$ that satisfy the properties of $V_{j+1} \subset V_j$, $\forall j \in Z$, and $f(x) \in V_j \leftrightarrow f(2^{(k-j)}x) \in V_k$. Let $\phi(x)$ be a scaling function, $\phi_{jn} = 2^{-(j/2)}\phi(2^{-j}x-n)$; $\psi(x)$ be a wavelet and $\psi_{jn} = 2^{-(j/2)}\psi(2^{-j}x-n)$. The wavelet transform of $f(x)$ at resolution 2^j is given by

$$W_j(k) = \langle f(x), \psi_{jk}(x) \rangle \sum g(2k-n)S_{j-1}(n)$$

which is the high frequency component of $f(x)$ at scale 2^j. Consequently, discrete wavelet transform can be expressed as Eq. (2).

$$DWT_{m, n(f)} = a_0^{-m/2} \int dt\, f(t)\, \psi\, (a_0^{-m}t-nb_0)$$
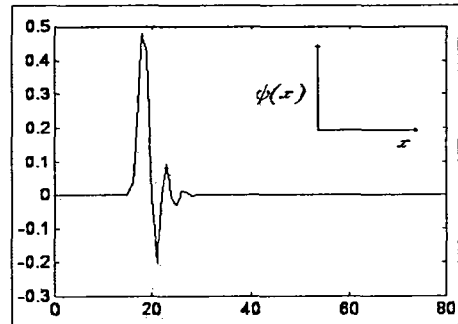
$$\int dt\, \psi(t) = 0 \quad (2)$$

Fig. 1. Example of mother wavelet in case of (a=1, b=16).

Example of mother wavelet is shown in Fig. 1. By useing Eq. (1), (2), DWT is performed so as to calculating the acoustic parameters such as loudness, pitch, bandwidth and harmonicity. loudness is approximated by RMS of signal. pitch is estimated by taking a series of short-time Fourier spectra. bandwidth is computed as the magnitude weighted average of the diference between the spectral components and centroid.

4. Neural model with perceptron

Every neuron model consists of a processing element with synaptic input connections and a single output. and then the neuron output signal can be modeld as follows:

$$o = f(w^t x), \text{ or}$$

$$o = f\left(\sum_{i=1}^{n} w_i x_i\right) \qquad (3)$$

where w is the weight vector defined as

$$w \triangleq [w_1 \ w_2 \cdots w_n]^t$$

and x, the input vector, is also expressed as :

$$x \triangleq [x_1 \ x_2 \cdots x_n]^t$$

Eq. (3) is often referred to as an activation function, and generally given by

$$f(net) \triangleq \frac{2}{1 + \exp(-\lambda net)} - 1 \qquad (4)$$

$$f(net) \triangleq sgn(net) = \begin{cases} +1, & net > 0 \\ -1, & net < 0 \end{cases} \qquad (5)$$

where $\lambda > 0$ in Eq. (4) is proportional to the neuron gain determining the steepness of the continuous function $f(net)$.

In general. the learning signal r is a function of w_i, x , and sometimes of the teacher's signal d_i. Thus learning signal r is defined as

$$r = r(w_i, x, d_i) \qquad (6)$$

The increment of the weight vector w_i produced by the learning step at time t according to the general learning rule is

$$\Delta w_i(t) = cr[w_i(t), x(t), d_i(t)] x(t) \qquad (7)$$

where c is a positive number called the learning constsnt that determines the rate of learning. The weight vector adapted at time t becomes at the next instant, or learning step,

$$w_i(t+1) = w_i(t) + cr[w_i(t), x(t), d_i(t)] x(t)$$

For the perceptron learning rule, the learning signal is the difference between the desired and actual neuron's response. Thus, learning is supervised and the learning signal is equal to

$$r \triangleq d_i - o_i$$

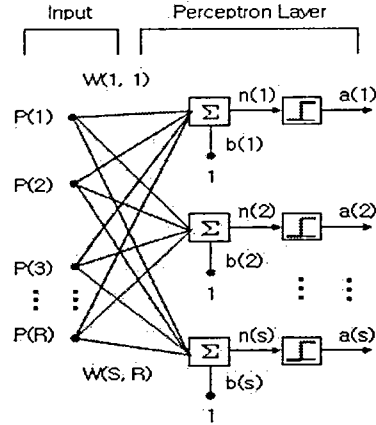where $o_i = sgn(w_i^t x)$, and d_i is the desired response as shown in Fig. (2)

Fig. 2. Perceptron learning rule

Weight adjustments in this scheme, Δw_i and Δw_{ij} are obtained as follows

$$\Delta w_i = c[d_i - sgn(w_i^t x)] x$$
$$\Delta w_{ij} = c[d_i - sgn(w_i^t x)] x_j, \qquad (8)$$
$$\text{for } j = 1, 2, \cdots, n$$

Under this rule, weights are adjusted if and only if o_i is incorrect. Error as a necessary condition of learning is inherently included in this training rule. Obviously, since the desired response is either 1 or -1, the weight Eq. (8) reduces to

$$\Delta w_i = \pm 2cx$$

where a plus sign is applicable when $d_i = 1$, and sgn $(w^t x) = -1$, and a minus sign is applicable when $d_i = -1$, and sgn $(w^t x) = 1$.

5. Experimental considerations

Basic procedure for extracting feature parameters after the performing DWT with a input image, has four stage, as follows:

step 1) : Perform the 1-level DWT and extract the coefficients.

step 2) : Analyze the acoustic parameters from the DWT coefficients.

step 3) : Arrange the harmonic chord pattern so as to use a perceptron neural input.

step 4) : Perform the learning function based on perceptron net.

step 5) : Calculate the recpgnition rate both in using perceptron and in acoustic parameters.

A neural network is a set of units which takes a linear combination of values from either an input vector or the output of other units. The linear combination then perform a nonlinear function, such as threshold or a sigmoid. Many kinds of neural network has been shown to offer potentially powerful, robust, and adaptive means of detecting and classifying objects under changing environmental conditions. One of the most desirable properties of these networks is to learn from examples and to generalize from training set to similar data, furthermore, they are inherently parallel and can solve problems faster than a serial system. We can see from table 2. that how the configuration of teaching vectors are constructed for each model. Recognition rates for each harmonic chord pattern are illustrated in Fig. (4).

Table 2. Teaching vectors of fig. 3 (a)

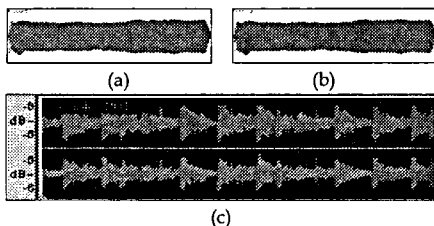| [Test model " Do "]	
0.5	-0.5
-0.5	-0.5

Fig. 3. Acoustic pattern: (a) Dormal pattern"Do", (b) # pattern "Do", (c) frequency pattern
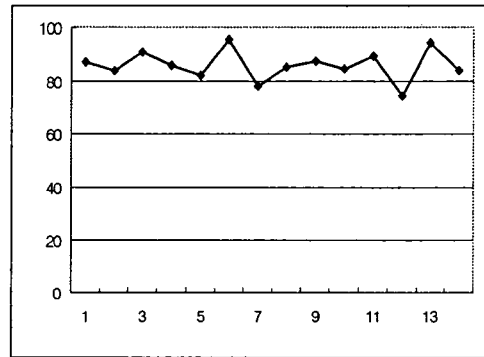

Fig. 4. Experimental result

6. Conclusions

We have introduced an algorithm for recognizing the acoustic pattern. Based on our experimental results and analysis, we draw the following general conclusions:

· Extraction of acoustic feature parameters are performed by using DWT.

.DWT based approach are relatively simple than other preprocessing which involved in space domain method.

· Experimental results of perceptron learning showed that proposed approach is able to use a acoustic correction system.

References

[1] Steven K. Rogers, Dennis W. Ruck Matthew Kabrisky Gregory L. Tarr, "Artificial Networks For Pattern Recognition", Intelligent Robots and Computer Vision VIII, pp. 410-417, 1989.

[2] Yun-Ho Kim, "Identity Decision of Dynamic Image by the Fuzzy Inference", Ph. D. thesis, The Gradute School, Chongju Univ., 1991.

[3] R. Schalkoff, "Pattern recognition statical ,structural, and neural aproaches", John Wiley & Son Inc., 1992.

[4] Myung-kil Lee and Joo-shin Lee, "A study on pattern recognition using DCT and neural network", KICS, Vol 22, No. 3, 1997.

[5] J. M. Zurada, Artificial Neural System, West publishing company, 1992.