

کاربرد الگوریتم تفریق طیفی ارتقایافته به منظور کاهش نویز به همراه بررسی تعداد کانال‌ها در پردازشگر سیگنال گفتاری سیستم کاشت حلزون

مرجان مظفری لقا^۱، دکتر سعید کرمانی^۲

مقاله پژوهشی

چکیده

مقدمه: کاشت حلزون شنوایی به عنوان یکی از مؤثرترین روش‌ها برای بازگرداندن درجه‌ای از شنوایی افراد مبتلا به کاهش شنوایی عمیق شناخته شده است. استراتژی‌های بهبود گفتار نقش بسیار مهمی را در بهینه‌سازی کاشت حلزون ایفا می‌کنند. در این مطالعه، یک الگوریتم کاهش نویز برای کاشت حلزون ارائه شده است.

روش‌ها: به منظور بهبود عملکرد کاشت حلزون در محیط‌های نویزی، یک الگوریتم کاهش نویز برای این سیستم ارائه شده است که شامل تفریق طیفی با استفاده از طبقه‌بندی بین غالب بودن گفتار و یا نویز در هر کانال می‌باشد. در طبقه‌بندی پیشنهادی از انحراف معیار طیف سیگنال مشاهده‌شده در هر کانال استفاده می‌گردد. عملکرد الگوریتم کاهش نویز پیشنهادی از طریق ارزیابی عینی توسط نسبت سیگنال به نویز (Signal-to-noise ratio یا SNR) قطعه‌ای با استفاده از جملات دادگان Noisy92 که توسط نویز همهمه، اتومبیل و قطار در SNR ۰ تا ۲۰ دسی‌بل مخدوش گردیده بودند و مشاهده‌ی طیف سیگنال و همچنین آزمون ذهنی توسط ۱۵ شنونده با شنوایی طبیعی با استفاده از شبیه‌ساز کاشت حلزون در کلینیک شنوایی دانشگاه علوم پزشکی اصفهان در مهر ماه سال ۱۳۹۱ انجام گرفت. به منظور بررسی تأثیر کانال‌ها بر روی درصد بازشناسی گفتار، از فیلتر بانک‌های ۶، ۸ و ۱۶ کاناله استفاده گردید.

یافته‌ها: در تمام SNRها، توسط الگوریتم پیشنهادی افزایش درکی در حدود ۴۵ درصد مشاهده شد.

نتیجه‌گیری: بر اساس ارزیابی‌های انجام‌شده، روش پیشنهادی موجب کاهش نویز سیگنال گفتاری می‌گردد.

واژگان کلیدی: کاشت حلزون، تفریق طیفی ارتقایافته، کاهش نویز

ارجاع: مظفری لقا مرجان، کرمانی سعید. کاربرد الگوریتم تفریق طیفی ارتقایافته به منظور کاهش نویز به همراه بررسی تعداد کانال‌ها در پردازشگر سیگنال گفتاری سیستم کاشت حلزون. مجله دانشکده پزشکی اصفهان ۱۳۹۲؛ ۳۱ (۲۵۷): ۱۶۷۴-۱۶۶۷

مقدمه

اعصاب شنوایی می‌باشد (۱). روش تبدیل سیگنال گفتاری ورودی به پالس‌های الکتریکی را پردازش گفتار و یا استراتژی کدگذاری می‌نامند (۲). استراتژی پردازش گفتار CI، عملکرد سیستم شنوایی محیطی را تقلید می‌کند (۳). اغلب دستگاه‌های مدرن CI برای

سیستم کاشت حلزون (Cochlear implant یا CI) یکی از پروتزهای شنوایی مؤثر برای بازگرداندن عملکرد شنوایی بیماران مبتلا به کاهش شنوایی حسی - عصبی شدید تا عمیق با استفاده از تحریک

* این مقاله حاصل پایان‌نامه‌ی دوره‌ی کارشناسی ارشد به شماره‌ی ۳۹۱۱۵۹ در دانشگاه علوم پزشکی اصفهان است.

۱- دانشجوی کارشناسی ارشد، گروه فیزیک و مهندسی پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۲- استادیار، گروه فیزیک و مهندسی پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

Email: kermani@med.mui.ac.ir

نویسنده‌ی مسؤول: دکتر سعید کرمانی

در سیستم‌های پردازش سیگنال امروزی، می‌باشیم. Loizou به بررسی مزایای استفاده از یک الگوریتم کاهش نویز به منظور بهبود CI در محیط نویزی پرداخت، که در آن از یک الگوریتم زیر بانندی استفاده کرد (۸).

در این مطالعه، برای بهبود عملکرد پردازشگر سیگنال گفتاری در محیط نویزی، یک الگوریتم تفریق طیفی بهبودیافته پیشنهاد گردیده است که در مقایسه با استراتژی نمونه‌برداری بین پالسی پیوسته (CIS یا Continuous interleaved sampling) موجود، موجب بهبود بیشتری در عملکرد سیستم CI می‌گردد.

روش‌ها

هدف اصلی این مقاله ارائه‌ی یک استراتژی جدید در پردازشگر سیگنال گفتاری سیستم CI بر اساس طبقه‌بندی بین گفتار و نویز غالب در هر کانال که شامل چندین باند فرکانسی است، می‌باشد (۹). مطابق با این طبقه‌بندی، تخمین طیف نویز بر اساس نتایج تصمیم‌گیری و ویژگی‌های ماسکینگ سیستم شنوایی انجام می‌گیرد (۱۰). سپس نویز از طریق روش تفریق طیفی حذف می‌گردد. به منظور این طبقه‌بندی از انحراف معیار (Standard deviation یا SD) طیف سیگنال مشاهده‌شده در هر کانال استفاده گردید. در صورت وجود نویز و سیگنال در یک سیگنال، SD بالا می‌باشد. به عبارت دیگر، اگر تنها نویز در یک کانال موجود باشد SD پایین است (۱۱). بنابراین، طبقه‌بندی سیگنال و نویز با استفاده از یک آستانه‌ی مناسب امکان‌پذیر است. در اینجا دو فاکتور برای سیگنال‌های نویزی و گفتار ارائه شده است که در ادامه مورد بررسی دقیق، قرار می‌گیرد.

تجزیه‌ی فرکانسی سیگنال ورودی، از فیلتر بانک استفاده می‌کنند که در واقع شبیه‌سازی عملکرد تجزیه‌ی فرکانسی توسط حلزون شنوایی بر اساس تئوری جا به جایی می‌باشد. خروجی هر یک از کانال‌های فیلتر بانک برای مدولاسیون دامنه‌ی پالس‌های الکتریکی تحریکی استفاده می‌گردد. در وسایل چند کاناله‌ی موجود، از فیلترهای میان‌گذر ساده استفاده می‌شود (۴). در حال حاضر، عملکرد CI در محیط‌های آرام جهت یک مکالمه‌ی موفق، بهبود یافته است. با این وجود زمانی که سطح نویز محیط افزایش می‌یابد، عملکرد CI به طور قابل توجهی کاهش می‌یابد.

چندین روش به منظور بهبود عملکرد CI تحت شرایط نویزی ارائه شده است (۵). الگوریتم‌های کاهش نویز تک کاناله، یکی از این روش‌ها است. اگر چه تعداد کمی از این استراتژی‌ها برای الگوریتم کاهش نویز در سیستم CI ارائه شده‌اند، با این وجود، این استراتژی‌ها در سیستم‌های CI قدیمی که بر اساس استراتژی استخراج ویژگی بودند، استفاده می‌گردید. Weiss یک روش پیش پردازش سیگنال به همراه یک الگوریتم کاهش نویز استاندارد را ارائه داد که موجب کاهش خطای استخراج فرمت‌ها می‌گردید (۶).

پردازشگرهای سیگنال گفتار امروزی بر اساس استراتژی‌های استخراج ویژگی نمی‌باشند، بلکه تنها سیگنال ورودی از یک فیلتر بانک عبور می‌کند (n کاناله و $n > 22$) و سپس پوش سیگنال در هر کانال گرفته می‌شود (۷). در حقیقت در برخی موارد این پیش پردازش‌ها ممکن است موجب دیستورشن (Distortion) شوند. در حالت ایده‌آل، ما به دنبال یک الگوریتم کاهش نویز حتی‌الامکان ساده و قابل استفاده

روش ارائه شده

این مطالعه به بررسی استفاده از یک روش جدید به منظور بهبود سیگنال در سیستم CI پرداخت. بهبود سیگنال گفتاری، در هر باند زمانی انجام گرفت. در CI، سیگنال شنیداری توسط میکروفن دریافت گردید و سپس از یک فیلتر بالاگذر به منظور تأکید بر روی فرکانس‌های بالا عبور کرد. پس از آن سیگنال از یک فیلتر بانک که به منظور شبیه‌سازی غشای بازیلار حلزون گوش استفاده گردید، عبور کرد. خروجی کانال‌ها از الگوریتم تفریق طیفی بهبودیافته عبور نمودند. به منظور تخمین نویز از روش کوانتایل استفاده گردید، در نهایت پوش سیگنال اکوستیکی استخراج گردید و از یک فیلتر پایین گذر عبور کرد. سیگنال $s(k)$ توسط نویز $n(k)$ مخدوش شد. سیگنال مشاهده شده به صورت معادله‌ی ۱ بود:

$$y(k) = s(k) + n(k) \quad ۱.$$

با گرفتن تبدیل فوریه

$$Y(w, r) = S(w, r) + N(w, r) \quad ۲.$$

اگر تخمین نویز را $|N^{\wedge}(w, r)|$ در نظر بگیریم و سیگنال گفتاری تخمینی را $|S^{\wedge}(w, r)|$ در نظر بگیریم، نتیجه به صورت زیر می‌شود:

$$|\hat{S}(w, r)| = H(w, r) |Y(w, r)| \quad ۳.$$

معادله‌ی ۴. H

$$H(w, r) = \begin{cases} \sqrt{1 - \alpha \text{SNR}_{\text{post}}(w, r)^2} & \text{if } J \geq 0 \\ \beta \sqrt{\text{SNR}_{\text{post}}(w, r)^2} & \text{otherwise} \end{cases}$$

$$J = \frac{1}{(\alpha + \beta)} - \text{SNR}_{\text{post}}(w, r)^2 \quad ۵.$$

$$\text{SNR}_{\text{post}}(w, r) = \frac{|N(w, r)|}{|Y(w, r)|} \quad ۶.$$

که در آن $H(w, r)$ تابع بهره‌ی $(\alpha \leq 1)$ فاکتور تفریقی و $(\beta \leq 0)$ فاکتور سطح پایه است. وقتی J بزرگتر از صفر است تفریق طیفی انجام می‌گردد. از

سوی دیگر، زمانی که $J > 0$ فلور طیفی انجام می‌گردد.

سیگنال مشاهده شده $y(k)$ توسط پنجره همینگ پنجره‌بندی گردید و سپس این سیگنال توسط تبدیل فوریه به حوزه‌ی فرکانس منتقل گردید. طبقه‌بندی سیگنال و نویز با استفاده از طیف سیگنال مشاهده شده انجام گرفت. در بخش بعد این طبقه‌بندی با جزئیات ارائه شده است. تخمین طیف نویز با استفاده از روش کوانتایل انجام گرفت. پس از آن طبقه‌بندی تفریق طیفی با استفاده از معادله‌های ۳ و ۴ انجام گرفت. دو فاکتور تفریق طیفی به منظور کاهش دیستورشن انتخاب گردید. در نهایت با افزودن فاز به سیگنال $\hat{s}(k)$ و گرفتن تبدیل فوریه‌ی معکوس سیگنال بهبودیافته به دست آمد.

$$\hat{S}(k) = \text{IFFT} \left[|\hat{S}(w, r)| \cdot e^{j \angle S(w, r)} \right] \quad ۷.$$

طبقه‌بندی غالب بودن گفتار و یا نویز

در این مقاله یک استراتژی جدید در پردازشگر سیگنال گفتاری سیستم CI برای طبقه‌بندی گفتار و نویز در هر کانال که شامل چندین باند فرکانسی است، ارائه شد. به منظور این طبقه‌بندی از SD طیف سیگنال مشاهده شده در هر کانال استفاده گردید. برای محاسبه‌ی انحراف معیار طیف سیگنال مشاهده شده از معادلات زیر استفاده کردیم:

$$Y_{w, r} - |Y(l, r)|^2$$

$$|Y(l, r)| = \frac{1}{B_i} \cdot \sum_{w \in C_{B_i}} |Y(w, r)| \quad ۹.$$

که در آن i تعداد باندهای بحرانی، C_{B_i} ست باندهای فرکانسی متعلق به باند بحرانی i ام و B_i تعداد باندهای فرکانسی در هر باند بحرانی C_{B_i} است. طبقه‌بندی سیگنال و نویز با استفاده از یک آستانه‌ی مناسب به صورت زیر امکان‌پذیر است:

معادله‌ی ۱۰.

$$\begin{cases} \text{speech} - \text{dominant} & \text{if } SD(i, r) > Th \\ \text{noise} - \text{dominant} & \text{otherwise} \end{cases}$$

یافته‌ها

در این بخش به ارزیابی الگوریتم ارائه‌شده می‌پردازیم. پارامترهای انتخاب‌شده برای این مقاله به صورت زیر است:

(۱) طول پنجره‌ی همینگ (32ms) $N = 512$ و 50

درصد هم‌پوشانی

(۲) $\alpha_1 = 2.5$, $\alpha_2 = 6$, $\beta = 0.001$, $Th = 5$

$q = 0.7$

به منظور بررسی تأثیر کانال‌ها بر روی درصد

بازشناسی گفتار، از فیلتر بانک‌های ۶، ۸ و ۱۶ کاناله استفاده گردید.

ارزیابی عینی

عملکرد الگوریتم کاهش نویز پیشنهادشده از طریق ارزیابی عینی توسط SNR قطعه‌ای با استفاده از جملات Noisy92 که توسط نویز همهمه، اتومبیل و قطار در SNR ۰ تا ۲۰ دسی‌بل مخدوش گردیده بودند، انجام شد.

به منظور محاسبه‌ی SNR قطعه‌ای داریم:

معادله‌ی ۱۱.

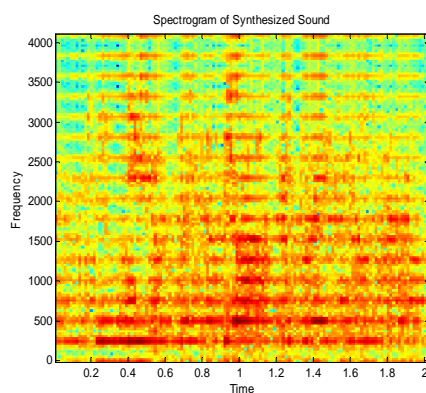
$$\text{seg. SNR} = \frac{10}{M} \sum_{m=0}^{M-1} \log_{10} \left(\frac{\sum_{n=N_m}^{M+N-1} s^2(n)}{\sum_{k=N_m}^{M+N-1} (s(n) - \hat{s}(n))^2} \right)$$

جدول ۱. بررسی SNR قطعه‌ای بر روی نویز همهمه، اتومبیل و قطار

۲۰	۱۵	۱۰	۵	۰	
۱۳/۳	۹/۷	۴/۶۷	۲/۱۱	۱/۱	بدون الگوریتم پیشنهادی
۲۴/۶	۲۱/۸	۱۴/۵۵	۱۰	۳/۳	الگوریتم پیشنهادی با نویز همهمه
۱۹/۵	۱۴/۰۳	۶/۹۰	۲/۲	۱/۱	الگوریتم پیشنهادی با نویز ماشین
۲۲/۸	۲۰/۶	۱۲/۲۴	۹/۸	۳/۱	الگوریتم پیشنهادی با نویز قطار

در جدول ۱، بهبود SNR قطعه‌ای بر روی نویز همهمه، اتومبیل و قطار نشان داده شده است. از جدول ۱، در می‌یابیم که روش پیشنهادی برای پردازشگر گفتار CI مناسب بود.

همچنین برای مقایسه‌ی بهتر، اسپکتروگرام سیگنال سنتز شده، بدون و با استراتژی پیشنهادی حذف نویز ترسیم شده است. شکل ۱ اسپکتروگرام سیگنال سنتز شده برای نویز همهمه را نشان می‌دهد.



شکل ۱. اسپکتروگرام سیگنال سنتز شده برای نویز همهمه

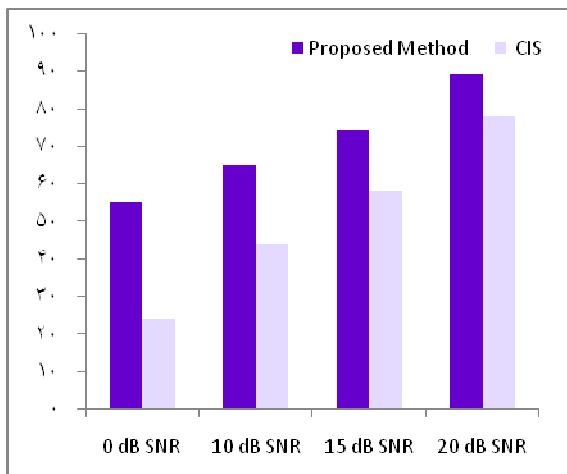
در شکل ۲، اسپکتروگرام سیگنال سنتز شده، با استفاده از الگوریتم ارائه‌شده برای نویز همهمه نشان داده شده است.

اسپکتروگرام سیگنال سنتز شده برای نویز قطار را می‌توان در شکل ۳ دید.

اسپکتروگرام به منظور نمایش طیف زمان-فرکانس سیگنال گفتاری استفاده می‌شود. همان طور که مشاهده می‌شود، استفاده از الگوریتم ارائه شده موجب کاهش اغتشاش فرکانسی موجود در کانال‌ها گردیده است و باعث تفکیک فرکانسی کانال‌ها می‌شود.

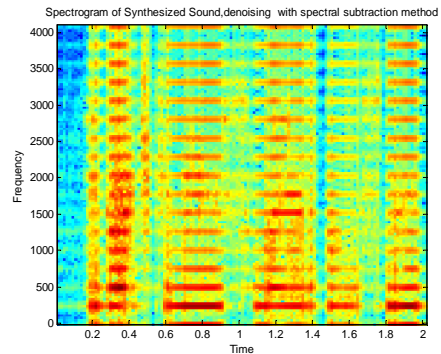
ارزیابی ذهنی

شکل ۵ درصد بازشناسی گفتار را تابع SNR برای دو شرایط پردازشی، با الگوریتم‌های پیشرفته‌ی پیشنهادی و CIS را نشان می‌دهد. به طور کلی، در تمام SNRها، افراد با الگوریتم پیشنهادی افزایش درکی در حدود ۴۵ درصد را نشان می‌دهند. آزمون ذهنی توسط ۱۵ شنونده با شنوایی طبیعی با استفاده از یک شبیه‌ساز CI در کلینیک شنوایی‌شناسی دانشگاه علوم پزشکی و خدمات بهداشتی درمانی اصفهان در مهر ماه سال ۱۳۹۱ انجام گرفت.

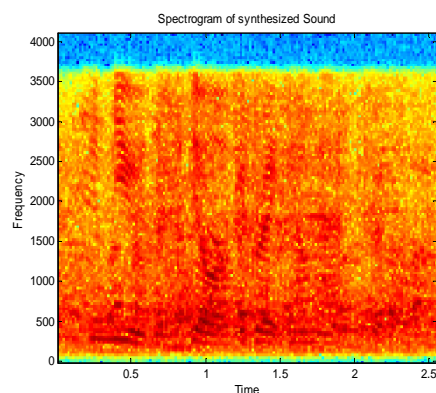


شکل ۵. درصد بازشناسی گفتار برای دو شرایط پردازشی الگوریتم‌های پیشرفته‌ی پیشنهادی و CIS

همچنین به منظور بررسی تأثیر کانال‌ها بر روی درصد بازشناسی گفتار، از فیلتر بانک‌های ۶، ۸ و ۱۶ کاناله استفاده گردید. شکل ۶ درصد بازشناسی گفتار

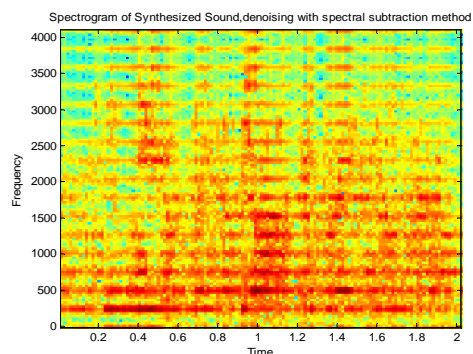


شکل ۲. اسپکتروگرام سیگنال سنتز شده، با استفاده از الگوریتم ارائه شده برای نویز همهمه



شکل ۳. اسپکتروگرام سیگنال سنتز شده برای نویز قطار

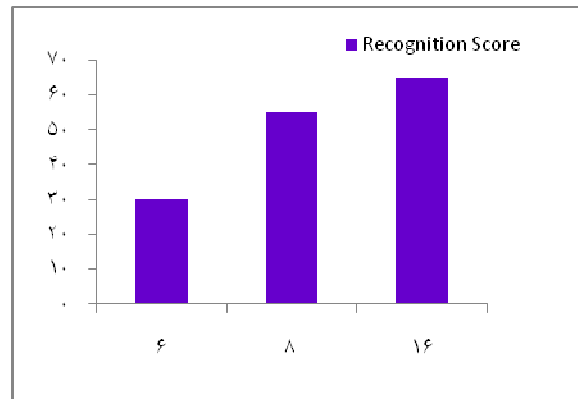
در نهایت در شکل ۴ اسپکتروگرام سیگنال سنتز شده، با استفاده از الگوریتم ارائه شده برای نویز قطار آمده است.



شکل ۴. اسپکتروگرام سیگنال سنتز شده، با استفاده از الگوریتم ارائه شده برای نویز قطار

طیفی با استفاده از طبقه‌بندی بین غالب بودن گفتار و یا نویز در هر کانال بود. در طبقه‌بندی پیشنهادی از انحراف معیار طیف سیگنال مشاهده شده در هر کانال استفاده گردید. در نهایت، تفریق طیفی، پس از طبقه‌بندی انجام شد. عملکرد الگوریتم کاهش نویز پیشنهاد شده از طریق ارزیابی عینی، توسط SNR قطعه‌ای با استفاده از جملات Noisy92 که توسط نویز مهمه، اتومبیل و قطار در SNR ۰ تا ۲۰ دسی‌بل مخدوش گردیده بودند و مشاهده‌ی طیف سیگنال و همچنین آزمون ذهنی توسط ۱۵ شنونده با شنوایی طبیعی با استفاده از یک شبیه‌ساز CI در کلینیک شنوایی‌شناسی دانشگاه علوم پزشکی و خدمات بهداشتی درمانی اصفهان در مهر ماه سال ۱۳۹۱ انجام گرفت. در تمام SNRها توسط الگوریتم پیشنهادی، افزایش درکی در حدود ۶۵ درصد مشاهده شد. بر اساس ارزیابی‌های انجام شده، روش پیشنهادی موجب کاهش نویز و اعوجاج گفتاری می‌گردد.

را تابع تغییر تعداد کانال‌ها را نشان می‌دهد. همان طور که مشاهده می‌شود، هر چه تعداد باندهای فرکانسی افزایش یابد، درصد بازشناسی گفتار نیز افزایش خواهد یافت.



شکل ۶. درصد بازشناسی گفتار برای از فیلتر بانک‌های ۶، ۸ و ۱۶ کاناله

بحث

در این مقاله، به منظور بهبود عملکرد کاشت حلزون در محیط‌های نویزی، یک الگوریتم کاهش نویز برای سیستم کاشت حلزون ارائه شده است که شامل تفریق

References

1. Nie K, Stickney G, Zeng FG. Encoding frequency modulation to improve cochlear implant performance in noise. *IEEE Trans Biomed Eng* 2005; 52(1): 64-73.
2. Wilson BS, Lawson DT, Muller JM, Tyler RS, Kiefer J. Cochlear implants: some likely next steps. *Annu Rev Biomed Eng* 2003; 5: 207-49.
3. Xu L, Tsai Y, Pflingst BE. Features of stimulation affecting tonal-speech perception: implications for cochlear prostheses. *J Acoust Soc Am* 2002; 112(1): 247-58.
4. Yao J, Zhang YT. The application of bionic wavelet transform to speech signal processing in cochlear implants using neural network simulations. *IEEE Trans Biomed Eng* 2002; 49(11): 1299-309.
5. Zeng FG, Nie K, Stickney GS, Kong YY, Vongphoe M, Bhargava A, et al. Speech recognition with amplitude and frequency modulations. *Proc Natl Acad Sci USA* 2005; 102(7): 2293-8.
6. Weiss MR. Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor. *J Rehabil Res Dev* 1993; 30(1): 117-28.
7. Hochberg I, Boothroyd A, Weiss M, Hellman S. Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear Hear* 1992; 13(4): 263-71.
8. McDermott HJ, McKay CM, Vandali AE. A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *J Acoust Soc Am* 1992; 91(6): 3367-71.
9. Hu Y, Loizou PC. A subspace approach for enhancing speech corrupted with colored noise. *IEEE Signal Processing Letters* 2002; 9(7): 204-6.

10. Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M. Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *J Rehabil Res Dev* 1993; 30(1): 110-6.
11. Dorman MF, Loizou PC. Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with continuous interleaved sampling (CIS) processors. *Ear Hear* 1997; 18(2): 147-55.

An Improved Spectral Subtraction Algorithm for Noise Reduction in Cochlear Implants with Increasing Number of Channels

Marjan Mozaffarilegha MSc¹, Saeed Kermani PhD²

Original Article

Abstract

Background: Cochlear implants are widely accepted as the unique and most effective way for individuals with severe to profound hearing loss to restore some degree of hearing. Speech enhancement strategies play an extremely important role in optimizing the cochlear implant. In this study, a noise reduction algorithm was proposed for cochlear implant.

Methods: To improve the performance of cochlear implants in noisy environments, a noise reduction algorithm was proposed for cochlear implant that applied a spectral subtraction using the classifications between the speech and the noise dominants in each channel. The proposed classifications use the standard deviation of the spectrum of observation signal in each channel. The performance of the proposed noise reduction algorithm was evaluated by segmental signal-to-noise ratio (SNR) using Noisy92 sentences embedded in babble, car noise and train at 0–20 dB. SNR and subjective listening tests were assessed with 15 normal hearing listeners using a specific cochlear implant (CI) simulator in Clinical audiology of Isfahan University of Medical Sciences, Iran, in June 2012. In addition, to evaluate the effect of the channel on speech recognition, 6, 8 and 16 channels of filter banks were used.

Findings: At all SNR values, subjects performed better with proposed enhanced algorithm by 45 percentage points.

Conclusion: Based on comparing segmental SNR with spectral subtraction, and visually inspecting the enhanced spectrograms and subjective listening test, the proposed method was found to effectively reduce noise while minimizing distortion to speech.

Keywords: Cochlear implants, Improved spectral subtraction, Speech enhancement

Citation: Mozaffarilegha M, Kermani S. **An Improved Spectral Subtraction Algorithm for Noise Reduction in Cochlear Implants with Increasing Number of Channels.** J Isfahan Med Sch 2013; 31(257): 1667-74

* This paper is derived from a MSc thesis No. 392159 in Isfahan University of Medical Sciences.

1- MSc Student, Department of Biomedical Engineering, School of Medicine AND Student Research Committee, Isfahan University of Medical Sciences, Isfahan, Iran

2- Associate Professor, Department of Biomedical Engineering, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Corresponding Author: Saeed Kermani PhD, Email: kermani@med.mui.ac.ir