الله
بسم الله الرحمن الرحيم

دانشکده: برق و رباتیک

گروه: الکترونیک

پایان نامه دوره کارشناسی ارشد مهندسی برق- الکترونیک

کاربرد مدل آماری مخلوط لاپلاس در بهسازی گفتار

زینب محمدپوری

اساتید راهنما:

جناب آقای دکتر حسین مروی

جناب آقای دکتر امید رضا معروضی

استاد مشاور:

جناب آقای دکترعلی سلیمانی

زمستان ۱۳۸۹

تقدیم به

پدر و مادرم مهربانم

که هرچه دارم از وجود پرمهر آنهاست

و همسرم

که همواره حامی و مشوق من بودند.

تشکر و قدردانی

خداوند متعال را به جهت الطاف بی پایان و خاصه اش که همـوار شـامل حـال من کرده بسیار شاکرم.

پس از آن بر خود لازم می دانم از زحمات و راهنمایی هـای اسـاتید گرانقـدرم جناب آقای دکتر مروی، دکتر معروضی و دکتر سلیمانی تشکر نمایم.

در ضمن از جناب آقای هادی ویسی به جهت راهنمایی ها و کمک هایشـان و جناب آقای امیرحسین آفریدون هم به جهت زحماتشان برای این پایـان نامـه، کمال تشکر را دارم.

چکیده

بهسازی گفتار یا Speech enhancement بیانگر گروه بزرگی از روش هاست کـه بـا انجـام پـردازش هایی روی سیگنال های نویزی، نهایتا منجر به بهبود کیفیت و قابلیت فهم گفتار می شود.

از مهم ترین روش های بهسازی گفتار، روش های آماری هستند که به لحاظ کارایی بالاتر نسبت به سایر روش ها و ایجاد اعوجاج کمتر در سیگنال نهایی بیشتر مورد توجه می باشد. در این روش یک مدل آماری برای گفتار و نویز فرض می شود و پردازش ها بر مبنای این مدل صورت می گیرد. این مدل آماری برای ضرایب فوریه یا پارامترهای دیگر هر قاب سیگنال گفتار یا نویز فرض می شود.

در این پایان نامه، روش آماری جدیدی برای تخمین سیگنال تمیز از روی سیگنال نویزی در حضور نویز جمع شونده و مستقل از سیگنال تمیز، در حوزه فرکانس ارائه شده است. تا کنون از مدل های گوسی، لاپلاس، گاما و مخلوط گوسی ، برای مدل کردن طیف سیگنال گفتار استفاده شده؛ اما در این پایان نامه، توزیع مخلوط لاپلاس برای مدل کردن بخش های حقیقی و موهومی طیف گفتار پیشنهاد شده است. پارامترهای این مدل (میانگین ها، واریانس ها و ضرایب) به صورت برون خطی و با استفاده از الگوریتم EM و سیگنال صحبت بدون نویز از پایگاه داده TIMIT محاسبه شده است. سپس رابطه تخمین گرMMSE یا کمترین میانگین مربعات خطا، با توجه به توزیع مخلوط لاپلاس به دست آورده شده است. واریانس نویز به صورت درون خطی و با استفاده از روش ردیابی کمینه ها به دست آمده است.

در نهایت برای ارزیابی و عملکرد روش پیشنهادی از سه معیار سیگنال به نویز قطعه‌ای، PESQ و Log Likellihood Ratio (LLR) استفاده شده؛ و روش پیشنهادی با روشهایی که مدل گوسی و لاپلاس را برای سیگنال تمیز فرض کرده اند، مقایسه شده است. بررسی نتایج بیانگر عملکرد قابل قبول روش پیشنهادی است.

کلمات کلیدی:

بهسازی گفتار، توزیع مخلوط لاپلاس، EM الگوریتم، تخمین گر MMSE، آمارگان کمینه

فهرست مقالات پذیرفته شده

۱- The Application Of Laplacian Mixture Model In Speech Enhancement: ICSAP CONFERENCE ۲۰۱۱, Singapore

۲- Speech Spectral Estimation Based on MMSE Estimator and Laplacian Mixture Model: ICECE CONFERENCE ۲۰۱۱, Dubai

۳- Speech Enhancement Using Mixture Of Laplacian Based MMSE Estimator: International Conference on Comunication System, Zahedan, ۲۰۱۰

فهرست مطالب

عنوان صفحه

۱- فصل اول: مقدمه

۱-۱- مقدمه ای بر بهسازی گفتار ۲

۱-۲- کاربردهای سیستم های بهسازی گفتار ۳

۱-۳- تاریخچه روش های بهسازی گفتار ۳

۱-۴- طبقه بندی روش های بهسازی گفتار ۵

۱-۵- اهداف پایان نامه ۶

۱-۶- ساختار پایان نامه ۷

۲- فصل دوم: بهسازی گفتار

۲-۱- مقدمه ۹

۲-۲- بهسازی گفتار ۹

۲-۳- طبقه بندی روش های بهسازی گفتار ۱۱

۲-۴- عوامل موثر در طراحی سیستم های بهسازی گفتار ۱۳

۲-۵- ارزیابی سیستم های گفتار ۱۴

۲-۵-۱- معیار ارزیابی سیگنال به نویز (SNR) ۱۵

۲-۵-۲- معیار ارزیابی PESQ ۱۸

۲-۵-۳- معیار ارزیابی LLR ۱۹

۲-۶- منابع و مشخصات نویزهای صوتی ۲۰

۳- فصل سوم: مروری بر روش های بهسازی گفتار تک کاناله

۳-۱- مقدمه ۲۳

۳-۲- تفریق طیفی ۲۳

۳-۳- زیر فضای سیگنال ۲۶

۳-۴- فیلتر کالمن ۲۷

۳-۵- روش های آماری ۲۹

۴- فصل چهارم: بهسازی گفتار با روش های آماری

۴-۱- مقدمه ۳۲

۴-۲- روش های آماری محض ۳۲

۴-۳- تخمین گر بیشترین شباهت ۳۳

۴-۴- تخمین گرهای بیز ۳۵

۴-۵- تخمین گر MMSE ۳۶

۴-۵-۱- فیلتر وینر ۳۷

۴-۵-۲- تخمین اندازه طیف (STSA-MMSE) ۳۹

۴-۵-۳- تخمین فاز با کمترین میانگین مربعات خطا ۴۲

۴-۵-۴- محاسبه سیگنال به نویز پیشین ۴۳

۴-۵-۴-۱- روش بیشترین شباهت ۴۴

۴-۵-۴-۲- روش تصمیم گرا ۴۵

۴-۵-۵- تخمین MMSE با توزیع های غیر گوسی ۴۶

۴-۶- تخمین گر بیشترین احتمال پسین MAP ۴۸

۴- ۷- به کارگیری عدم قطعیت گفتار (SPU) در بهسازی گفتار ۴۹

۴-۷- ۱- تخمین گر MMSE با SPU ۵۰
۴-۸- روش های تخمین نویز در بهسازی گفتار ۵۰
۴-۸-۱- روش میانگین گیری بازگشتی ۵۱
۴-۸-۲- روش مبتنی بر هیستوگرام ۵۲
۴-۸-۳- روش ردیابی کمینه ها ۵۴
۴-۹- روش های آماری مبتنی بر مدل ۵۷

۵- فصل پنجم: روش پیشنهادی بر مبنای توزیع مخلوطی از لاپلاس ها

۵-۱- مقدمه ۶۱
۵-۲- معرفی پایگاه داده ۶۱
۵-۳- تخمین طیف سیگنال تمیز ۶۳
۵-۳-۱- فرضیات ۶۳
۵-۳-۲- مدل کردن سیگنال تمیز با توزیع مخلوط لاپلاس ۶۴
۵-۳-۲-۱- توزیع مخلوط لاپلاس(LMM) ۶۵
۵-۳-۳- تخمین طیف سیگنال تمیز با تخمین گر MMSE ۶۸
۵-۴- تخمین پارامترهای توزیع سیگنال تمیز با EM الگوریتم ۷۱
۵-۴-۱- بیشترین شباهت ۷۲
۵-۴-۲- EM الگوریتم ۷۳
۵-۴-۳- محاسبه پارامترهای توابع توزیع مخلوط ۷۵
۵-۴-۴- محاسبه پارامترهای LMM ۷۸
۵-۴-۵- پیاده سازی EM الگوریتم ۷۹
۵-۵- تخمین نویز ۸۶
۵-۶- ارزیابی روش پیشنهادی ۸۷

۵-۶-۱- معیار ارزیابی سیگنال به نویز قطعه ای ۸۹
۵-۶-۲- معیار ارزیابی LLR ۹۳
۵-۶-۳- معیار ارزیابی PESQ ۹۶
۵-۷- مقایسه روش پیشنهادی با روش های دیگر ۹۹
۵-۷-۱- معیار سیگنال به نویز قطعه ای ۱۰۰
۵-۷-۲- معیار LLR ۱۰۲
۵-۷-۳- معیار PESQ ۱۰۴

۶- فصل ششم: نتیجه گیری و پیشنهادات

۶-۱- مقدمه ۱۰۸
۶-۲- نتیجه گیری و جمع بندی ۱۰۸
۶-۳- پیشنهاد برای کارهای آینده ۱۱۱

فهرست منابع ۱۱۲

فهرست شکل ها

عنوان صفحه

شکل (۳-۱): نمای کلی الگوریتم تفریق طیفی . ۲۵

شکل (۳-۲) نمای کلی روش های آماری. ۳۰

شکل (۴-۱): پریودوگرام $|Y(\lambda,k)|^2$، پریودوگرام هموار شده $P(\lambda,k)$ و تخمین نویز برای فرکانس $k=25$. ۵۶

شکل(۵-۱): بلوک دیاگرام روش پیشنهادی. ۶۲

شکل (۵-۲): توزیع لاپلاس به ازای واریانس و میانگین های متفاوت. ۶۵

شکل (۵-۳): هیستوگرام بخش حقیقی تبدیل فوریه زمان کوتا سیگنال تمیز. ۶۶

شکل (۵-۴): هیستوگرام بخش موهومی تبدیل فوریه زمان کوتا سیگنال تمیز. ۶۷

شکل (۵-۵): نقطه چین:تخمین توزیع بخش حقیقی تبدیل فوریه زمان کوتاه سیگنال تمیز با یک لاپلاس، خط یکپارچه: تخمین توزیع بخش حقیقی تبدیل فوریه زمان کوتاه سیگنال تمیز با مخلوطی از شش لاپلاس، خط چین: تخمین توزیع بخش حقیقی تبدیل فوریه زمان کوتاه سیگنال تمیز با شش گوسی. ۶۸

شکل (۵-۶): تخمین EM الگوریتم برای بخش حقیقی طیف سیگنال تمیز، (الف) پس از ۱۰ بار تکرار، (ب) پس از ۳۰ بار تکرار، (ج) پس از ۴۰ بار تکرار، (چ) پس از ۵۰ بار تکرار. ۸۴

شکل (۵-۷): تخمین EM الگوریتم برای بخش حقیقی طیف سیگنال تمیز : (الف) به ازای یک لاپلاس، (ب) به ازای ده لاپلاس. ۸۵

شکل (۵-۸): تخمین EM الگوریتم برای بخش موهومی طیف سیگنال تمیز:
(الف) به ازای یک لاپلاس، (ب) به ازای ده لاپلاس. ۸۶
شکل (۵-۹): خط چین: پریودوگرام سیگنال نویزی که با نویز F ۱۶ و
در سیگنال به نویز ۵ دسی بل, آغشته شده است. خط یکپارچه نازک:
پریودوگرام هموارشده و خط یکپارچه کلفت: تخمین نویز. ۸۸
شکل (۵-۱۰): نمودار سیگنال به نویز قطعه ای سیگنال بهسازی شده،
پس از اعمال روش پیشنهادی روی سیگنال آغشته به نویز سفید. ۹۱
شکل (۵-۱۱): نمودار سیگنال به نویز قطعه ای سیگنال بهسازی شده،
پس از اعمال روش پیشنهادی روی سیگنال آغشته به نویز F۱۶. ۹۱
شکل (۵-۱۲): نمودار سیگنال به نویز قطعه ای سیگنال بهسازی شده،
پس از اعمال روش پیشنهادی روی سیگنال آغشته به نویزهمهمه. ۹۲
شکل (۵-۱۳): نمودار LLR برای سیگنال بهسازی شده، پس از اعمال
روش پیشنهادی روی سیگنال آغشته به نویز سفید. ۹۴
شکل (۵-۱۴): نمودار LLR برای سیگنال بهسازی شده، پس از اعمال
روش پیشنهادی روی سیگنال آغشته به نویز F۱۶. ۹۵
شکل (۵-۱۵): نمودار LLR برای سیگنال بهسازی شده، پس از اعمال
روش پیشنهادی روی سیگنال آغشته به نویز همهمه. ۹۵
شکل (۵-۱۶): نمودار PESQ برای سیگنال بهسازی شده، پس از اعمال
روش پیشنهادی روی سیگنال آغشته به نویز سفید. ۹۷
شکل (۵-۱۷): نمودار PESQ برای سیگنال بهسازی شده، پس از اعمال
روش پیشنهادی روی سیگنال آغشته به نویز F۱۶. ۹۷
شکل (۵-۱۸): نمودار PESQ برای سیگنال بهسازی شده، پس از اعمال

روش پیشنهادی روی سیگنال آغشته به نویز همهمه. ۹۸
شکل(۵-۱۹): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب سیگنال
به نویز قطعه ای و برای نویز F۱۶. ۱۰۱
شکل(۵-۲۰): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب سیگنال
به نویز قطعه ای و برای نویز همهمه. ۱۰۱
شکل(۵-۲۱): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب
LLR و برای نویز F۱۶. ۱۰۲
شکل(۵-۲۲): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب
LLR و برای نویز همهمه. ۱۰۳
شکل(۵-۲۳): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب
PESQ و برای نویز F۱۶. ۱۰۴
شکل(۵-۲۴): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب
PESQ و برای نویز همهمه. ۱۰۵

فهرست جدول ها

عنوان صفحه

جدول(۲-۱) : مهم ترین روشهای ارزیابی سیستم های بهسازی گفتار. ۱۶

جدول(۳-۱): مروری برروش تفریق طیفی. ۲۶

جدول(۳-۲): مروری برروش زیر فضای سیگنال. ۲۷

جدول(۳-۳): خلاصه روش های آماری. ۳۰

جدول (۵-۱): پارامترهای LMM و I_{KL}، برای بخش حقیقی تبدیل
فوریه زمان کوتاه داده های TIMIT به ازای N=۳ و بعد از ۱۰ ،۳۰
،۴۰ و ۵۰ تکرار EM الگوریتم. ۸۱

جدول (۵-۲) : مقایسه سیگنال به نویزهای قطعه ای برای سه نویز همهمه،
سفید و F۱۶ در سه سیگنال به نویز ۰، ۵، ۱۰ و N های متفاوت. ۹۲

جدول(۵-۳) : مقایسه LLR برای سه نویز همهمه، سفید و F۱۶ در سه سیگنال
به نویز ۰، ۵، ۱۰ و N های متفاوت. ۹۶

جدول(۵-۴): مقایسه PESQ برای سه نویز همهمه، سفید و F۱۶ در سه سیگنال
به نویز ۰، ۵، ۱۰ و N های متفاوت. ۹۸

جدول(۵-۵): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب سیگنال
به نویز قطعه ای و برای نویز همهمه و F۱۶. ۱۰۲

جدول(۵-۶): مقایسه روش پیشنهادی با روش های
Lap-MMSE,Log-MMSE,MMSE, بر حسب

LLR و برای نویز همهمه و F۱۶. ۱۰۳

جدول(۵-۷): مقایسه روش پیشنهادی با روش های

Lap-MMSE,Log-MMSE,MMSE, بر حسب

PESQ و برای نویز همهمه و F۱۶. ۱۰۵

فصل اول

مقدمه

۱-۱- مقدمه ای بر بهسازی گفتار

همان طور که می دانیم مهمترین وسیله ارتباطی افراد از طریق صدا و گفتار می‌باشد. به همین جهت پردازش گفتار، به جهت کاربردهای فراوان در زندگی بشر، از سالیان قبل مد نظر دانشمندان و محققین بوده است. یکی از اساسی ترین بخش های پردازش گفتار، بهسازی گفتار و تلاش برای بهبود کیفیت عملکرد سیستم های ارتباط گفتاری می‌باشد.

شاید تا به حال برای شما پیش آمده باشد که از یک نوار کاست قدیمی که حاوی سخنرانی یا اطلاعات مورد نظر شماست، استفاده کرده باشید و احتمالا به علت گذشت زمان یا مشکلات زمان ضبط، صداهای روی آن به سختی شنیده شود. قطعا شما دوست دارید که به طریقی کیفیت صداهای ضبط شده روی نوار کاست را بالا ببرید. یا حالتی را در نظر بگیرید که در داخل خودرو با تلفن همراه صحبت می کنید. قطعا صداهای مزاحم نیز مانند صدای باد، موتور ماشین و سرنشینان همراه با صدای شما وارد میکروفون می‌شوند.

این موارد تنها مثال های ساده ای از کاربردهای فراوان روش های کاهش یا حذف اثرات نویز بر سیگنال صوت است که به روش های بهسازی گفتار معروف هستند.

در قالب یک تعریف کلی، بهسازی گفتار عبارت است از تلاش برای بهبود کیفیت عملکرد سیستم های ارتباط گفتاری در مواردی که سیگنال گفتار تحت تاثیر نویز، انعکاس[۱] و سایر عوامل تخریبی واقع گردیده است.

نیاز به بهسازی گفتار از آنجا ناشی می شود که سیگنال گفتار معمولا، یا از منبعی تولید می‌شود که در محیط نویزی قرار دارد، یا تحت تاثیر کانال انتشار و در اثر نویز یا انعکاس دچار تخریب می‌شود، یا در گیرنده به نویز آلوده می‌شود و یا ممکن است منشا نویز ترکیبی از سه حالت فوق باشد [۶۶].

---

۱ - Echo

۱-۲- کاربردهای سیستم های بهسازی گفتار

سیستم های بهسازی گفتار کاربردهای وسیع و متفاوتی در حوزه های مختلف دارند که به طور کلی می‌توان آنها را به صورت زیر طبقه بندی کرد:

- تشخیص خودکار گفتار[۱] و تشخیص هویت گوینده[۲] : در سیستم های تشخیص خودکار گفتار و هویت گوینده ، عملکرد سیستم در محیط هایی با سیگنال به نویز پایین مناسب نمی‌باشد. ثابت شده که نرخ تشخیص با به کارگیری الگوریتم های بهسازی گفتار بهبود می یابد [۵۲].
- سیستم های تلفن: که در آنها گفتار اصلی به وسیله نویز زمینه یا نویز موجود در مسیر مخابره و نیز در اثر انعکاس صدای طرفین مکالمه خراب می‌شود.
- تلفن های عمومی: که در محیط های پر سر و صدا و شلوغ واقع اند.
- سیستم های مخابرات هوا به زمین: که در آنها نویز اتاق خلبان، پیام ارسال شده از سوی خلبان را خراب می کند.
- سمعک: که به عنوان یک تقویت کننده، هم سیگنال گفتار و هم نویز موجود در محیط را تقویت نموده و موجب ناراحتی کاربر می‌شود.
- حذف نویز در گفتارهای ضبط شده: که در آنها صدای ضبط شده اشخاص آغشته به نویز است و جهت بهبود کیفیت آنها، باید نویز را حذف کرد.
- کدینگ اطلاعات: که اثر نویز بر گفتار به شدت عملکرد آنها را تحت تاثیر قرار می‌دهد[۶۶].

۱-۳- تاریخچه روش های بهسازی گفتار

با توجه به کاربرد های مهم و فراوانی که برای روش های بهسازی گفتار بیان شد، و اهمیت آن، روش های مختلفی برای این کار پیشنهاد شده اند.

---

۱ - Automatic Speech Recognition
۲ - Speaker Identification

از مهمترین و قدیمی ترین روش های بهسازی گفتار می توان به فیلتر وینر[۱] [۶۰] و مشتقات آن اشاره کرد. در این روش ها، که برای اولین بار در سال ۱۹۴۹ وارد حوزه بهسازی گفتار گردید، سعی می‌شود که تخمین سیگنال تمیز بر اساس یک معیار بهینه سازی باشد. یکی دیگر از روش های قدیمی و بسیار معروف بهسازی گفتار، تفریق طیفی [۲] [۴] و [۶]، است که در سال ۱۹۷۵ پایه گذاری شد و در سال ۱۹۷۹ با پردازش در حوزه فرکانس مطرح شد. پس از آن بهبودهای زیادی بر روی این روش صورت گرفت. فیلتر کالمن[۳] هم در سال ۱۹۸۷با مقاله Paliwal & Basu [۴۷] وارد این حوزه گردید.

روش های آماری محض[۴] به عنوان یکی از مهم ترین روش های بهسازی گفتار، با کارهای Lim & Oppenheim [۳۵] و McAullay & Malpass [۴۵] و مقاله معروف Malah & Ephraim [۲۲] پیشنهاد شدند و بسیار مورد توجه محققین قرار گرفتند. روش های آماری مبتنی بر مدل[۵] نیز با کارهای Ephraim در [۲۴] و [۱۸] مطرح شدند و امروزه به عنوان یکی مهم ترین زمینه های علمی و کاربردی در بهسازی گفتار مورد توجه می‌باشند. می توان ادعا کرد که در کل روش های آماری چه از لحاظ کیفی و چه از لحاظ کمی نتایج بهتری نسبت به اغلب روشهای دیگر ارائه می‌دهند.

روش مهم دیگر بهسازی گفتار که مبتنی بر تئوری جبر خطی می‌باشد، زیر فضای سیگنال[۶] است که در ابتدا در [۱۴] و [۱۶] و [۲۵] و در سال ۱۹۹۱ارائه گردید و سپس بهبودهای زیادی روی آن صورت پذیرفت. اخیرا روش های مبتنی بر کپسترال[۷]، تبدیل موجک[۸] و تبدیل فوریه کسینوسی[۹] نیز نیز وارد حوزه بهسازی گفتار شده اند.

علاوه بر این روش ها، روش های دیگری نیز برای بهسازی گفتار پیشنهاد شده اند که در آنها از دو یا

---

۱ -Wiener Filter
۲ - Spectral Subtraction
۳ - Kalman Filter
۴ - Pure Statistical Method
۵ - Statistical Model-Based
۶ - Signal Subspace
۷ - Cepstral
۸ - Wavelet
۹ -Discerete Cosine Transform

چند میکروفن برای ضبط سیگنال نویزی استفاده می‌شود. از مهم ترین این روش ها می توان به روش فیلتر وفقی دوکاناله که در سال ۲۰۰۳ مطرح شد، اشاره کرد.

تا امروز مقالات و روش های بسیاری در زمینه بهسازی گفتار ارائه شده است، اما تخمین سیگنال تمیز از روی سیگنال نویزی هنوز یک مساله حل نشده در پردازش گفتار است و چالش های متعددی را در خود نگه داشته است. یکی از دلایل ساده برای اثبات این ادعا این است که علی رغم قدمت بیشتر بهسازی گفتار نسبت به برخی کاربردهای دیگر پردازش گفتار مثل بازشناسی گفتار، هنوز روش قطعی و غالبی در این زمینه وجود ندارد. از دلایل اصلی مشکل بودن حل این مساله، نامشخص بودن جواب دقیق برخی از سوالات کلیدی و ابهامات در این حوزه است. از جمله این سوالات عبارتند از:

- آیا می‌توان یک رابطه مشخص را برای معیارهای کیفیت و قابلیت فهم، به صورتی که انسان درک می کند، به دست آورد؟ پاسخ به این سوال، به نوبه خود منجر به یافتن معیار بهینه سازی معنی داری می شود که با ادراک انسان متناسب است. با داشتن روابط مشخص برای این دو، می توان یکی از آنها را بدون داشتن اثر منفی بر دیگری بهبود بخشید و می توان به این سوال نیز پاسخ داد که چه روش هایی برای بهسازی گفتار در کاربرد کمک شنیداری مناسب است؟ کدام یک برای کاربرد بازشناسی گفتار و کدام یک برای کدینگ؟
- چه پردازش هایی در سطوح مختلف سیستم شنوایی انسان انجام می‌گیرد و چگونه می‌توان آنها را مدل کرد؟ پاسخ این سوال نیز منجر به روشن شدن شبهات دیگری مانند نحوه ترکیب تکنیک های پردازش سیگنال با روش های ادراکی انسان می‌شود.

۱-۴- طبقه بندی روش های بهسازی گفتار

با توجه به تعدد روش های بهسازی گفتار، آنها را می‌توان به طرق مختلفی طبقه بندی کرد [۲۸] که در این فصل فقط آنها را نام می بریم. این طبقه بندی در فصل بعد توضیح داده خواهد شد.

۱- طبقه بندی بر اساس تعداد کانال ورودی