A REPORT ABOUT THE APPLICATION AND DEVELOPMENT ON DSP – 5500 AND ASP SYSTEM

Zu, Yiqing

DSP – 5500 和 ASL 软件系统的使用及开发报告

祖漪清

[提要]　一.DSP – 5500 语图仪和 ASL 软件系统

　　DSP – 5500 语图仪是美国 KAY 电器公司近几年推出的产品,它是一个综合性的实时数字信号处理工作站,它功能齐全,可独立使用,也可与计算机连接,作为计算机的输入输出子系统。ASL 是语音分析、合成软件,它的输入,输出都要通过 DSP – 5500。目前,国内已有好几家单位购买了 DSP – 5500,其中有的同时购买了 ASL(或 CSL)软件,我们中国社会科学院语言研究所于 1992 年初购置了 DSP – 5500 和 ASL 软件,在这一年多的工作中,它们的使用率是相当高的,已成为我所语音分析、合成、识别不可缺少的设备。实践证明,DSP – 5500 用 ASL 是一套十分理想的语音分析工具。

1.DSP – 5500 的主要功能(1,2)

2.ASL 软件的主要功能

二.DSP – 5500 和 ASL 的使用和开发情况

　　DSP – 5500 及 ASL 系统虽然是理想的语音分析工具,但它们毕竟是通用的,对我们的特殊要求不可能全部满足。ASL 存放的语音波形是以两个字节,即二进制补码形式存放在文件中的,操作手册提供了文件头的格式,我们可方便地读取。但 ASL 的分析数据是以文本方式输出给用户的,可直接用 DOS 系统的 TYPE 命令打印出来,我们将它转换成二进制数据,供给我们的程序使用。另外,我们装入 PZP 图形驱动程序,可做各种规格的,各种图形拷贝。

1.利用 DSP – 5500 和 ASL 的若干实例

(1)语音感知实验

(2)切音和编辑

(3)语音学研究

2.利用 DSP – 5500 和 ASL 实现的语音学参数数据管理程序

　　我们利用 DSP – 5500 及 ASL 的输出结果编制的这套声学参数数据库管理程序已投入了使用,它所起的作用是减少手工劳动,提高分析速度,使对大量的数据进行统计成为可能。

　　然而,众所周知,LPC 分析是建立在声道的声管模型基础上的,并假设声道与声源无耦合作用,这些与实际情况是不符合的,因此,尽管 ASL 采用了基频同步等技术,分析结果也不可能是完全如人意的,它的辅音部分(噪声段)也用共振峰及其带宽描述,使我们很难建立它和语图的对应关系,它的过渡部分显得规律不够明显,鼻尾和鼻化的情况下,不是表现出多个共振峰,而是出现峰兼并,以带宽加宽的形式反映给我们,这也为分析造成了障碍。

ABSTRACT

DSP – 5500 and ASL software, the working stations of digital signal processing with complete functions, are new produce by Kay Elemetrics USA. They can be operated independently as well as combined with each other. We have done many phonetic experiments and have been managing monosyllable, disyllable data in our database by using DSP – 5500 and ASL with high efficiency and less handwork. One and half year's application in Lab. proves that DSP – 5500 and ASL are more ideal equipments than others for speech processing.

Ⅰ. AN INTRODUCTION TO DSP – 5500 AND ASL

1. The main functions of DSP – 5500(1)

DSP – 5500, a working station of digital signal processing with complete functions, is new produce of Kay Elemetrics USA. It can be operated independently as well as combined with a computer.

ASL is a software system on speech signal processing, its input and output must use DSP – 5500. We purchased DSP – 5500 and ASL in early 1992. One and half year's application proves that DSP – 5500 and ASL are ideal equipments for speech processing.

DSP – 5500 can display three dimensional sonagraph which is called golden display in real time by means of FFT, the analysis frequency is ranged from DC to 32kHz, its input has two passageways and 16 bit/word. The processing speed is 200 times as fast as IBM/ PCAT and 15 times as fast as VAX. Under the condition of 20KHz sampling frequency, a signal 95sec long can be stored in the buffer. Except for the width band and narrow band sonagraph, DSP – 5500 can also extract pitch, amplitude, power spectrum, zero – acrossing and so on. DSP – 5500 provides 20 combined displays of acoustic signals mentioned above, and provide user with some places to modified those combination. On the screen of DSP – 5500, user can measure values of frequency and time freely by cursors.

DSP – 5500 has powerful function on gating and editing, so the signals in buffer can be gated and put together.

DSP – 5500 can communicate with computer by interface (SCSI and IEEE – 488).

DSP – 5500 is applicable to not only speech signal but also every time – vairing signals.

2. The main functions of ASL(2)

ASL is a software on speech analysis and synthesis, it is a part of CSL (Computer Speech Lab.). ASL needs TMS – C25 and must be used with DSP – 5500 which is regarded as input and output system. ASL can extract pitch, amplitude, formants frequencies and bandwidth by LPC analysis. Those parameters can be used to synthesis speech. User can also modify analysis data with keyboard as well as mouse.

The parameters analysized by ASL are satisfactory. As a result of pitch sychronous technique, the loci of formants frequencies are clear enough;owing to that the LPC residual

signal is used as excitation, synthesis speech hears almost the same as the original one.

ASL is friendly to user, one can easily follow the menues step by step.

In biref, DSP – 5500 and ASL are more ideal equipments than others for speech analysis.

II. THE APPLICATION AND DEVELOPMENT OF DSP – 5500 AND ASL

Although DSP – 5500 and ASL are excellent, they are in common use, so, we must develope their functions for our special needs.

The wave data provided for user are in binary, we can easily read it into our program, but the analysis data are in ASCII which can be displied on screen by TYPE command in DOS. We have a program to convert ASCII data into binary data. Besides, we have installed PZP graph drivering program which can help us copy all sorts of graphs.

1. Some examples of using DSP – 5500 and ASL

(1) speech perception test

Put a speech signal into DSP – 5500, gate the usful section by cursor, then translate it into ASL and export to a file with a name arranged in order. By this way, we get a serious stimuli. Using the number of sum of stimuli and repeat time defined by perception test, a program of ours can disorder the arranged files randomly and give a new order. Using the new order, those stimuli are taken one by one by through ASL and translated into the buffer of DSP – 5500, and in the state of "gating and editing" of DSP – 5500, the repeat time of every stimulus and interval between two stimuli in auditory discrimination can be controled. At last, the stimuli with new order are played and recorded into a tape which we can use to carry out discrimination test.

(2) Gating and editing

DSP – 5500 has a powerful function of gating and editing, but it must be under the state of "gating/editing", if you escape from there, the treaded signal disappears. Combining with ASL, we can realize speech signal editing. There are two ways for this work, one is wave editing, we can gate speech signal at DSP – 5500 and store them by throuth ASL, and then edit the gated signal by programming. The other way is parameter editing, in this way, we put gated signal into ASL and have LPC analysis done, we must choose "Appende" under the sub – menu of "Current LPC data" in the ASL menu of "option" before the next operation, i. e. the analysis data will be preserved in the buffer, and the next analysis data will be added to the previous one.

(3) Phonetic research

ASL provides the function of modifying pitch, formant frequencies and bandwidth by a mouse; user can change the curves of the parameters freely and use the parameters to synthesis speech, so, can examine the results by auditory test for the meticulous research on phonetics.

2. A program for management and measurement about acoustic parameter of database by

using DSP‑5500 and ASL

We have started to build a acoustic parameters database of standard Chinese since 1990 which include speech wave and acoustic parameters. By now, we have recorded 1275 monosyllables spoken by 15 males (3) and 3000(4) pairs disyllables spoken by one male. The data are preserved in tapes and floppy disks. There are too much data for us to do phonetic research. Traditionally, we make sonagram at first, then recognize and measure the formants frequencies and so on by hand. It is a research processing that many phonetists have rich experience. But for a great amount of data, it must take too long time and can't meet the needs of the subjected in new period. So we decided to get acoustic parameters by LPC analysis in ASL and refer to the sonagram displied on DSP‑5500. Fig.2 is a flow chart of our program for management and measurement on monosyllables and disyllables. Fig.3 shows acoustic parameters patterns of 15 speakers and 4 tones co‑articulation between two syllables, so we choose formant frequencies at five key points and relative eight durations (5, 6, 7), we only need to measure the position of 5 points, the acoustic parameters will be given automatically. Fig.4 is the measurment screen.

It is necessary to give explaination to some problem about format of wave data (.nsp) and analysis data (.1pr) provided by ASL, the wave data is in binary and it needs postfix ".nsp". User has better to save analysis data (ASCII) in the state "exporting" under the sub‑menu of "data", otherwise we can't easily read them from data files. If the formant data are required, key "F9" must be pressed to exchang k1,k2···km for F1,B1,F2,B2, ··· Fm,Bm, where m is the order of LPC analysis.

The program on management and measurement has been put into operation, it has reduced the bankwork and increased the speed. By this means it is possible to deal with a great amount of data.

As is kown to all LPC is a linear model, the analysis data is impossiablly exact as sonagram, especially the consonant and transition part of signal, two or three formants may merge into one. Those parts need further research.

CONCLUSION

DSP‑5500 is a excellent tool for speech processing, we can efficiently do our experiment with it. But we must be very careful with processing the acoustic data in ASL.

REFERENCES:

Cao,Jianfen (1991), "Temporal structure of diphthongal finals in Standard Chinese" RPR‑IL(CASS).

Chen Xiaoxia (1991), "Acouslic analysis and temporal Perception test of lateral in Standard Chinese" RPR‑IL(CASS).

Instruction /Tutorial Manual, LPC Parameter Manipulation/Synthesis Program, Model 5635 (DSP Sona‑Craph) Software Version 1.X (Issue A,1991), (Issue B,1992)

Maintenance manual, DSP Sona – Graph Model 5500/5500 – 1 (in Chinese)(1990)

Sun Guohua (1991), "Experimental of the acoustic Patterns and perception tests of retroflex vowels in Standard Chinese" RPR – IL(CASS).

Yan, Jingzhu & Sun, Guohua (1991), "Data bank of acoustic parameters in monosyllables of Standard Chhinese" RPR – IL(CASS).
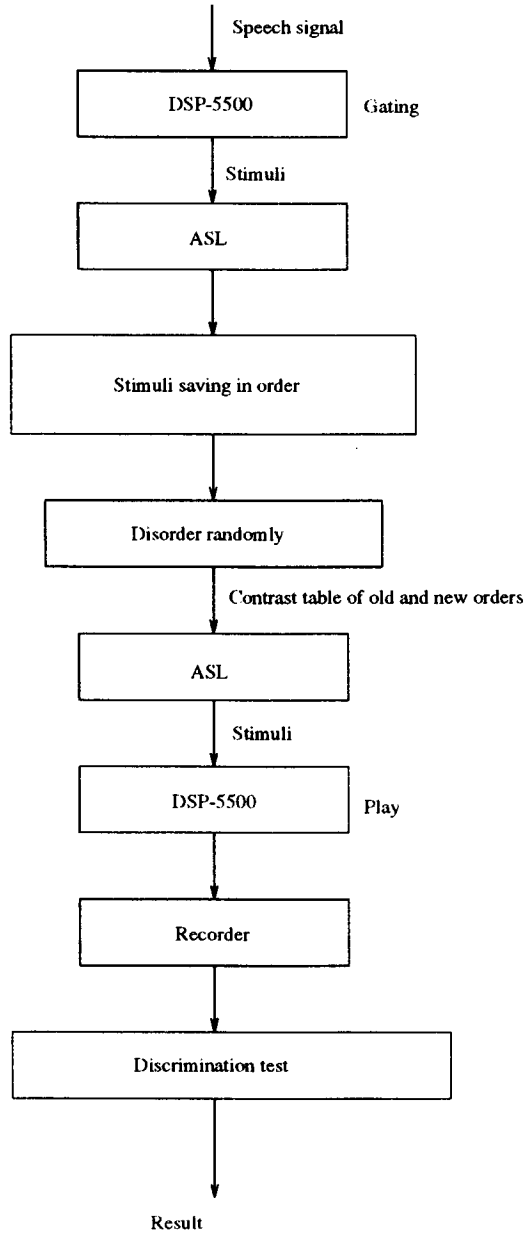
Fig.1 Flow chart of perception test

Speech signal

```
┌──────────────┐
│   DSP-5500   │
└──────────────┘
        │
        ▼
┌──────────────┐         Analysis data        ┌──────────────┐   Disyllabic words
│     ASL      │────────────────────────────▶│ ASCII to Binary│──────────────────┐
└──────────────┘                              └──────────────┘                   │
        │                                            │                            │
        ▼                                   Monosyllabic words                    ▼
┌──────────────┐                                     │                   ┌──────────────┐
│   Database   │                                     ▼                   │  Measurement │
└──────────────┘                            ┌──────────────┐             │    screen    │
                                            │Classify according           └──────────────┘
                                            │to consonants  │                    │
                                            └──────────────┘                     ▼
                                                     │              ┌──────────────┐
                          ┌──────────┐               ▼              │ Five key points│
                          │   ASL    │──▶ ┌──────────────────┐      │measuring by cursor│
                          │          │◀── │Durations calculating     └──────────────┘
                          └──────────┘    │according to the frame of         │
                                          │acoustic phonetics │              ▼
                                          └──────────────────┘      ┌──────────────┐
                                                     │              │Acoustic parameters│
                                                     ▼              │calculating at five│
                                          ┌──────────────┐          │key points     │
                                          │Insertation and│          └──────────────┘
                                          │   ANOVA      │                    │
                                          └──────────────┘                    ▼
                                                     │              ┌──────────────┐
                                                     ▼              │Durations calculating│
                                          ┌──────────────┐          │from five key points│
                                          │Formant patter │          └──────────────┘
                                          │(4 tones and 15│                   │
                                          │speakers)      │                   ▼
                                          └──────────────┘          ┌──────────────┐
                                                                    │  Data saving │
                                                                    └──────────────┘
                                                                            │
                                                                            ▼
                                                                    ┌──────────────┐
                                                                    │    ANOVA     │
                                                                    └──────────────┘
```
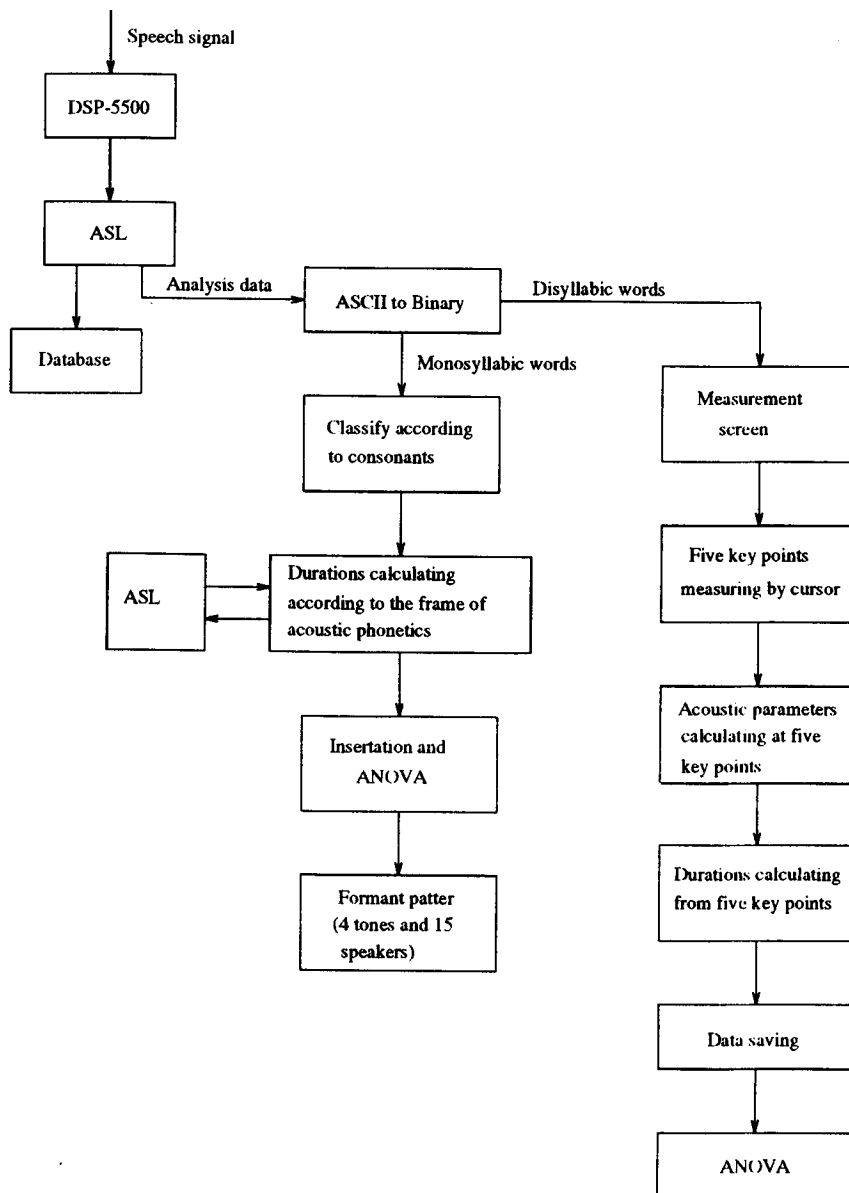
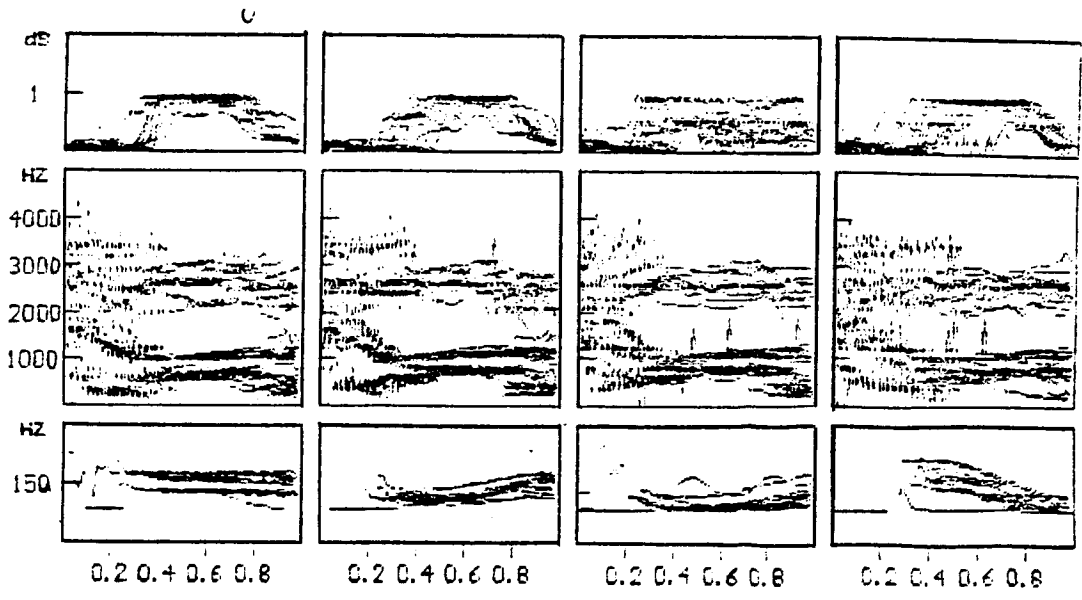Fig.2 Flow chart of management and measurement system of monosyllabic and disyllabic words

Fig.3 Formant pattern off monosyllabic words
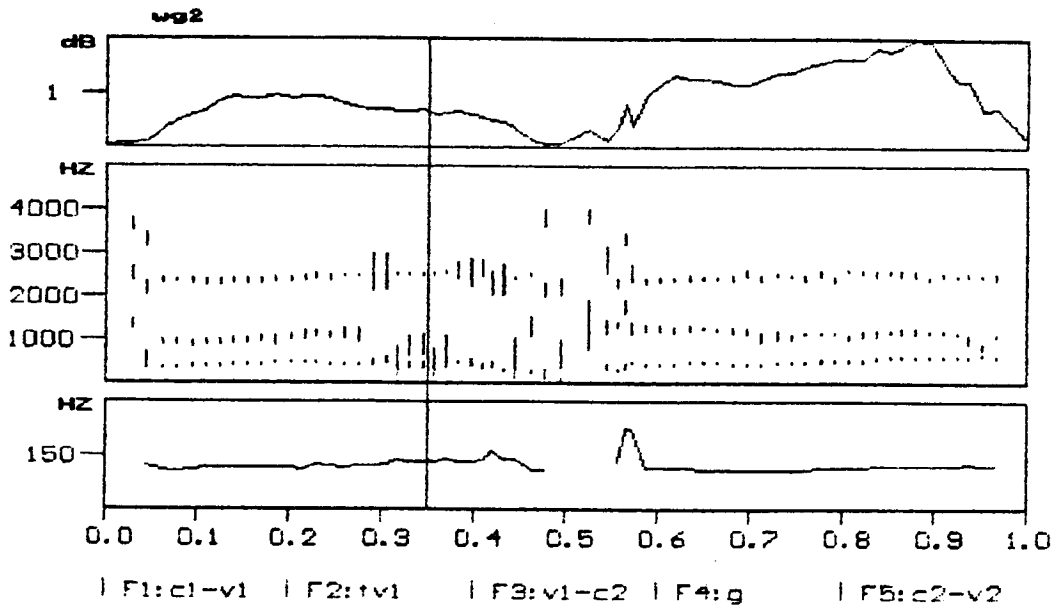(4 tones and 15 speakers in /chuang/)

| F1:c1-v1 | F2:tv1 | F3:v1-c2 | F4:g | F5:c2-v2

Fig.4 Measuring screen of disyllabic words