

Research on the Key Technologies of Broadcasting Audio Quality Automatic Detection

Qi Wang^{1,*} and Xiao Han²

¹InternetInformation Research Institute, Communication University of China, Beijing100024, China

²Collaborative Innovation Center, Communication University of China, Beijing, China

*Corresponding author

Abstract—As an important part of broadcast business, Broadcast audio quality testing guarantee the broadcast quality, it can find out audio quality problem during broadcasting process such as stereo phase inversion, too large of left and right level difference, mute and overload etc. The traditional detection method of artificial technical is time-consuming and labor-intensive, which can easily lead to miscarriage of justice and leakage of judgment. This paper is for the above problems, through the specific audio quality check business process analysis, combined with the relevant digital signal processing technology and computer-aided analysis of algorithm design, and through the test experiment completed the audio quality of the automated analysis and processing algorithms to achieve a better analysis and detection effect.

Keywords—broadcast audio, audio technology audit, audio quality detection

I. INTRODUCTION

With the widespread application and development of computer technology in the field of radio, audio quality audit business based on quality detection technology have been moving toward automation and precision. The quality of audio directly affects the broadcast quality of radio programs. The low-quality audio will cause the safety accident to be broadcasted. However, the traditional manual audit methods are time-consuming and labor-intensive. With computer-assisted means, the automatic detection of several key audio quality indexes can be greatly reduce labor costs and improve the detection efficiency and accuracy of detection. Audio quality detection mainly involves the testing of several key indicators of stereo phase inversion and left and right level difference detection, in order to achieve the precise positioning for audio quality problems.

II. STEREO REVERSE DETECTION

A. Stereo Reverse Introduction

Radio programs usually use the left and right channel stereo broadcast, for the stereo audio signal, there is a phase difference between the left and right channels, known as the phase difference [1]. The phase difference here is actually a poor polarity, only when the waveform signal is sine, it is called the phase difference. When two waves propagate in the same medium and overlap, according to the principle of superposition of waves : If the wave peaks (or troughs) of two waves reach the same place at the same time, the phase difference between the two waves is 0, the same phase, the

interference wave will have the greatest amplitude, known as the constructive interference; if one of the wave peak and the other waves at the same time arrives at the same location, then the two waves with a phase difference of 180° , said the two waves at this point Phase, interference wave will produce the smallest amplitude, known as destructive interference. In many cases, the phase difference between two waves is between 0° and 180° . At this time, the superimposed waveforms will be phase-matched or canceled according to the specific phases. For details, see the wave superposition principle.

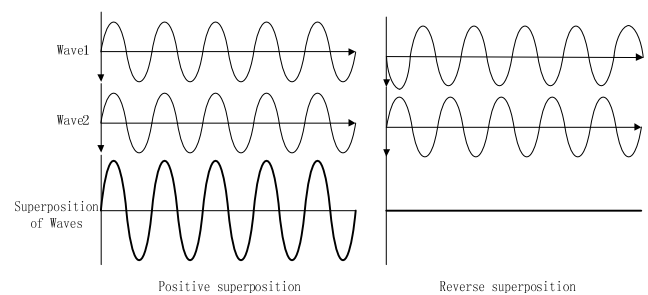


FIGURE I. WAVE SUPERPOSITION

For stereo, if the left and right channels are inverted, the energy of the sound between them will be weakened and canceled each other in the sound field, so that the total energy of the audio is decreased and the volume of the sound is not reached the standard, the sound intensity Worse, bass turbidity and other phenomena [2]. For the detection of stereo inversion, the phase difference between the left and right channels of the sampled data is mainly required. Common methods of measuring the phase are: zero comparison method, digital correlation method, Fourier transform spectrum analysis method.

Zero-crossing comparison is a hardware implementation of the detection method, by counting the two pulses through the zero-point time, T1 and T2, according to the time difference between the pulse period and the phase difference between the two signals [3]. Cross-zero comparison method is good linear, easy to digitize, but it is based on the hardware testing, and vulnerable to external factors, such as the impact of harmonic interference.

The digital correlation method finds the phase difference between two related signals by using the autocorrelation function and the cross-correlation function [4]. Fourier transform spectrum analysis Discrete Fourier Transform (DFT)

principle, the time domain signal is transformed to the frequency domain, and then calculated according to the phase-frequency signal phase difference.

The digital correlation method and the DFT spectrum analysis method have strong anti-interference ability, and both of them use the orthogonality of the trigonometric function. Under the synchronous sampling conditions, the accuracy is higher and the calculation is more suitable for accurate calculation [5]. DFT spectrum analysis method in seeking the phase difference, but also simultaneously determine the signal frequency, amplitude and other parameters, the project is more widely used.

Due to the huge amount of calculation of ordinary DFT, this paper uses the Fast Fourier Transform (FFT) method to improve the speed.

B. Fast Fourier Transform

The Fourier transform of finite length sequence $x(n)$ of length N is:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j \frac{2\pi}{N} nk) = \sum_{n=0}^{N-1} x(n) W_N^{kn}, k = 0, 1, 2, \dots, N-1 \quad (1)$$

From Equation 1.1, we know that for every k value in the sequence, the Fourier transform requires N complex multiplications and $N-1$ complex additions. For the entire sequence, you need N^2 complex multiplications and $N(N-1)$ complex additions. When N is large, the amount of computation is amazing, such as 1024576 multiplications when $N = 1024$. The use of ordinary DFT calculation stereo phase difference less efficient, often used in practical applications of improved Fast Fourier Transform (Fast Fourier Transform, FFT). Fast Fourier Transform (FFT) is not a new theory of Fourier analysis[6], but a general term of various algorithms that improve the efficiency of the algorithm by improving the Fourier transform, reducing the computational complexity, and improving the efficiency of the algorithm. Because the transformation factor W_N of DFT is periodic and symmetrical, that is:

$$\text{Periodicity: } W_N^{m+kN} = W_N^m, k \text{ is a positive integer} \quad (2)$$

$$\text{Symmetry: } W_N^{-m} = W_N^{N-m} \quad (3)$$

In FFT, a N -sequence (assuming $N = 2k$, k is a positive integer) is divided into two subsequences of $N/2$ terms using the periodicity and symmetry of the transform factors. DFT transforms are used on two subsequences of $N/2$ terms, each requiring $(N/2)^2$ operations. Then use N operations to combine the two $N/2$ DFT transforms into one N -point DFT transform. After this transformation, the total number of operations becomes $N + 2 * (N/2)^2 = N + N^2/2$. Thus, when $N = 1024$, the total number of multiplications becomes 525312 times, saving about 50% of the computation.

According to the above method, the two $N/2$ subsequences are further divided into four $N/4$ sub-sequences and eight $N/8$ sub-sequences until the two to two groups of DFT

operation units, Then N -point DFT transformations only require $N \log_2 N$ multiplications. When $N = 1024$, the complex number of calculations only 10240 times, compared with the DFT method increased by about 100 times. The advantage of FFT lies in that the larger N is, the more computation is saved, which is why more FFTs than DFTs are used in practical projects.

C. Detection Algorithm Process

Take the mono s48 file as an example, the phase reversal detection process can be described as follows:

- 1) According to the actual broadcast needs of radio stations, set the time threshold and phase threshold;
- 2) Set the flag bit LastAnti to identify whether the previous frame is in anti-phase, set the Anti structure, save the start position of anti-phase (Start) and end position of anti-phase (End), set AntiList to save all anti-phase Anti;
- 3) Set LastAnti to False;
- 4) Read each frame of audio in turn, perform the following operations on each frame of audio:
 - a) Using fast Fourier transform(FFT) to calculate the phase difference PhaseDiff between the left and right channel;
 - b) Determine PhaseDiff is less than the phase threshold;
 - c) When PhaseDiff is less than the phase threshold: If the previous frame is also smaller than the phase threshold, add the duration of the frame to the End of the AntiList, otherwise create a new Anti element at the end of the AntiList with the start time set to The start position of the frame, the end time is set to the beginning of the frame plus the duration of the frame, and finally the LastAnti set to True.
 - d) When PhaseDiff is greater than the phase threshold, the element is deleted if the previous frame is smaller than the phase threshold and the difference between start time and end time of the AntiList is less than the time threshold. At Last, LastAnti is set to False.
- 5) After the audio detection of the last frame ends, it is determined whether the end of the AntiList is different from the start threshold by less than a threshold. If the difference is smaller than the threshold, the element needs to be deleted.

The reverse detection process is shown in Figure 2.

D. Algorithm Test and Result

In order to verify the correctness and accuracy of the reverse phase detection algorithm, this paper uses Microsoft Visual Studio 2010 software to create the console application of reverse phase detection, using C# language to achieve the detection algorithm. The test audio sequence used in the experiment is shown in Table 1. The test audio sequences are from the network, after the audio editing software Audacity processing, the sequence of each audio random distribution of 10 inverted regions. In the experiment, set the time threshold to 5 seconds and the inversion threshold to -0.7.

TABLE I. STEREO INVERTING TEST AUDIO SEQUENCE

No.	Test Audio File	No.	Test Audio File
1	AntiphaseTest1.wav	6	AntiphaseTest1.s48
2	AntiphaseTest2.wav	7	AntiphaseTest2.s48
3	AntiphaseTest3.wav	8	AntiphaseTest3.s48
4	AntiphaseTest4.wav	9	AntiphaseTest4.s48
5	AntiphaseTest5.wav	10	AntiphaseTest5.s48

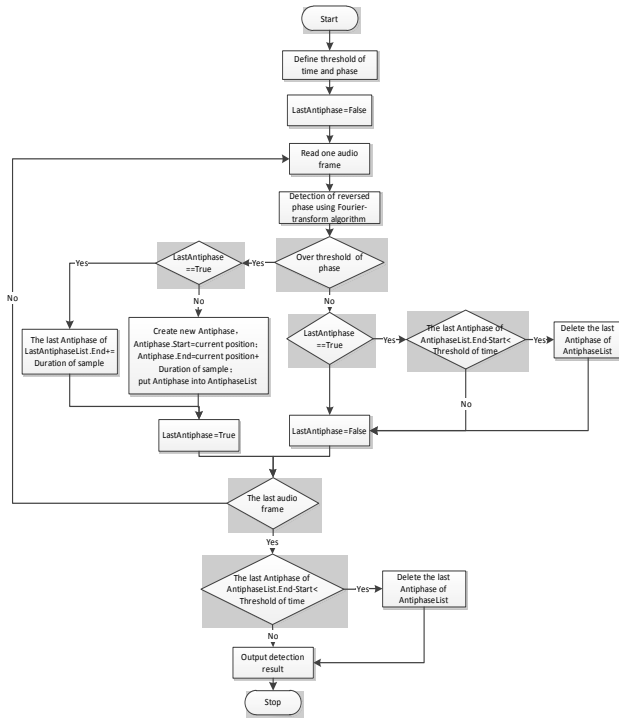


FIGURE II. REVERSE PHASE DETECTION ALGORITHM FLOW

Test audio AntiphaseTest1.wav anti-phase test results is shown in Figure 3, the test results in the reverse phase region and the actual comparison available experimentally detected the number of correct inversion region is 10, a false positive, missed 0, As shown in Figure 3.

The number of Reversed Phase Areas in Test Audio "AntiphaseTest1.wav": 11			
Reversed-Phase-Area1:	StartPosition: 2.925714	EndPosition: 11.7551	
Reversed-Phase-Area2:	StartPosition: 15.77796	EndPosition: 23.77132	
Reversed-Phase-Area3:	StartPosition: 29.25714	EndPosition: 36.98946	
Reversed-Phase-Area4:	StartPosition: 39.18367	EndPosition: 46.2369	
Reversed-Phase-Area5:	StartPosition: 53.86449	EndPosition: 61.12669	
Reversed-Phase-Area6:	StartPosition: 64.83591	EndPosition: 72.15037	
Reversed-Phase-Area7:	StartPosition: 82.54694	EndPosition: 90.95856	
Reversed-Phase-Area8:	StartPosition: 92.16	EndPosition: 99.78793	
Reversed-Phase-Area9:	StartPosition: 104.3331	EndPosition: 112.4835	
Reversed-Phase-Area10:	StartPosition: 122.3053	EndPosition: 130.6124	
Reversed-Phase-Area11:	StartPosition: 134.2694	EndPosition: 140.9569	

FIGURE III. DETECTION RESULT OF ANTIPHASETEST1.WAV

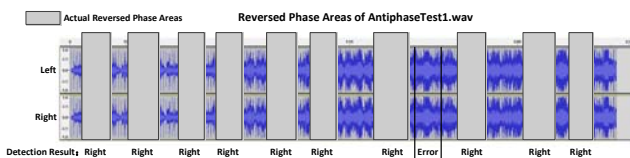


FIGURE IV. COMPARISON CHART BETWEEN DETECTION RESULT AND ACTUAL RESULT FOR ANTIPHASETEST1.WAV

Similarly, do the same for the other test audio in the sequence, count the test results against each test audio by comparing it to the actual inverse domain and plot it to form Table 2.

TABLE II. REVERSE PHASE DETECTION STATISTICS

No.	Error Detection Area	Missing Detection Area	No.	Error Detection Area	Missing Detection Area
1	1	0	6	0	1
2	0	0	7	1	0
3	1	0	8	0	0
4	0	0	9	0	0
5	0	0	10	0	0

As can be seen from Table 2, the false detection rate of reverse phase detection is 3%, the missed detection rate is 1%, and the reverse phase detection accuracy is high.

III. LEVEL DIFFERENCE DETECTION

A. Level Difference Introduction

Before introducing the level difference, first introduce the digital audio level. Digital audio level refers to the digital audio signal full-scale relative level value. DB Full Scale (dbFS) is a digital audio level unit, full scale (0dbFS) is the highest level of loudness that a digital device can reach. The maximum audio signal encoded by the digital audio signal is the maximum value that can be processed by the system. The ratio of the amplitude of the actual audio signal to the amplitude represented by the maximum code is the full-scale relative level value.

Stereo left and right level difference is about the full scale of the left and right channel signal relative level value of the difference. In stereo playback, if the level difference is too large, there will be sound drift, volume reduction, noise, distortion and other symptoms. The greater the difference between the left and right channel levels, the poorer the sound quality of the audio, and sometimes it can result in a soundtrack or overload failure on a channel [7].

For stereo left and right level difference detection, usually first seek the left and right level difference of each frame, by comparing with the set level difference threshold to determine whether the frame level difference is too large.

Because digital audio is achieved by quantizing the amplitude of the audio signal, taking a typical 16-bit sample as an example, the maximum quantized value is 0x7FFF (the most significant bit is used to indicate positive and negative direction). In this way, a sampling point of digital audio level can be expressed as:

$$Sample_{dbFS} = 20 * \log_{10} \frac{sv}{0.777777}, sv \text{ is the quantized value of the sample point} \quad (4)$$

Taking the s48 file as an example, for a frame of audio of the s48 file, the average quantization value of all the sampling points of the frame is used as the level value of the frame. Because of the positive or negative value of the quantized value, the root mean square ARMS of all sampling points of the frame is usually used to represent the average quantized

value of the frame. Assuming that the frame has N sampling points, there is:

$$A_{RMS}\{x(n)\} = \sqrt{\frac{\sum_{n=0}^{N-1} x(n)^2}{N}}, \quad (5)$$

x (n) represents the quantized value of the sampling point n, n = 0, 1, 2 ... N-1.

For a stereo audio file, the average quantized value DRMS of the left and right channel differences for each frame can be expressed as:

$$D_{RMS}\{x(n), y(n)\} = \sqrt{\frac{|\sum_{n=0}^{N-1} x(n)^2 - \sum_{n=0}^{N-1} y(n)^2|}{N}}, \quad (6)$$

Here, x (n) and y (n) denote quantized values of left and right channel sampling points, respectively, and n = 0, 1, 2 ... N-1. Accordingly, the left and right level difference Block_{dbFS} for each frame is:

$$Block_{dbFS} = 20 * \log_{10} \frac{D_{RMS}\{x(n), y(n)\}}{0x7FFF}, \quad (7)$$

B. Detection Algorithm Process

Taking mono s48 file as an example, the level difference detection process can be described as follows:

1) According to the actual broadcasting demand of radio station, the time threshold and level difference threshold can be set;

2) Set the flag LastDiff identifies whether the level difference between the left and right of the previous frame is too large, set the Diff structure, save the position around the start and end of the level difference, and set the DiffList to save all the oversized left and right level differences ;

3) Set LastDiff to False;

4) Read each frame of audio in turn, perform the following operations on each frame of audio:

a) According to the formula, find the frame around the level difference FrameDiff;

b) Judgment whetherFrameDiff is less than the level difference threshold;

c) When FrameDiff is less than the level difference threshold: If the previous frame is also smaller than the level difference threshold, then add the duration of the frame to the end of the element at the end of the DiffList (End), otherwise create a new Diff element at the end of the DiffList, Time is set to the beginning of the frame, the end time is set to the start of the frame plus the duration of the frame, and finally LastDiff is set to True;

d) When FrameDiff is greater than the level difference threshold, the element is deleted if the previous frame is less

than the level difference threshold and the end of the DiffList element's End and Start difference are less than the time threshold. LastDiff set to False.

5) After the audio detection in the last frame is completed, it is required to determine whether the end of the DiffList and the Start difference are less than the time threshold. If the difference is smaller than the threshold, the element needs to be deleted.

The level difference detection process are shown in Figure 5.

C. Algorithm Test and Result

The experiment uses the same programming environment as the reverse phase detection. The test audio sequence used in the experiment is shown in Table 3. The test audio sequence are from the network, after the audio editing software Audacity processing, the sequence of each audio are randomly distributed 10 level difference area. In the experiment, set the time threshold to 5 seconds and the level difference threshold to 5dbFS.

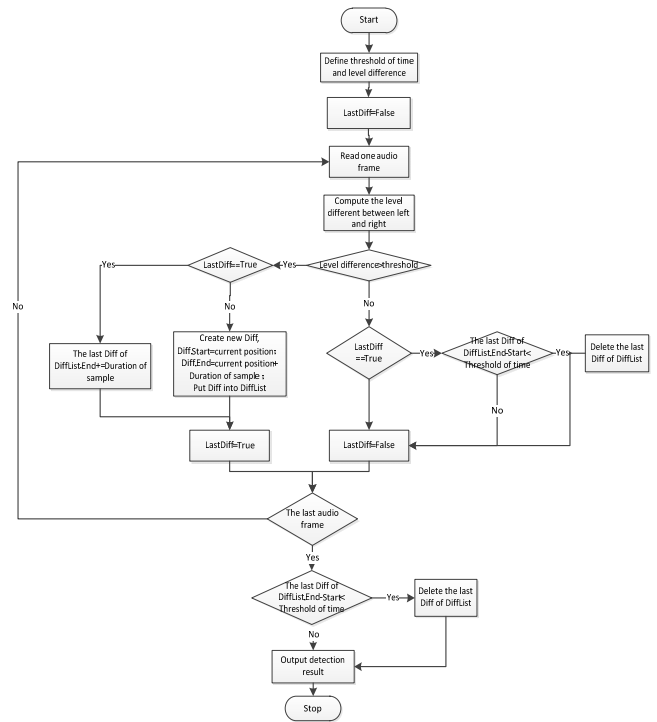


FIGURE V. LEFT AND RIGHT LEVEL DIFFERENCE DETECTION ALGORITHM FLOW

TABLE III. TEST AUDIO SEQUENCE OF LEFT AND RIGHT Level DIFFERENCE

No.	Test Audio	No.	Test Audio
1	DiffTest1.wav	6	DiffTest1.s48
2	DiffTest2.wav	7	DiffTest2.s48
3	DiffTest3.wav	8	DiffTest3.s48
4	DiffTest4.wav	9	DiffTest4.s48
5	DiffTest5.wav	10	DiffTest5.s48

Among them, the test audio DiffTest3.wav left and right level difference test results shown in Figure 6, the test results of the larger level difference with the actual comparison available experimentally detected the correct level difference is too large9 , False positives0, missing ones, as shown in Figure 7.

```
The number of Level Difference too Large in Test Audio "DiffTest3.wav": 9
Too-Large-Area1: StartPosition: 5.694694 EndPosition: 14.26288
Too-Large-Area2: StartPosition: 18.86041 EndPosition: 26.48806
Too-Large-Area3: StartPosition: 29.41388 EndPosition: 39.54953
Too-Large-Area4: StartPosition: 68.70204 EndPosition: 77.11366
Too-Large-Area5: StartPosition: 82.59918 EndPosition: 90.59283
Too-Large-Area6: StartPosition: 98.69061 EndPosition: 107.9382
Too-Large-Area7: StartPosition: 117.3943 EndPosition: 125.5447
Too-Large-Area8: StartPosition: 137.9265 EndPosition: 152.0779
Too-Large-Area9: StartPosition: 159.7649 EndPosition: 170.0051
```

FIGURE VI. LEVEL DIFFERENCE TEST RESULTS OF DIFFTEST3.WAV

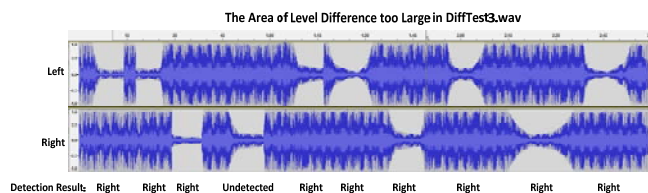


FIGURE VII. COMPARISON CHART BETWEEN DETECTION RESULT AND Actual Result FOR DIFFTEST3.WAV

Similarly, the same for the other test audio in the sequence can be done, and the test results against each test audio by plotting against the actual inverse domain is counted, and plot to form Table 4. It can be seen from the table that the error detection rate of left and right level difference detection is 1%, the rate of missed detection is 2%, and the accuracy of left and right level difference detection is high.

TABLE IV. DETECTION STATISTICS RESULT OF LEFT AND RIGHT LEVEL DIFFERENCE

No.	Error Detection Area	Missing Detection Area	No.	Error Detection Area	Missing Detection Area
1.	0.	0.	6.	0.	0.
2.	0.	0.	7.	0.	0.
3.	0.	1.	8.	0.	1.
4.	0.	0.	9.	1.	0.
5.	0.	0.	10.	0.	0.

IV. CONCLUSIONS

The quality control of the broadcast program for radio stations is the most basic requirement of all radio stations, it needs for the use of appropriate technology to control. The broadcasting of radio programs in our country is not only a technical issue, but strict political requirements. Program audio quality audits to ensure the quality of broadcast safety, which is crucial significance. Aiming at the defect of traditional manual audit, this paper researches the automatic detection of audio based on computer. Due to limited space, this paper focuses on the business process analysis and detection from two aspects of stereo audio reverse detection and left-right channel level difference detection. The design and experiment of the algorithm are compared with the actual situation and the detection accuracy is high. After further proofreading by artificial audio quality testing, the detection efficiency of audio quality of broadcast program will be greatly improved, and the productivity of broadcasting program production will be greatly improved.

ACKNOWLEDGMENT

This paper is supported by the Radio and Television Information Security Research Institute of the Communication University of China. And it is also applied to the Project of 2015 Beijing Science and Technology Project "Broadcast-oriented intelligent audio editing and broadcasting system"、Science and Technology Innovation Project of Ministry of Culture (WHB201520) and Engineering Planned Project of CUC(3132015XNG1523).

REFERENCES

- [1] Kaixiong Ding, Inverted Judgment and Processing of Audio Signal, Sound technology, 2009, 10:16-19
- [2] Hongyu Jia, Wenguo Wu, Hui Cao. Effects of Inverted Signals on FM Stereo Broadcasting, Inner Mongolia Radio and Television Technology, 2012, 29(3): 9-10.
- [3] Yu Lei, Yonghong Tao. Research on Digital Phase Difference Detection Algorithm, Journal of Sichuan University of Science and Technology, 2004, 23(S1): 110-112
- [4] Jing Wu, Haibin Jin. High accuracy phase difference measurement method, Proceedings of the CSEE, 2010 (13): 41-45.
- [5] Min Zhang, Application of Digital Audio Technology in Broadcasting and TV Field, Journalism & Media Studies, 2017(5):98-99
- [6] Meilin Xu, Heming Song. New Data Processing Algorithm and Application Based on FFT Theory, Industrial Control Computer, 2017(7):53-54
- [7] Huijing Miao, Guihua Sun, Hengwei Li. Stereo signal phase difference level difference tester, Electronic Design Engineering, 2012, 20(16): 102-105.