

## Research on the Peripheral Sound Visualization Using the Improved Ripple Mode

Bin Xu<sup>1, a</sup>, Dan Yang<sup>2, b</sup>, Yunyi Zhang<sup>2</sup> and Xu Wang<sup>2, c</sup>

<sup>1</sup>Computing Center of Northeastern University, Shenyang City, Liaoning Province, China, 110819

<sup>2</sup>School of Information Science and Engineering, Northeastern University, Shenyang City, Liaoning Province, China, 110819

<sup>a</sup>xubin@mail.neu.edu.cn, <sup>b</sup>yangdan@ise.neu.edu.cn, <sup>c</sup>wangxu@ise.neu.edu.cn

**Keywords:** Peripheral sounds, Improved ripple mode, the deaf, ART1 neural network

**Abstract.** In this paper, we proposed a peripheral sound visualization method based on improved ripple mode for the deaf. In proposed mode, we designed the processes of transforming sound intensity and exterminating the locations of sound sources. We used power spectrum function to determine the sound intensity. ART1 neural network was subtly applied to identify which kind of the real-time input sound signals and to display the locations of the sound sources. We present the software that aids the development of peripheral displays and four sample peripheral displays are used to demonstrate our toolkit's capabilities. The results show that the proposed ripple mode correctly showed the information of combination of the sound intensity and location of the sound source and ART1 neural network made accurate identifications for input audio signals. Moreover, we found that participants in the research were more likely to achieve more information of locations of sound sources.

### Introduction

There are many subtle ways in which people use sound to gain awareness of the state of the world around them. For example, in an office, the sound of officemates working provides awareness of whether you are alone in the room or if you have company. Similarly, many everyday devices use sound to communicate events. According to their use, peripheral sound can be divided into two types. One is background sound, the other is warning tone. Background sound tells people what is happening in the ambient environment such as whispers and typing sound. Announcement tone indicates the occurrence of an event, which can draw people's attention and interrupt their current work to take action for it such as phone ringing and fire alarms. People who are deaf have difficulty maintaining an awareness of these peripheral sounds. They employ a variety of alternative mechanisms to gain this information through other channels[1,2,3]. However, these techniques and devices have limitations to some extent. Hearing dogs don't often provide the information of the background sound in the particular environment and flashing lights never provide environmental information. Flashing lights has very high investments when the system has been established. Each device such as telephones and doorbells should be connected separately and able to be seen in every room. Hearing dogs need high-qualified maintenance throughout. Besides, many of these techniques require priori knowledge of the sound monitoring while they're not competent when some sudden sound appears.

Recently, many researchers spent hard efforts on positional ripples visualization of peripheral sound first proposed by Jennifer Mankoff [4, 5, 6]. In the ripple mode, the ambient sound was simplified to a plane graph, in which information of sound intensity was displayed by a series of centric circles with different radii and colors. The greater the sound intensity was, the larger the centric circle was. Each center of a circle represented a location of sound source. However, they only focused on the display form the deaf required and pointed out that microphone array could be used to identify locations of the sound sources based on the hardware platform. In this paper, we proposed the specific software method for identification of locations of sound sources using sound intensity and ART1 neural network. Demos we developed are presented to test the capabilities of the work.

### Improved Ripple Mode

Our proposed method for the displayed peripheral sounds can be described as the following:

**Audio signals power spectrum shape features extraction.** Power spectrum shapes vary by various peripheral sounds and each kind of sounds has similar power spectrum shapes. In the proposed method, the power spectrum values are extracted according to the listed steps:

- 1) Monitoring the input audio signals to find out the effective starting and end points,
- 2) Recording 15 values of the power spectrum from the starting point without noise to make the length of the sound remain 120ms,
- 3) Selecting a standard value Base. Binarizing the values of the power spectrum and set as input vector. If the value is larger than Base, it would be set 1, or it would be set 0. Making this result as the input network mode vectors.

Except for building the input vector, the calculated power spectrum values of input sound signals are also used for ripples color conversion. First we find the maximum value for each frame of the sound signal. Then we set the appreciate threshold for power spectrum. The radius of concentric circles can be expressed as the form of dB using  $20\lg(P(n,k))$  and different colors are designed to represent different ranges of values of  $dB(P)$ . The relationship of value of power spectrum and color is listing as table1.

Table1 Relationship between power density and color

Power density [dB]	0~10	10~20	20~30	30~40	40~50	50~
Color	black	blue	cabbage green	green	yellow	red

**Determining the Location of the Sound Source based ART1 neural network.** The Adaptive Resonance Theory was introduced by Grossberg[7]. ART nets are designed to control the degree of similarity of patterns place on the same cluster unit. An ART network composes of two layers: an input layer and an output layer. There are no hidden layers. This algorithm can automatically find the adaptive clusters based on training patterns. The structure and algorithm of ART1 are complicated and have been described and illuminated in detail in reference[8], the steps are described as follows:

- i) Modifying Format of input data to a special format according to the kind of ART-Network;
- ii) Setting Input layer activity to modified input data values; Compute output neurons activities (net activities,tj)
- iii) Find up-dated winner neuron(J) with higher activity comparing with other neurons
- iv) Comparing input data with saved templates and finding the best fitting template ( $W_j$ );
- v) Check similarity to achieve resonance in similarity  $\geq$  vigilance  $\rho$
- vi) If similarity between best fitting template and input data belongs to  $[p,1]$  after resonance, adaption of template  $W_j$  towards modified input, otherwise create a new cluster that current input data is the first template of it.

Under normal circumstances, the locations of sound objects are usually fixed in a particular environment. If the users can identify the sound they monitor, the locations of sound sources can be easily found. In this paper, the process for determining the location of the sound source was realized by ART1 neural network. For some sound sources of which locations can be determined, we draw them on the corresponding positions in the ripple. The others that we don't know their locations are displayed in a fixed position.

The basic software tool we used for this work is programmed by ourselves with Visual C++ in Windows XP. Other MATLAB programs were specifically developed to check and validate the accuracy of our computational results. In the experiment, we use four kinds of sound to make demos of ART1 neural network identification. They are telephones, doorbells, clock alarms and other unknown people conversation voice. The four kinds of sound monitored by power spectrum are

displayed in the following Fig.1.(a) is the result of the phone ring tones.(b) is clock alarms.(c) is the doorbell rings.(d) is voice. From the display results, we could clearly see the distinction shapes of four kinds of sound. The non-speech sound can be distinguished by sample power spectrum easily.

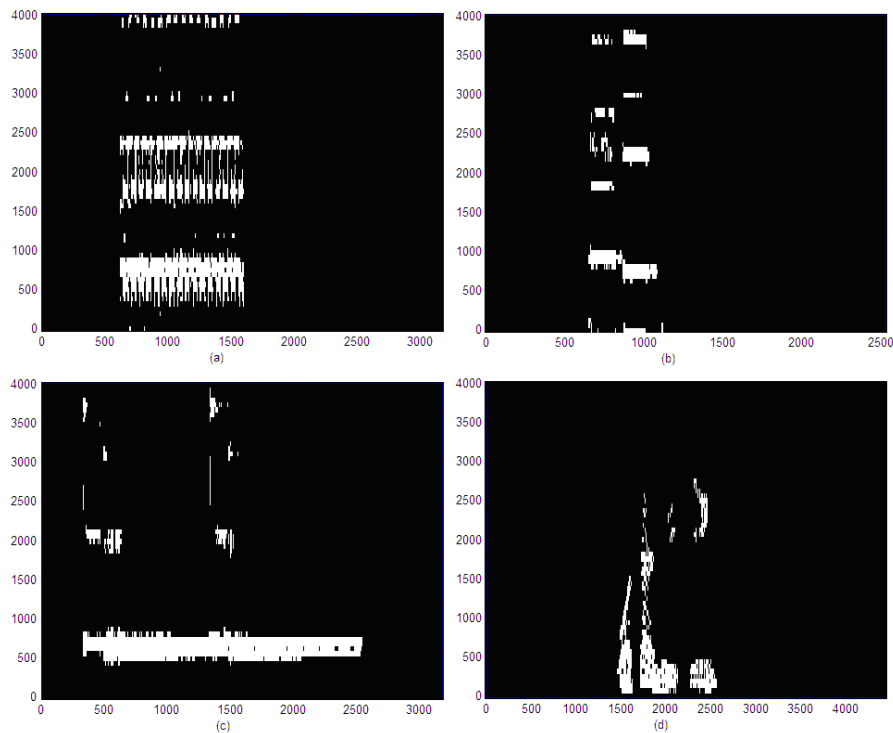


Fig. 1. Spectrogram shapes of various sounds

In the experiment, we define the fixed location for different peripheral sound. The users exist in the specific room as Fig.2. The different location is set as different sound happened place, where location 1 is defined as telephone exist place, location 2 is defined as clock, location 3 is defined as door place and location 4 is defined as other sound such as people's voice.



Fig. 2. Top view of room

In the experiment, we put these four categories of sound samples into ART1 network for training. The analysis frame length WINSIZE of the audio sound is 128. The number of analysis frame NUM is 15. The number of input-layer nodes of the network is set as  $(WINSIZE/2+1)*NUM$ . The number of output layer node is 50. The warning parameter  $\rho$  of ART1 neural network is defined by several trials. Every kind of sound is selected 20 sets of training data for network. The four kinds of sound monitored by the ripple are displayed in the following Fig.3.(a) is the result of the phone ring tones.(b) is clock alarms.(c) is the doorbell rings.(d) is voice. From the display results, the ripple showed the monitoring results envisaged. This suggested that the proposed shape feature of the power spectrum is valid and ART1 neural network has made correct identifications according to this feature.



Fig. 3. Monitoring and displaying results in ripples mode

## Conclusion

Based on a survey of existing peripheral displays and cognitive science literature, we have characterized peripheral displays according to their use. We mainly studied the method of real-time monitoring and displaying the sound for the deaf in the existing environment. We developed a tool to support the room of peripheral displays, based on power spectrum and ART1 neural network. In addition, we also do some perceptual experiments to verify the effect of the proposed method.

We gratefully acknowledge the support from the Fundamental Research Funds for the Central Universities(No. N100304008), and also thank all the colleagues for their contributions to experiments and data analysis.

## References

- [1] Somervell J., An evaluation of information visualization in attention-limited environments, In Proceedings of IEEE TCVG Symposium on Visualization. (2002) 211-217.
- [2] Information on <http://www.hearingdogs.org.uk>
- [3] Harvey Dillon. Hearing aids, Thieme, New York, 2001.
- [4] Hsieh G., Mankoff J., A Comparison of Two Peripheral Displays for Monitoring Email: Measuring Usability, Awareness, and Distraction, University of California, Berkeley Technotes. 3(2003) 286-293.
- [5] Tara Matthews, Janette Fong, F. Wai-ling Ho-Ching, and Jennifer Mankoff, Evaluating non-speech sound visualizations for the deaf, In Behaviour and Information Technology. 25(2006) 333-351.
- [6] Tara Matthews, Scott Carter, Carol Pai, Janette Fong, Jennifer Mankoff. Evaluating a mobile sound transcription tool for the deaf, In Proceedings of the international conference on Ubiquitous Computing. (2006) 159-176.
- [7] S. Grossberg. Adaptive pattern classification and universal recording: I. Parallel development and coding of neural feature detectors. Biological Cybernetics. 23(1976) 121-134.
- [8] S. Grossberg. Adaptive pattern classification and universal recording: II. Feedback, expectation, olfaction, illusions. Biological Cybernetics. 23(1976) 187-202.