۱۷۱۰۵

« بنام خدا »

طراحی آزمایشگاه پردازش دیجیتال سیگنالهای صوتی

رامین صفری

پایان نامه تحصیلی برای اخذ درجه کارشناسی ارشد

در رشته

مهندسی مخابرات

مرکز اطلاعات و مدارک علمی ایران
شعبه مدارک

دانشگاه صنعتی اصفهان

دانشکده برق و کامپیوتر

مرداد ۱۳۶۸

۲۴۷

کیفیت وارزش گزارش حاضربعنوان پایان نامه کارشناسی ارشدموردتأئیداست .

دکتربهروزفرهنگ (استادراهنمای پروژه)

کیفیت وارزش گزارش حاضربعنوان پایان نامه کارشناسی ارشدموردتأئیداست .

دکترمهردادفهیمی

" قـدردا نـــی "

بدینوسیلـه‌ا زا ستا دمحترم آ قا ی دکتربـهروزفرهنگ بروجنی کـه‌درتما م مرا حل انجام پروژه‌ا زرا هنما ئیـها وکمکها ی بیدریغ ا یشا ن بـهره‌مندبودم وپا یا ن نا مهٔ حا ضرنیزبـــا همکا ری ا یشا ن تدوین شده‌ا ست ، تشکروقدردا نی کرده‌وا زدرگا ه‌خدا وندمتعا ل ، بـــرای ا یشا ن موفقیت روزا فزون مسئلت دا رم .

همچنین ا زآ قا ی دکترمهردا دفهیمی کـه‌بعنوا ن ا ستا ددوم پروژه‌درا مرتدویـــن و تکمیل پا یا ن نا مه ، نظرا ت وپیشنها دا ت سودمندی ا را ئه‌نموده‌ا ند ، سپا سگزا ری کرده‌و موفقیت وپیروزی ا یشا ن راآ رزومندم .

ضمنا" ا زآ قا ی دکترحسین علوی کـه‌علیرغم مشغولیتها ی فرا وا ن، برا ی ا علام نظر ، درسمینا را ینجا نب شرکت دا شتند ، تشکرمی‌نما یم .

را مین صفـــری

تابستان ۱۳۶۸

« فهرست مطالب »

عنوان	صفحه
فهرست جداول	شش
فهرست شکلها	هفت
چکیده	ده
فصل اول : بررسی مقدمات انجام پروژه	۱
۱-۱ : مقدمه	۱
۱-۲ : ساختار رساله	۶
فصل دوم : طراحی یک سیستم باهوش جهت نمونه برداری و بازسازی سیگنالهای صوتی	۱۰
۲-۱ : مقدمه	۱۰
۲-۲ : سیستم نمونه برداری- بازسازی کننده‌ٔ باهوش	۱۰
۲-۲-۱ : معرفی بلوک دیاگرام سیستم	۱۰
۲-۲-۲ : جزئیات سخت افزاری سیستم	۱۲
۲-۳ : سخت افزار رابط بین کامپیوتر شخصی و سیستم نمونه برداری- بازسازی کننده‌ٔ باهوش	۱۸
۲-۴ : امکانات موجود جهت گسترش سیستم	۲۰
فصل سوم : نمونه برداری و بازسازی	۲۲
۳-۱ : مقدمه	۲۲
۳-۲ : معرفی تراشه 8255	۲۲
۳-۳ : معرفی تراشه Z-80A CTC	۲۷
۳-۴ : اتصال تراشه‌های 8255 به یکدیگر و نرم افزارهای نمونه برداری و بازسازی	۲۸
۳-۵ : نمونه برداری	۳۱
۳-۶ : بازسازی	۳۲
فصل چهارم : مکانیزم سیستم صوتی و آنالیز زمانی	۳۳

۴-۱ : مقدمه ۳۳
۴-۲ : توصیف مکانیزم سیستم صوتی ومعرفی پارامترهای آن ۳۴
۴-۳ : آنالیززمانی سیگنال صحبت ۳۷
۴-۳-۱ : آنالیزانرژی دورهٔ زمانی کوتاه ۳۸
۴-۳-۲ : متوسط نرخ عبورازصفردردورهٔ زمانی کوتاه ۴۱
۴-۳-۳ : تابع خودهمبستگی با دورهٔ زمانی محدود ۴۳
۴-۳-۴ : تخمین پریود Pitch ۵۰
فصل پنجم : آنالیزفرکانسی دورهٔ زمانی محدود ۷۲
۵-۱ : مقدمه ۷۲
۵-۲ : تأثیرپنجره درآنالیزفرکانسی ۷۴
۵-۳ : الگوریتم FFT درمحاسبهٔ نمونه‌های فرکانسی سیگنال صحبت ۷۹
۵-۴ : نمایش اسپکتروگرام سیگنال صحبت ۸۲
فصل ششم : ارتباط استفاده‌کننده‌وبخش نرم افزارسیستم ۹۲
۶-۱ : مقدمه ۹۲
۶-۲ : نرم افزارهای‌سیستم نمونه‌بردار-بازسازی کنندهٔ باهوش ۹۳
۶-۳ : نرم افزارهای پردازش سیگنال صحبت ۹۳
۶-۴ : نحوهٔ استفاده‌ازحافظهٔ RAM کامپیوترشخصی IBM ۹۸
۶-۵ : نحوهٔ برقراری ارتباط بین استفاده‌کننده‌وبخش نرم‌افزاری‌سیستم ۱۰۱
فصل هفتم : نتیجه‌گیری وپیشنهادات برای ادامهٔ کار ۱۰۷
۷-۱ : جمع بندی ونتیجه‌گیری ۱۰۷
۷-۲ : معرفی امکانات سیستم ۱۰۷
۷-۳ : پیشنهادات برای ادامهٔ کاروتکمیل پروژه ۱۰۸
مراجع ۱۱۱

« فهرست جداول »

شماره جدول	عنوان	صفحه
۱ـ۲ :	نقشهٔ حافظهٔ سیستم نمونه بردار ـ بازسازی کنندهٔ باهوش	۱۶
۲ـ۲ :	نقشهٔ ورودی ـ خروجی سیستم نمونه بردار ـ بازسازی کنندهٔ باهوش ..	۱۶
۱ـ۶ :	لیست نرم افزارهای پردازشی کامپیوترشخصی	۹۵
۲ـ۶ :	نقشهٔ حافظهٔ کامپیوترشخصی	۹۹

" فهرست شکلها "

شماره شکل	عنوان	صفحه
۱-۱ :	حداقل امکانات لازم جهت راه اندازی آزمایشگاه صوت	۷
۱-۲ :	بلوک دیاگرام سخت افزاری سیستم نمونه بردار- بازسازی کنندهٔ باهوش	۱۱
۲-۲ :	نحوهٔ بافرنمودن خطوط ریزپردازنده، پالس ساعت ومدار فرونشاننده	۱۳
۳-۲ :	جزئیات سخت افزاری سیستم نمونه بردار- بازسازی کنندهٔ باهوش	۱۴
۴-۲ :	سخت افزار مدار رابط.....	۱۹
۱-۳ :	دیاگرام زمانی حالت ورودی	۲۶
۲-۳ :	دیاگرام زمانی حالت خروجی	۲۶
۳-۳ :	نحوهٔ ارتباط سخت افزاری دوبخش موجوددرمد Handshaking	۲۹
۱-۴ :	مدل سازی آنالیزانرژی دورهٔ زمانی کوتاه	۳۹
۲-۴ :	تغییرات انرژی فریمهای متوالی.....	۴۰
۳-۴ :	نرخ عبورازصفرفریمهای متوالی.....	۴۴
الف-۴-۴:	نمونه‌ای ازسیگنال Voiced	۴۶
ب-۴-۴ :	نمونه‌ای ازسیگنال Unvoiced	۴۷
الف-۵-۴:	تابع خودهمبستگی شکل (الف-۴-۴)	۴۸
ب-۵-۴:	تابع خودهمبستگی شکل (ب- ۴-۴)	۴۹
۶-۴ :	مدل سازی تخمین زنندهٔ Pitch	۵۳
الف-۷-۴:	تابع برش دهندهٔ مرکزی فشرده.....	۵۳
ب-۷-۴:	تابع برش دهندهٔ مرکزی ساده	۵۳

ج ـ۷ـ۴ : تابع علامت..... ۵۳

الف ـ۸ـ۴ : خروجی بلوک غیرخطی برش دهندهٔ مرکزی فشرده به ازای ورودی شکل (الف ـ۴ـ۴) ۵۵

ب ـ۸ـ۴ : خروجی بلوک غیرخطی برش دهندهٔ مرکزی فشرده به ازای ورودی شکل (ب ـ۴ـ۴) ۵۶

الف ـ۹ـ۴: خروجی بلوک غیرخطی برش دهندهٔ مرکزی ساده به ازای ورودی شکل (الف ـ۴ـ۴) ۵۷

ب ـ۹ـ۴ : خروجی بلوک غیرخطی برش دهندهٔ مرکزی ساده به ازای ورودی شکل (ب ـ۴ـ۴) ۵۸

الف ـ۱۰ـ۴: خروجی بلوک غیرخطی تابع علامت به ازای ورودی شکل (الف ـ۴ـ۴) ۵۹

ب ـ۱۰ـ۴: خروجی بلوک غیرخطی تابع علامت به ازای ورودی شکل (ب ـ ۴ ـ ۴) ۶۰

الف ـ۱۱ـ۴: تابع خودهمبستگی شکل (الف ـ۸ـ۴) ۶۱

ب ـ۱۱ـ۴: تابع خودهمبستگی شکل (ب ـ ۸ـ۴) ۶۲

الف ـ۱۲ـ۴: تابع خودهمبستگی شکل (الف ـ۹ـ۴) ۶۳

ب ـ۱۲ـ۴: تابع خودهمبستگی شکل (ب ـ ۹ ـ ۴) ۶۴

الف ـ۱۳ـ۴: تابع خودهمبستگی شکل (الف ـ۱۰ـ۴) ۶۵

ب ـ۱۳ـ۴ : تابع خودهمبستگی شکل (ب ـ۱۰ ـ ۴) ۶۶

الف ـ۱۴ـ۴: پریود Pitch پس از اعمال تابع غیرخطی برش دهندهٔ مرکزی فشرده روی ۶/۵ ثانیه متن صحبت ۶۸

ب ـ۱۴ـ۴: پریود Pitch پس از اعمال تابع غیرخطی برش دهندهٔ مرکزی ساده روی ۶/۵ ثانیه متن صحبت ۶۹

ج ـ۱۴ـ۴: پریود Pitch پس از اعمال تابع غیرخطی تابع علامت روی ۶/۵ ثانیه متن صحبت ۷۰

د ـ ۱۴ـ۴ : پریود Pitch بدون اعمال بلوکهای غیرخطی روی ۶/۵ ثانیه متن صحبت ۷۱

۱ ـ ۵ : تغییرات زمانی پنجره‌های مختلف ۷۵

۲ ـ ۵ : طیف فرکانسی پنجره‌های مختلف ۷۷

۳ ـ ۵ : بلوک دیاگرام ترسیم کننده‌ٔ اسپکتروگرام به‌روش آنالوگ ۸۴

۴ ـ ۵ : اسپکتروگرام باندعریض ۸۶

۵ ـ ۵ : اسپکتروگرام باندباریک ۸۶

۶ ـ ۵ : اسپکتروگرام باندباریک برروی چاپگر ۸۹

۷ ـ ۵ : اسپکتروگرام باندعریض برروی چاپگر ۹۰

۱ ـ ۶ : فلوچارت نرم افزاری سیستم ۱۰۲

" چکیـــده "

هدف از رسالهٔ حاضر طراحی و ساخت امکانات لازم جهت پردازش دیجیتال سیگنالهای صوتی است . در این راستا از یک کامپیوتر شخصی ، تحت عنوان کامپیوتر پردازشگر ، که نقش پردازش دیجیتالی صوت را بعهده دارد استفاده شده است . سیستمی نیز به منظور نمونه‌برداری و بازسازی سیگنال صوت ، مبتنی بر ریزپردازندهٔ Z-80A ، طراحی و ساخته شده است . طراحی و ساخت این سیستم بصورتی بوده است که به فرم کاملا" هوشمند عمل نموده و قابلیت انعطاف نرم افزاری و سخت افزاری مناسبی را دارا می باشد. سیستم مذکور توسط کامپیوتر پردازشگر برنامه ریزی شده و در مدمناسب عملیاتیش قرار می‌گیرد . ارتباط این سیستم با کامپیوتر پردازشگر بواسطهٔ یک مدار رابط ، که درون کامپیوتر پردازشگر تعبیه شده است ، صورت می‌گیرد . سیستم موجود قابلیت نمونه برداری و بازسازی سیگنال آنالوگ صوت را تحت فرکانسهای ۸ الی ۳۰ کیلوهرتز دارد . در طراحی موجود قابلیت ضبط حداکثر 2^{16} نمونهٔ هشت بیتی صوت ، درون حافظهٔ کامپیوتر پردازشگر ، وجود دارد . بواسطهٔ وجود این نمونه‌ها درون حافظهٔ کامپیوتر پردازشگر ، امکان تحقیق در زمینه‌های متنوع پردازش دیجیتالی صوت فراهم می‌آید . نرم افزارهای پردازشی مناسبی نیز بر روی نمونه‌های صوتی صورت گرفته‌اند . این نرم افزارهای پردازشی نمونه‌هائی از الگوریتمهای آنالیز زمانی و آنالیز فرکانسی ، جهت دستیابی به پارامترهای سیستم صوتی می‌باشند .

** فصل اول

" بررسی مقدمات پروژه "

۱-۱ : مقدمه

هدف از رسالهٔ حاضر ، طراحی وپیاده‌سازی سخت افزاری ونرم افزاری یک آزمایشگاه پردازش دیجیتالی صوت می‌باشد . در ابتدا برای آماده‌نمودن ذهن خواننده مقدمه‌ای را در رابطه با سیگنال صوت ارائه می‌دهیم .

سیگنال صوت یک سیگنال باندعریض[1] است که‌درحالت خاص سیگنال صحبت می‌تواند محتوای فرکانسی آنرا درمحدودهٔ صفرتا چهارکیلوهرتزدانست . سیگنال صحبت توسط ارتعاش تارهای صوتی ویا اعمال فشارهوادرون سیستم صوتی ، بعنوان منبع محرک ،پس ازعبورازحنجره ، گلوودهان همراه‌با تأثیرپذیری اززبان، لب ودندان‌تولیدمی‌شود [۱] .

ازدیدگاه‌تئوری اطلاعات که‌بصورت بسیارکیفی با مسئله‌برخوردمی‌کند ، نمایش مناسبی ازسیگنال صحبت محتوای پیام یا اطلاعات موجوددرآن می‌باشد* [۱] . با این دیدمی‌توان حداقل‌سرعت ارسال برای یک پیام را ، با درنظرگرفتن تعداد بیتهای لازم برای آن پیام ، وقتی حروف مربوط به‌آن کدمی‌شوندوبا توجه‌به‌سرعت معمول خواندن متن ، حدود ۵۰ الی ۶۰ بیت درثانیه‌بدست‌آورد . البته‌بایستی توجه‌داشته‌باشیم که نتیجهٔ اخیربدون توجه‌به‌لحن صحبت ، بلندیاکوتاهی آن ویا تأکیدبرکلمات خاص و بالاخره‌هویت صحبت کننده‌گرفته‌شده‌است .

درآنالیزوپردازش‌سیگنال صوت دوروش‌کلی آنالوگ ودیجیتال وجوددارد. روشهای

---

1. Wideband

* این مسئله‌توسط Shannon موردتوجه‌قرارگرفته‌است .

آنالوگ دارای پیچیدگی زیادی هستند. بدلیل تنوع تکنیکهای پردازش دیجیتالی سیگنال و امکان دسترسی به‌کامپیوترها درجهت پیشرفت تکنولوژی پردازش دیجیتال، در دودههٔ اخیرتوجه‌بسیاری ازمحققین به‌سمت پردازش دیجیتالی سیگنال صحبت معطوف شده است. درپردازش دیجیتالی سیگنال صحبت، با نمونه‌برداری ازسیگنال صحبت و ذخیره‌نمودن نمونه‌ها درحافظهٔ کامپیوتربه‌قابلیت انعطاف زیادی درجهت پردازش و استخراج مشخصات سیگنال صوت می‌توان دست یافت[۱۰].

دلایل دیگری نیزبرای استفاده‌ازتکنیک دیجیتال درپردازش سیگنال صحبت وجود دارد[۱]، بطورمثال اگرکدینگ مناسبی موردا ستفاده‌قرارگیرد، سیگنال صحبت درفرم دیجیتال را می‌توان بفرم قابل اطمینانی درکانالهای نویزی ارسال نمود. ازطرفی سیگنال صحبتی که‌بفرم دیجیتال درآمده‌باشد، فرمی مشابه‌دیگراطلاعات دیجیتال داشته وازاینرومدارمخابراتی موردا ستفاده می‌توانند هردونوع اطلاعات راارسال نماید. در ضمن اگرامنیت مخابرهٔ سیگنال صحبت موردتوجه‌باشد، فرم دیجیتالی سیگنال صحبت بهترین وراحت‌ترین فرم موجودبرای ارسال اطلاعات امن خواهدبود.

درنمایش سیگنال صحبت بفرم دیجیتال، اساس کارتبعیت ازقضیهٔ نمونه‌برداری[1] است. این قضیه‌بیان کنندهٔ این موضوع می‌باشد که سیگنالهای باباند محدود را می‌توان توسط نمونه‌های متوالی زمانی این سیگنال، که‌با سرعت مناسبی نمونه‌برداری شده‌اند، نمایش داد[۹].

روشهای متنوعی برای نمایش گسستهٔ سیگنال صحبت وجوددارد. این روشها را می‌توان به دو بخش عمدهٔ زیرتقسیم نمود[۱]:

الف) نمایش شکل موج: دراین نمایش هدف حفظ شکل موج سیگنال صحبت در پروسهٔ نمونه‌برداری وکوانتیزاسیون می‌باشد.

---

1. Sampling theorem

ب) نمایش پارامتری : در این نوع نمایش هدف اصلی نمایش سیگنال صحبت بفرم خروجی یک مدل که نقش تولید صوت را دارد می باشد .

درحالت اخیر ، پس از نمونه برداری سیگنال صحبت ، براساس نمونه های حاصل پارامترهای لازم برای توصیف سیستم مولد صوت محاسبه و معرفی می گردند . پارامترهای این مدل را بسادگی می توان به دو قسمت پارامترهای تحریک ، در ارتباط با منبع محرکهٔ تولید صوت ، و پارامترهای پاسخ سیستم صوتی ، در ارتباط با تولید اصوات مختلف و خود مدل، تقسیم بندی نمود*.

در مورد تکنیکهای پردازش دیجیتال ذکر این نکته ضروری است که بسته به مورد استفادهٔ موجود و حالت خاص مورد نظر بایستی انتخابی صحیح روی تکنیکهای متنوع موجود صورت گیرد .

پردازش دیجیتال سیگنال صحبت در زمینه های فراوانی کاربرد دارد که از آن جمله اند:

۱ـ انتقال و ذخیره سازی سیگنال صحبت بفرم دیجیتال.

۲ـ سنتز سیگنال صحبت [1].

۳ـ تشخیص صحبت [2].

۴ـ تصدیق و تعیین هویت سخنگو [3].

۵ـ افزودن کیفیت سیگنال صوت .

۶ـ کمک در جهت غلبه بر برخی ناتوانیهای جسمی [4].

---

* در مورد مکانیزم تولید سیگنال صحبت و نحوهٔ مدل سازی آن در فصل چهارم به تفصیل صحبت خواهیم کرد .

1. Speech synthesis

2. Speech recognition

3. Speaker verification & Identification

4. Aids-to-the-handicapped

انتقال وذخیره‌سازی سیگنال صحبت به‌فرم دیجیتال :

یکی از قدیمی‌ترین و مهمترین کاربردهای پردازش صحبت درطراحی وکدر[1] (کدکنندهٔ صوت) می‌باشد. هدف این سیستمها کاهش باندفرکانسی موردنیاز برای ارسال سیگنال صحبت می‌باشد. علیرغم افزایش حجم کانالهای مخابراتی دردودههٔ اخیر، که‌به‌جهت امکان استفاده‌ازمایکروویو، ماهواره‌ومخابرات نوری بوجودآمده است، هنوزهم دراکثرمواردنیازبه‌کاهش سرعت ارسال[2] تاحدامکان برای مخابرهٔ اطلاعات صوتی وجوددارد. باکدکردن سیگنال صحبت به‌روش بهینه می‌توان حجم اطلاعات ذخیره‌شدهٔ صوتی را نیزتاحدامکان به‌مقدارقابل توجهی پائین آوردواین مسئله‌فاکتور مهمی درذخیرهٔ سیگنال صحبت نیزمی‌باشد. روشهای کدینگ متنوعی برای دستیابی به این کاهش‌حجم اطلاعات وجوددارندکه‌متدهای PCM[3]، ADM[4]، DPCM[5] و ... اززمرهٔ این متدها می‌باشند.

سنتزسیگنال صحبت :

سیستمهائی که‌قابلیت سنتزسیگنال صحبت را دارا می‌باشنددرواقع توانائی ارتباط لفظی با استفاده‌کننده‌را دارا می‌باشند. دلیل عمدهٔ توجه‌به‌سیستمهای‌سنتزکنندهٔ صوت، ذخیرهٔ مقرون بصرفهٔ اطلاعات درآنها می‌باشد. درسیستمهای اخیر اطلاعاتی که‌بصورت لفظی ازماشین خارج می‌شوند، بشکل متون عادی ذخیره‌شده‌اند. علاوه‌براین، سیستمهای سنتزکنندهٔ صوت نقش مهمی را درفراگیری نحوهٔ تولید صوت توسط انسان بازی می‌کنند.

---

1. Vocoder(Voice coder)
2. Bit rate
3. Pulse coded modulation
4. Adaptive delta modulation
5. Differential pulse coded modulation

تشخیص صحبت :

تشخیص صوت بطورا عم به‌معنا ی تبدیل شکل موج صوت به معا دل دست نویس یا اطلاعات نهفته‌در پیغا م صوتی می‌با شد . کا ربردها ی سیستم تشخیص صوت بسیا رندکه ازآن جمله‌اند : دستگا ه‌تا یپ مبتنی برسیگنا ل صحبت ، ا رتبا ط صوتی با کا مپیوتروسیستمها ی کنترلی‌که دستورا ت خودرا بفرم صوتی دریا فت می‌کنند . ا ین سیستمها با یستی قا بلیت تشخیص کلما ت زیا دی را مستقل ا زصدا ی سخنگودا شته‌با شند . ترکیب یک سیستم تشخیص صوت ویک سیستم تولیدصوت می‌توا ند بوجودآ ورندهٔ یک سیستم مخا براتی با سرعت ا رسا ل بسیا رکم با شد .

تصدیق وتعیین هویت سخنگو :

تکنیکها ی متنوع موجودبرا ی تصدیق سخنگوکا ربردها ی متنوعی درتصدیق فسرددر میا ن ا فرا ددیگردا رد . درواقع سیستم تصدیق سخنگوبا یستی تصمیم بگیردکه‌آ یا سخنگو هما ن شخص موردنظرا ست یا خیر . ا ین سیستمها کا ربردها ئی درکنترل سیستمها ی ا نحصاری وا ستفا ده‌ا زسیستمها ی شخصی دا رد . بطورمثا ل می‌توا ن با نکها ی تلفنی یا قفل‌در صوتی را بعنوا ن نمونه‌ها ئی ا زکا ربردا ین تکنیک معرفی نمود .

سیستمها ی تعیین هویت سخنگوبا یستی با دقت با لائی بتوا نندفردی را ا زتعداد زیا دی ا فرا دمظنون دیگربا توجه‌به‌خصوصیا ت صوتی ا وجدا ساخته‌وبدین ترتیب به‌هویت ا و پی ببرند . یکی ا زکا ربردها ی سیستم تعیین هویت سخنگودرتعیین هویت ا فرا دبا توجه به‌نمونه‌ا ی ا زصحبت شخص می‌با شد که‌درمسا ئل جنا ئی کا رببردا رد . درواقع ا ین تکنیک ا همیتی معا دل ا ثرانگشت ا فرا درا دا رد .

ذکرا ین نکته‌ضروری ا ست که‌با توجه‌به موقعیتها ی متنوع فیزیکی وروا نی که سخنگودرحین تولیدصوت درآ ن قرا ردا رد ، سیستم صوتی ا ومتفا وت عمل خوا هدکردوا ین فاکتورهای مهم‌تأثیر زیا دی بردقت عملکرد سیستمهای معرفی‌شده‌بخصوص‌در مورد دوکا ربرد اخیرمی‌گذارد .