Phone Keypad Voice Recognition (PKVR): An Integrated Experiment for Digital Signal Processing Education

Xiaoyan Luo
Image Processing Center
School of Astronautics
Beihang University
Beijing, China
luoxy@buaa.edu.cn

Lu Li
Image Processing Center
School of Astronautics
Beihang University
Beijing, China
lilu@buaa.edu.cn

Han Wan
School of Computer Science and Engineering
Beihang University
Beijing, China
wanhan@buaa.edu.cn

Fugen Zhou
Image Processing Center
School of Astronautics
Beihang University
Beijing, China
zhfugen@buaa.edu.cn

***Abstract*—This Innovative Practice Work-In-Progress presents an integrated signal processing experiment, which can cover most knowledge points of digital signal processing (DSP) course. Since the DSP course focuses on one dimension signal processing, voice signal has a great advantage. We provide an integrated voice signal processing experiment named as Phone Keypad Voice Recognition (PKVR), including the following parts: phone keypad voice collection, Discrete Fourier Transform (DFT) and analysis, filter design, digital query table establishment, number recognition of any keypad voice. Through the improvement of the practice training, the classroom teaching theory can be better understood in an interesting way for our students.**

Keywords—digital signal processing (DSP); curriculum design; practice; experiment design

I. INTRODUCTION

This WIP practice paper describes the experiment training design of digital signal processing (DSP) course for undergraduate students. With the development of technology in the world, DSP is widely used to obtain valuable information in many fields such as communication, automatic control, imaging acquisition and many other information related fields [1-3]. Therefore, DSP course is always a core course for information engineering related majors.

For DSP discipline, many educators develop the educational projects to keep the curriculum to technological current [4-5]. In DSP teaching, it is the most challenge to find an effective way which can attract the interest of undergraduate students [6]. In general, the students are more motivated to deal with the material in real world, but various commercial system is hard to employ in classroom teaching owing to the system complexity. Thus, how to construct interesting experiment with the help of limited resource is very important for DSP course teaching. Facing the classical DSP theory, its corresponding practical training is fragmented and difficult to design. To make the practical training of DSP course more interesting, we had considered and discussed how to design the practical experiment with an integrated case in our DSP course from AutumnSim 2019.

Compared with traditional classroom teaching method, the students show a more positive attitude to the practical training, especially on their familiar platform such as mobile [7]. In addition, the DSP usually focuses on the signal in one dimension, and then voice signal has a great advantage in DSP teaching projects. For example, a surround sound system is developed by DSP boards [8]. Therefore, in this WIP practice work, we present an integrated voice signal processing experiment named as Phone Keypad Voice Recognition (PKVR), which is an integrated practice framework for training experiment in DSP course.

II. PHONE KEYPAD VOICE RECOGNITION (PKVR) STRATEGY

Considering that the DSP course mainly introduces digital signal acquisition and processing, PKVR is designed to include the following parts: phone keypad press voice signal collection, Discrete Fourier Transform (DFT) and analysis, filter design, digital query table establishment, number recognition of any keypad press voice, whose system framework as shown in Fig. 1. With reference to each part, this PKVR experiment design can cover all the important knowledge points of whole DSP course. For instance, the keypad voice acquisition step is used to illustrate the concept and properties of digital signal. The DFT operation can aid to understand the principle and effect of the frequency transformation and frequency spectrum analysis. Filter design part also helps students reveal the essence of filter. Finally, the DSP system integrity and complete application case is embodied into the establishment of digital query table and the number recognition of any keypad voice part. Moreover, the introduction of these knowledge points is progressive, which is consistent with the experiment progress requirement.

A. Keypad Voice

To acquire keypad press voice, the students can easily use their own mobile phone as experiment acquisition equipment. On one hand, they press the mobile keypad from 0 to 9. On the other hand, they use mobile recorder to collect the corresponding keypad press voice.

To facilitate further processing, each voice signal of one pressed number should be kept in same length and done some

Fig.1. The PKVR system framework

signal denoising preprocessing. For one number, its keypad press voice signal is denoted as $x_i(n)$ with length *N*, and then all the ten numbers can be collected as voice reference set $\{x_i(N), i = 0, 1, \dots, 9\}$.

Through this signal collection operation, the students can be further familiar with the concept and understand the discrete characteristic and representation of digital signals.

B. *Frequency Analysis*

As we known, the keypad press voice is hard to distinguish in time domain, so frequency domain analysis is necessary to the digital signal of each number.

Following the following DFT formula,

$$\mathrm{X}_i(k) = \sum_{n=0}^{N-1} x_i(n)\, e^{-j\frac{2\pi}{N}kn},\ \ k = 0, \dots, N-1 \quad (1)$$

each voice segment $x_i(n)$ can be transformed into frequency domain as $\mathrm{X}_i(k)$.

In our DSP course, students can use any software platform without limitation, and then their course autonomy in the practical programming is ensured and increased. Since this experiment only focuses on the amplitude of frequency domain signal, the DFT signal is calculated and only considered amplitude value. Here, we take the MATLAB platform as an example. It has the fast implementation DFT function fft(*) and amplitude value function abs(*). Therefore, the amplitude value of keypad press voice signal is obtained for each number via the simple MATLAB operation from digital signal.

$$|\mathrm{X}_i| = abs(fft(x_i)) \quad (2)$$

Based on the DFT especially FFT realization, the different meaning of signal in time domain and frequency domain is clear to students.

In this part, we should clarify the relationship between DFT and FFT to students. During using the corresponding MATLAB functions, the students are required to know the programming operation principle and understand the transformation theory well.

	1209 Hz	1336 Hz	1477 Hz	1633 Hz
697 Hz	1	2	3	A
770 Hz	4	5	6	B
852 Hz	7	8	9	C
941 Hz	*	0	#	D

Fig.2. The frequency mode table for DTMF.

C. *Look-up Table*

In DSP teaching, the course display example of practical application is generally lacking. Dual-tone multifrequency (DTMF) is a simple but good example of frequency application, which is a signaling method developed by Bell Laboratories and is a widely used way to transmit telephone dialing information [9]. DTMF is composed of high frequency group and low frequency group. Each frequency group contains four frequencies dividing around at 1000Hz. With a high frequency signal and a low frequency signal, a combined signal is formed via their superposing, which can be used to represent a number. DTMF signal has 16 codes to construct a matrix with the integrating of low and high frequencies as shown in Fig. 2., in which digits from 0 and 9 are included. DTMF is an intuitive frequency application example, which makes the purpose of DFT transformation clear. We introduce it into our course teaching, and then students can understand the significance of frequency transformation easily.

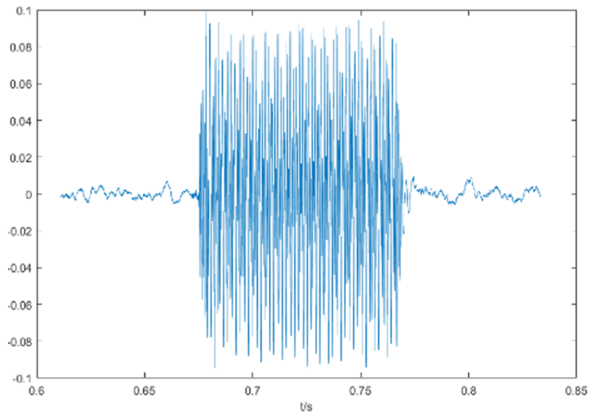
Therefore, to find out the low frequency and high frequency of each keypad press voice, designing filters is an useful method on amplitude value. For this goal, teaching filter design methods is a valuable step, and the students mostly are interested in different filters. Additionally, some software implementation of filters is introduced and then the filtering effect is displayed properly after theory teaching, so that students will have a good grasp of the principle and implementation of the filters.

Based on different filter design method and frequency selection, the students design low-pass and high-pass filters. The amplitude signal $|\mathrm{X}_i|$ is filtered out in low frequency part and high frequency part. For low frequency par and high frequency part, one representative frequency is located respectively to represent the frequency energy concentration point. With the representative frequency point, the students can find out low-high pairs $\{(FL, FH)_i, i = 0, 1, \dots, 9\}$ for all the keypad press voice signals $\{x_i(N), i = 0, 1, \dots, 9\}$. According to these high frequency group and low frequency group, each student can construct their own DTMF look-up table. Referring to the standard DTMF table, the students can compare their own table to analyze the operation process. Fig. 3. shows a toy example to illustrate the representative frequency case, which fits the look-up table in Fig.2.

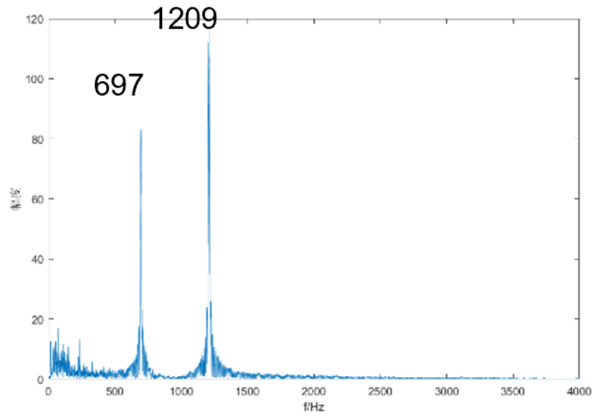
III. STUDENT PERFORMANCE EVALUATION

To ignore the impact of noise, we measure the practical performance focusing on the operation implementation for each student. With the same mobile to each student, the teacher test the PKVR with different keypad press voice input. For each student, the practice performance is measured by the consistency between the actual input numbers with the recognized numbers.

Given the actual input numbers $\mathrm{T} = \{t_j, j = 0, 1, \dots, L\}$ with length L+1, their referred low-high frequency pairs are $\mathrm{FP} = \{(FL, FH)_j, j = 0, 1, \dots, L\}$ on the corresponding look-up table. After the frequency analysis operation with the student program, the numbers are recognized as $\mathrm{T}' =$



(a) A time-domain keypad voice



(b) The frequency amplitude signal of the time domain signal in (a), and with the representative low and high frequency $(FL = 697Hz, FH = 1209Hz)$

Fig. 3. One toy example for time-domain signal and frequency analysis.

$\{t'_j, j = 0, 1, \ldots, L\}$ with estimated low-high frequency pairs $\mathrm{FP}' = \{(FL', FH')_j, j = 0, 1, \ldots, L\}$. To ensure the final evaluation score of student practical performance between 0 and 1, each variable are normalized by their maximum value.

To measure the difference between the actual and estimated number variables A and B with same length M, we define their distance function based on Euclidean distance.

$$\mathrm{D(A, B)} = \frac{\|A-B\|_2}{M} \tag{3}$$

The PKVR system performance firstly depends on the difference from recognized numbers to the actual ones, and then depends on the frequency difference between the found representative low-high frequency pair and the constructed frequency pair in DTMF table. Therefore, the final student performance evaluation is defined as

$$\mathrm{S} = [1 - \mathrm{D(T, T')}] \times [2 - \mathrm{D}(FL, FL') - \mathrm{D}(FH, FH')]/2 \tag{4}$$

The higher the value of S, the better the score of the PKVR for students. Ideally, S=1 is highest score to represent the perfect PKVR system.

From formula (4), we can obviously find that even if the recognized numbers are completely identified correctly, the difference in low-high frequency pair also brings different score, which is very important to evaluate the student performance. That is to say, when the recognized numbers are same, if the filtered low-high frequency pairs have different distance to the referred low-high frequency pair in DTMF table, the score could be different. This evaluation strategy can give higher score to the one nearer to the referred frequency even if their recognized numbers are the same. Therefore, this can be fine to assess the training performance.

IV. Conclusion

Facing the fragmented and difficult problem of the classical DSP course, this WIP work aims to seek one experiment training strategy which can easily attract the interest of student, embed the theoretical knowledge with classroom teaching, operate conveniently with less limitation no matter the software platform and the used equipment.

Based on these considerations, we present an integrated practice framework named as PKVR for training experiment in DSP course, which includes the following parts: phone keypad press voice collection, signal denoising preprocessing, Discrete Fourier Transform (DFT) and analysis, filter design, digital query table establishment, number recognition of any keypad press voice. This experiment design can cover all the important knowledge points of whole DSP course. For examples, the voice acquisition step is used to illustrate the concept and properties of digital signal. The DFT operation can aid to understand the principle and effect of the frequency transformation and frequency spectrum analysis. Filter design part also helps students reveal the essence of filter. Finally, the DSP system integrity and complete application case is embodied into the establishment of digital query table and the number recognition of any keypad press voice part. Moreover, the introduction of these knowledge points is progressive, which is consistent with the experiment progress requirement.

In our DSP course, students can use any software platform without limitation, whose course autonomy in the practical programming is ensured and increased. Simply, they can use their own mobile phone as experiment acquisition equipment to collect their keypad press voice. Each number voice signal should be kept in same length, which is transformed into frequency domain by DFT. According to the frequency distribution, the students need to design low-pass and high-pass filters. Based on dual-tone multi-frequency strategy, digital query table is constructed referring to the filtered high frequency group and low frequency group. After these operations, the teacher can test the PKVR of each student with different keypad press voice input. For each student, the practice performance is measured by the consistency between the actual input numbers with the recognized numbers.

This paper presents the rationality and detail about PKVR experiment setup. We only tried with some students in AutumnSim 2019, and we plan to use it in our course in AutumnSim 2020. The students who joined our DSP-PKVR experiment test reflected that PKVR could be very interesting and very helpful to understand the theory. In our future work, the connections to best teaching and learning practices should be pay more attention in our DSP course.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (Grand No. 61907002). The authors wish to thank our research team and collaborators for their contributions to the project. The authors also wish to thank the course Assistant: Zhang Kexin, for her contributions to the classroom teaching and practical training of DSP course.

REFERENCES

[1] E. Konguvel, M. Kannanv, "A Survey on FFT/IFFT Processors for Next Generation Telecommunication Systems" [J]. Journal of Circuits Systems & Computers, 2018, 27(3).

[2] L. Song, "Image recognition based on SVM information fusion and DSP parallel realization"[J]. Computer Engineering and Application（CEA）, 2009, 3(1):56-7.

[3] C. H. G. Wright, T. B. Welch, D. M. Etter, and M. G. Morrow, "A systematic model for teaching DSP," In Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 4140–4143, 2002.

[4] E. Richter and A. Nehorai, "Enriching the undergraduate program with research projects," IEEE Signal Processing Magazine, vol. 33, no. 6, pp. 123–127, 2016.

[5] A. Moniri, I. Kisil, A. G. Constantinides and D. P. Mandic, "Refreshing DSP Courses through Biopresence in the Curriculum: A Successful Paradigm," IEEE 23rd International Conference on Digital Signal Processing (DSP), 2018

[6] J. P. Hoffbeck, M. M. Sugiyama, "Real-time FM radio for teaching DSP and communication systems," IEEE Frontiers in Education Conference, 2013.

[7] M. K. Banavar, H. Gan, B. Robistow, A. Spanias, "Signal processing and machine learning concepts using the reflections echolocation app." IEEE Frontiers in Education Conference, 2017.

[8] J. Jiang, "Audio processing with channel filtering using DSP techniques," IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), pp. 545–550, 2018.

[9] P. Mock, "Add DTMF generation and decoding to DSP-uP designs," Digital signal processing applications with the TMS320 family, Vol. 1. 1987.