

بررسی مدل‌های آماری سیگنال‌های صحبت و ارائه نرم‌افزاری برای جداسازی منابع کور

مهسا برقی مهمان‌داری
دانشگاه آزاد اسلامی واحد تبریز
sama_68@yahoo.com

نیلوفر رجبیون
دانشگاه آزاد اسلامی واحد تبریز
باشگاه پژوهشگران جوان واحد تبریز
n_rajabion84@yahoo.com

چکیده- انتخاب دقیق تابع غیرخطی برای منبع سیگنال جهت اجرای هرچه بهتر الگوریتم‌های فرایند جداسازی منابع صحبت کور ضروری است. در این مقاله، یک روش سریع برای پیاده سازی جداسازی منابع کور (BSS) بر اساس گرادیان طبیعی ارائه شده است که هدف اصلی استفاده از یک تابع غیرخطی جدید، مطابق با تابع چگالی احتمال (pdf) صحیحی از سیگنال‌های صحبت است؛ که با انتخاب صحیح shape parameters می‌توان منابع سیگنال صحبت را با توزیع مرکب نرمال-لاپلاس مدل کرد و با مشتق گیری pdf آنرا بدست آورد و تابعی غیرخطی که برای الگوریتم BSS مناسبتر است از این مدل توزیع، مشتق کرد. نتایج شبیه سازی با یک همگرایی خوب روش پیشنهادی را تأیید می‌کند. کلمات کلیدی- منابع صحبت، الگوریتم گرادیان طبیعی، تابع مرکب نرمال-لاپلاسن، BSS

۱- معرفی

جداسازی منابع کور روشی است که اخیراً در زمینه‌های مختلف پژوهشی توجه زیادی به آن شده است و وظیفه اصلی BSS، بازیابی سیگنال‌های حاصل از ترکیب منابع با استفاده از مقادیر آماری آنها می‌باشد. در اکثر تحقیقاتی که در این زمینه انجام شده است مجموعه ای از سیگنال‌های سنس شده که شامل ترکیب خطی از چندین منبع سیگنال است به عنوان پیش فرض در نظر گرفته شده است، هدف اصلی پردازش این سیگنال‌ها، دستیابی به منابع سیگنالی است که هیچگونه اطلاعی از این منابع و شرایط ترکیب آنها در دست نیست.

یک روش متداول برای جداسازی منابع کور، انتخاب توابع غیرخطی مناسبی است که بیشترین شباهت را با pdf منابع دارد. از آنجایی که pdf منابع برای انجام الگوریتم‌های BSS ناشناخته است برای هر منبع یک مدل توزیعی فرضی در نظر گرفته شده و تابع غیرخطی از pdf آن مشتق می‌شود. برای مثال، منابع سیگنال سوپر گوسی مانند سیگنال‌های صحبت که توزیع لاپلاسنی دارند تابع sign و برای سیگنال‌های ساب‌گوسی تابع cubic انتخاب مناسبی خواهد بود.

۲- الگوریتم گرادیان طبیعی برای جداسازی منابع کور

n منبع سیگنال مستقلی را فرض کنید که توسط ماتریسی مربعی و نامعلوم که ماتریس ترکیب نامیده می‌شود با هم ترکیب شده‌اند و توسط ماتریس $s(t) = [s_1(t), s_2(t), \dots, s_n(t)]^T$ نشان داده می‌شوند. m سیگنال ترکیبی با بردار $x(t) = [x_1(t), x_2(t), \dots, x_m(t)]^T$ بیان می‌شوند که این ماتریس مبین سیگنال‌های سنس شده است.

در این مقاله سیگنال‌های صحبت را در حوزه زمان فرض کرده و شبیه سازی‌هایمان را بر پایه تابع غیرخطی جدیدی

$$x(t) = H.s(t) + n(t) \quad ; x(t) \in \mathbb{R}^m \quad (1)$$

۳- مدل توزیع مرکب نرمال لاپلاسین برای منابع صحبت

با توجه به اینکه فرم تابع غیر خطی که با فرمول ۴ بیان می شود وابسته به توزیع احتمال منابع است. می توان مدل های چگالی فرضی متنوعی برای آنها در نظر گرفت، بطور مثال برای سیگنالهای صحبت که توزیع سوپر گوسی دارند مدل توزیع لاپلاسی مدلی مناسب است که تابع چگالی احتمال آن $(p(y))$ بصورت زیر تعریف می شود:

$$p(y) = \frac{1}{2a} \exp\left(\frac{-|y|}{a}\right) \quad (7)$$

که می توان تابع غیرخطی sign را از آن نتیجه گرفت:

$$\varphi_i(y_i) = \text{sign}(y_i) \quad (8)$$

و $a)$ همان تابع معروف علامت است.

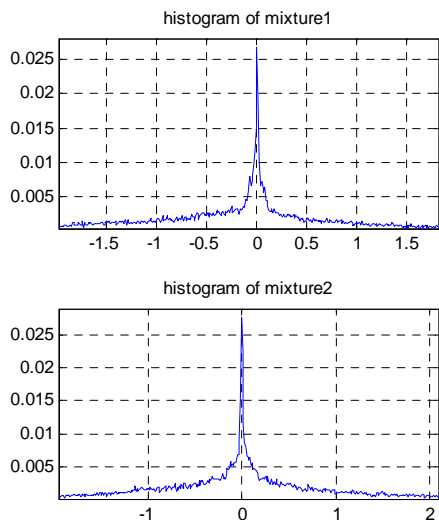
در این مقاله توزیع مرکب نرمال- لاپلاسین برای سیگنالهای صحبت در نظر گرفته شده است، که می توان تابع غیرخطی جدیدی از آن مشتق گرفت. Pdf این توزیع بصورت زیر تعریف می شود:

$$p(y) = \frac{k}{2\varphi} \exp\left(-\frac{|y-\theta|}{\varphi}\right) + \frac{1-k}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-\theta)^2}{2\sigma^2}\right] \quad (9)$$

θ مقدار متوسط، φ مقدار متوسط مطلق و k ,

shapeparameter ای است که نرخ نمایی نزول را نشان می دهد. برای موارد خاص که $k=1$ یا $k=2$ ، تابع $p(y)$ برای توزیع مرکب نرمال- لاپلاسین همان pdf توزیع لاپلاسی یا گوسی است. با دقت در نمونه های حاصل از نمونه برداری سیگنال های صحبت می توان دریافت این نمونه ها همانند متغیرهای تصادفی هستند.

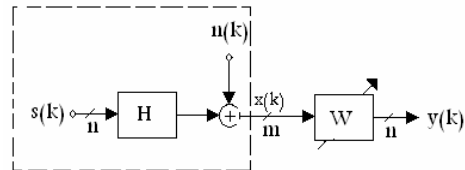
شکل ۲ هیستوگرام سیگنالهای دریافتی را نشان می دهد.



در این مقاله سیگنال دریافتی صحبت بدون در نظر گرفتن نویز بیان شده و تعداد منابع و سیگنالهای دریافتی یکسان فرض شده است.

هدف اصلی تکنیکهای جداسازی منابع کور، تعیین ماتریس جداساز w است که همانند عملگری معکوس در فرایند ترکیب عمل می کند.

شکل ۱ بلوک دیاگرام این مدل را نشان می دهد.



شکل ۱) بلوک دیاگرام مربوط به BSS

روشهای گوناگونی بدین منظور با استفاده از مقادیر آماری منابع توسط محققین در ریاضیات کاربردی، شبکه های عصبی، تشخیص الگو... ارائه شده است.

حال شبکه ای خطی را در نظر بگیرید که خروجی $y(t)$ آن به شکل زیر تعریف می شود:

$$y(t) = Wx(t) \quad (2)$$

که (i, j) -امین درایه از ماتریس $(W_{i,j})$ نشان دهنده وزن سیناپسی بین $x_i(t)$ و $y_j(t)$ است.

تابع تلفات را نیز مطابق فرمول زیر تعریف می کنیم:

$$L(W) = -\log|\det W| - \sum_{i=1}^n \log p_i(y_i) \quad (3)$$

که $p_i(\cdot)$ معرف تابع چگالی احتمال است.

حال با فرض

$$\varphi_i(y_i) = -\frac{d \log p_i(y_i)}{dy_i} \quad (4)$$

گرادین تابع تلفات بصورت زیر بیان می شود:

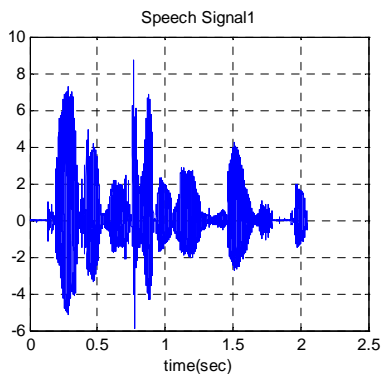
$$\nabla L(W) = \frac{\partial L(W)}{\partial W} = -W^{-T} + \varphi(y)x^T \quad (5)$$

با استفاده از گرادین طبیعی $(\nabla L(W))$ به الگوریتم زیر برای یافتن W در فرایند BSS می رسیم.

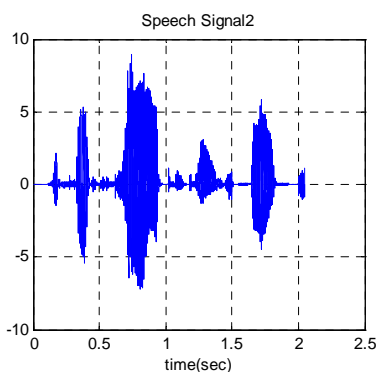
$$W(t+1) = W(t) - \eta_t \nabla L(W) \quad (6)$$

$$\begin{aligned} & W(t) - \eta_t \nabla L(W) W^T(t) W(t) \\ & W(t) - \eta_t (-W^{-T} + \varphi(y)x^T) W^T(t) W(t) \\ & W(t) + \eta_t \{I - \varphi(y(t))y^T(t)\} W(t) \end{aligned}$$

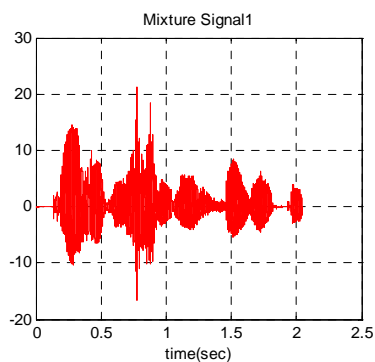
شکل ۲) هیستوگرام سیگنال های دریافتی



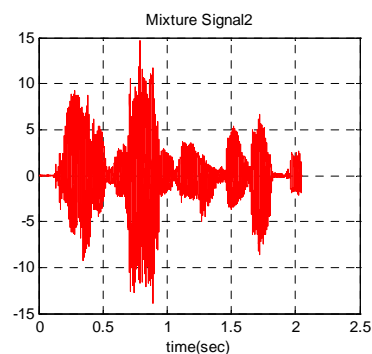
شکل ۳



شکل ۴



شکل ۵



شکل ۶

حال اگر از معادله ۹ مشتق جزئی نسبت به y بگیریم، تابع غیرخطی جدیدی بصورت زیر برای جداسازی سیگنالهای کور تعریف می شود:

$$\varphi_i(y_i) = -\frac{d \log(p_i(y_i))}{dy_i} \quad (10)$$

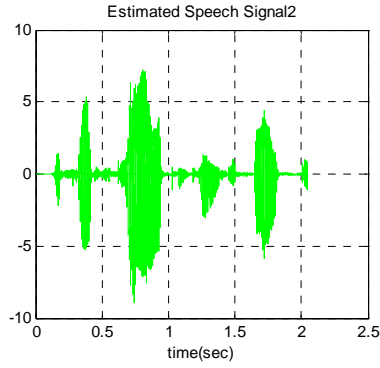
$$= \frac{\sqrt{2\pi} \exp(-|y_i|) \text{sign}(y_i) k + 2 \exp\left(\frac{-y_i^2}{2}\right) y_i (1-k)}{\sqrt{2\pi} \exp(-|y_i|) + 2 \exp\left(\frac{-y_i^2}{2}\right) (1-k)}$$

۴- نتایج شبیه سازی

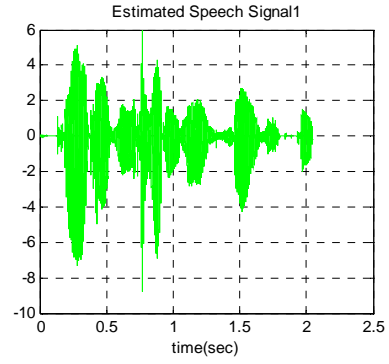
در این بخش از مقاله، عملکرد تابع غیرخطی جدید (۱۰) برای BSS با فرض ماتریس A به عنوان ماتریس ترکیب توسط نرم افزار Matlab مورد بررسی قرار گرفته و نتایج شبیه سازی نمایش داده شده است.

$$A = \begin{bmatrix} 2 & 1 \\ 1.25 & -1.5 \end{bmatrix}$$

شکل های ۳ و ۴ سیگنالهای منبع اصلی، شکل های ۵ و ۶ سیگنالهای دریافتی و شکل ۷ و ۸ سیگنالهای تخمینی توسط الگوریتم ارائه شده را نشان می دهد. مقایسه شکل های ۳ و ۴ با شکل های ۷ و ۸ عملکرد صحیح این الگوریتم را تأیید می کند.

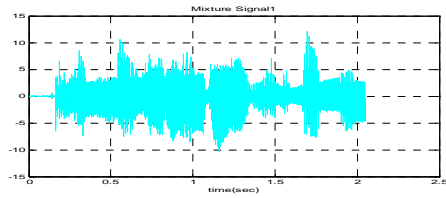


شکل ۸

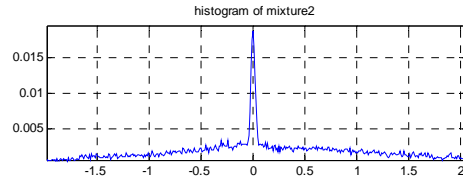
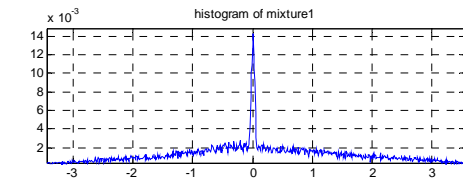


شکل ۷

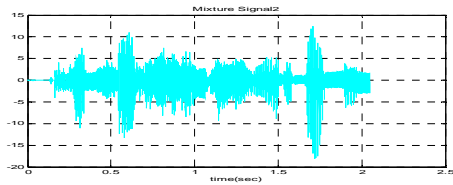
نتایج شبیه‌سازی برای دو سیگنال دیگر در زیر آورده شده است:



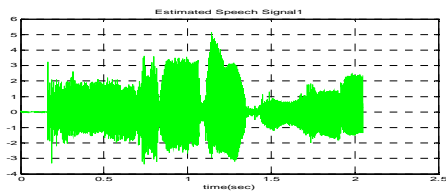
شکل ۱۲) mixture signal 1



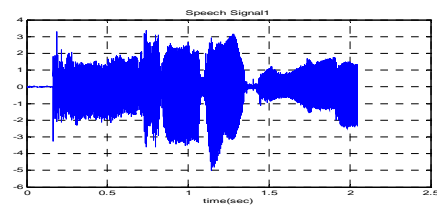
شکل ۹



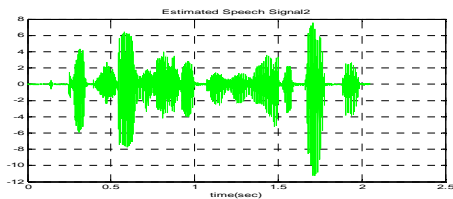
شکل ۱۳) mixture signal 2



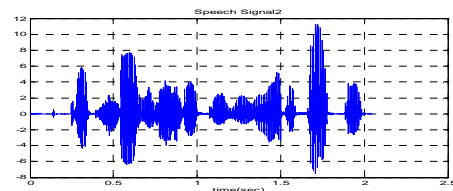
شکل ۱۴) estimated speech signal 1



شکل ۱۰) speech signal 1



شکل ۱۵) estimated speech signal 2



شکل ۱۱) speech signal 2

۵- نتیجه‌گیری

در این مقاله با بررسی حول مدل‌های توزیع مختلف برای منابع صحبت این نتیجه حاصل شد که علی‌رغم استفاده متداول از مدل توزیع لاپلاسی، مدل توزیع مرکب نرمال - لاپلاسیین مدل مناسبتری برای سیگنال‌های صحبتی است که حالت سوپر گوسی دارند. با تکیه بر این نتیجه و با مشتق-گیری از pdf مدل توزیع مرکب نرمال - لاپلاسیین، تابع جدیدی برای BSS ارائه شد که با فرض مقادیر مناسب برای σ و θ شبیه‌سازیها توسط نرم‌افزار matlab انجام شد که خود تائیدی بر گفته‌های تئوری است.

مراجع

- [1] Bell A and Sejnowski T. An information maximization approach to blind separation and blind deconvolution. 1995,7: 1129-1159.
- [2] Cichocki A and Amari S. Adaptive blind signal and image processing: learning methods and application.
- [3] Hu Y , and Loizou P C. "A generalized subspace approach for enhancing speech corrupted by colored noise". IEEE Transactions on Speech and Audio Processing ,2003 ,11(4):334 -341.
- [4] Choi S, Cichocki A and Amari . Flexible Independent Component Analysis, Journal of VLSI Signal Processing, 2000, 26(1/2).