نوزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران
دانشکده مهندسی کامپیوتر، دانشگاه شهید بهشتی، تهران، ۱۵-۱۳ اسفندماه ۱۳۹۲

کاهش نویز در گفتار با استفاده از کدگذاری تنک و تابع انقباض

علیرضا شیرخدائی[۱]، محمدمهدی همایون‌پور[۲]

[۱] کارشناس ارشد، مهندسی کامپیوترگرایش هوش مصنوعی، دانشگاه صنعتی امیرکبیر، تهران
alireza.shirkhodaei@gmail.com

[۲] دانشیار، دانشکده مهندسی کامپیوتر و فنآوری اطلاعات، دانشگاه صنعتی امیرکبیر ، تهران
homayoun@aut.ac.ir

چکیده

در این مقاله به روشی برای حذف و کاهش نویز از گفتار می‌پردازیم که بر کدگذاری تنک و اصول انقباضی بنا شده است. کل عملیات در حوزه زمان انجام می‌گیرد. از مهمترین کاربردهای این روش می‌توان به استفاده در سیستم‌های تشخیص گفتار و نیز در ساخت سمعک برای کم‌شنواها، اشاره نمود که کاهش نویز در این سیستم‌ها حیاتی هستند. نکته‌ی مهم در روش کدگذاری تنک چگونگی تعیین پارامترهای لازم برای تابع انقباض و تخمین واریانس نویز است. در این روش تخمین واریانس نویز با روشی ابتکاری در زمان آزمایش انجام می‌پذیرد، و تعیین سایر پارامترهای لازم در زمان آموزش از روی دادگان تمیز به خوبی انجام می‌پذیرد و سبب می‌گردد که نه‌تنها نویز بسیار خوب کاهش یابد، بلکه ساختار کلیدی گفتار نیز حفظ گردد. کارایی روش فوق با الگوریتم‌های شناخته شده‌ای چون فیلترگذاری وینر و تفریق طیفی با سه مقیاس نسبت سیگنال به نویز قطعه‌ای، ارزیابی ادراکی کیفیت گفتار و میانگین امتیاز نظرات ارزیابی و مقایسه می‌گردد و نتیجه گرفته می‌شود که برتری قابل توجهی نسبت به سایر روش ها ارائه داده است. به‌عنوان نمونه در مقیاس نسبت سیگنال به نویزقطعه‌ای، کدگذاری تنک انقباضی در نسبت سیگنال به نویز ۱۰ دسی‌بل ۶/۲ بوده درحالی‌که فیلترگذاری‌وینر ۲/۸و تفریق طیفی نتیجه بهتری از ۲/۲ ارائه نداده است و در شدت‌های دیگر نویز نیز به همین منوال می‌باشد.

کلمات کلیدی

کاهش نویز، کدگذاری تنک، تابع انقباض، بهسازی خودکار گفتار

۱- مقدمه

امواج صوتی در کاربردهای واقعی، اغلب توسط کانال‌های مخابراتی یا نویزهای پس‌زمینه‌ی جمع‌شونده، خراب می‌شوند. هدف بهسازی خودکار، تخمین سیگنال گفتار تمیز از سیگنال مشاهده شده‌ی نویزی است. تفریق طیفی[۱] [1] و فیلترگذاری وینر[۲] [2,3] جزو اولین فن‌آوری‌های بهسازی گفتار محسوب می‌گردند و هنوز به‌دلیل کارایی قابل توجه و پیچیدگی کم محاسباتی، پرمصرف هستند. تفریق طیفی با کسر تخمین بزرگی طیف نویز از بزرگی طیف گفتار نویزی به‌دست می‌آید. الگوریتم فیلترگذاری وینر سیگنال گفتار نویزی را توسط اعمال فیلتری که مبتنی بر معیار کمینه نمودن مربع میانگین خطا است، بهسازی می‌کند.

برخی از فن‌آوری‌های بهسازی تلاش می‌کنند تا خصوصیات پوششی سیستم شنوایی انسان را به‌دست آورند زیرا این سیستم سبب می‌شود تا نویز جمع‌شونده، با بزرگی زیر آستانه‌ی پوششی، توسط گوش انسان قابل شنیدن نباشد. همین مساله در چهارچوب تفریق طیفی تعمیم‌یافته توسط ویراگ[۳] [4] در سال ۱۹۹۹ و در الگوریتم بهسازی گفتار مرتبه‌ی بتای کمینه‌ی، کمینه سازی مربع میانگین خطا[۴] توسط یو و همکاران[۵] [5] در سال ۲۰۰۶ مورد بررسی قرار گرفته است. با استفاده از وزن‌دهی ادراکی ارائه شده توسط هو و لوویزو[۶] [6] در سال ۲۰۰۳ نسخه‌ی تصحیح‌شده‌ای از فیلترگذاری وینر عرضه شد.