Arshive of SID

Medical Journal
of
Tabriz University of Medical Sciences and Health Services

Original Article

The Intelligent Diagnosis of Speech Disorder: Using of Audio Signal Processing System

Iman Abaspur Kazerouni, Hasan Moslemi*, Fatemeh Hourali

Electrical and Computer Engineering Department, Esfarayen University of Technology, Esfarayen, Iran.
*Corresponding author; E-mail: h.moslemi@esfarayen.ac.ir

Received: 5 October 2016 Accepted: 15 January 2017 First Published online: 13 December 2018
Med J Tabriz Uni Med Sciences Health Services. 2018 December - 2019 January; 40(5):55-62

Abstract

Background: Voice is one aspect of speech. In this study, evaluation of the audio signal is used to determine the severity of the speech disorder and intelligibility. Until now, methods studied the quality of speech.

Methods: In consultation with the specialists a sentence with 5 words and 8 syllables has been determined to test 80 sample random persons were randomly selected among healthy people with different age groups and gender. Unhealthy people randomly selected from among the pupils centers. Among the different grading unhealthy and speech disorders were present. The cases were 50 healthy people and 30 patients with speech disorders. Using the model proposed in this paper, an intelligent system designed base on the principles of speech and image processing for detection speech disorder.

Conclusion: The proposed system has been tested on the different voice signals and statistical data have been extracted. The results represent a breakdown of people with an accuracy of more than 92% has been compared with the results of other reports, the result is very acceptable. The percentage of the proposed system error that leads to a repeat of the experiment is also about 7%.

Keywords: Speech Intelligibility, Voice Quality, Classification, Data Collection

How to cite this article: Abaspur Kazerouni I, Moslemi H, Hourali F. [The Intelligent Diagnosis of Speech Disorder: Using of Audio Signal Processing System]. Med J Tabriz Uni Med Sciences Health Services. 2018 December - 2019 January;40(5):55-62. Persian.

© 2018 The Author(s). This is an Open Access article published by Tabriz University of Medical Sciences under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

www.SID.ir

Arshive of SID

مقاله پژوهشی

تشخیص هوشمند اختلال گفتاری با استفاده از سیستم پردازش سیگنال صوت

ایمان عباس‌پور کازرونی ، حسن مسلمی* ، فاطمه حورعلی

دانشکده مهندسی برق و کامپیوتر، مجتمع آموزش عالی فنی و مهندسی اسفراین، اسفراین، ایران
*نویسنده مسئول؛ ایمیل h.moslemi@esfarayen.ac.ir

دریافت: ۱۳۹۵/۷/۱۴ پذیرش: ۱۳۹۵/۱۰/۲۶ انتشار برخط: ۱۳۹۷/۹/۲۲
مجله پزشکی دانشگاه علوم پزشکی و خدمات بهداشتی ـ درمانی تبریز. ۱۳۹۷ آذر و دی؛ ۴۰(۵):۵۵-۶۲

چکیده

زمینه: صوت یکی از جنبه های گفتار است. در این پژوهش، هدف بررسی و ارزیابی سیگنال صوتی جهت تعیین شدت اختلال گفتاری و همچنین وضوح گفتار است. تاکنون روش‌ها جهت ارزیابی گفتار به صورت کیفی بوده است.

روش‌کار: با مشورت متخصصان گفتار درمانی یک جمله شامل ۵ کلمه و ۸ هجا تعیین شد و از ۸۰ نفر نمونه گیری شد که به صورت تصادفی در میان افراد سالم و از گروه های سنی و جنسیت مختلفی انتخاب شده اند. افراد ناسالم نیز از میان دانش آموزان مراکز استثنایی به صورت تصادفی انتخاب شده اند. از میان نمونه های انتخاب شده ۵۰ شخص سالم و ۳۰ نفر دارای اختلالات گفتاری هستند. در این مقاله، سیستم هوشمندی طراحی شده که بر اساس اصول پردازش گفتار و تصویر، می تواند افراد سالم را از افراد دارای اختلالات گفتاری تفکیک کند.

یافته‌ها: از آنجا که تشخیص وجود اختلالات گفتاری، معمولا تنها بر اساس تشخیص متخصصان گفتار درمانی و بر اساس اصول شنیداری انجام می شده است لذا با سیستم پیشنهادی طراحی شده در این مقاله، می توان عملیات تشخیص را به صورت هوشمند و با دقت بسیار بالا و بدون خطای انسانی انجام داد.

نتیجه‌گیری: نتایج حاصل نشان دهنده تفکیک افراد با دقت بیش از ۹۲٪ بوده است که در مقایسه با نتایج سایر گزارشات نتیجه بسیار قابل قبولی می باشد. درصد خطای سیستم پیشنهادی که منجر به تکرار آزمایش شده است هم در حدود ۷٪ می باشد.

کلید واژه‌ها: وضوح گفتار، کیفیت صوت، دسته بندی، جمع آوری داده

نحوه استناد به این مقاله: عباسپور کازرونی ا، مسلمی ح، حورعلی ف. تشخیص هوشمند اختلال گفتاری با استفاده از سیستم پردازش سیگنال صوت. مجله پزشکی دانشگاه علوم پزشکی و خدمات بهداشتی ـ درمانی تبریز. ۱۳۹۷؛ ۴۰(۵):۵۵-۶۲

حق تألیف برای مؤلفان محفوظ است.
این مقاله با دسترسی آزاد توسط دانشگاه علوم پزشکی و خدمات بهداشتی ـ درمانی تبریز تحت مجوز کریتیو کامنز (http://creativecommons.org/licenses/by/4.0) منتشر شده که طبق مفاد آن هرگونه استفاده تنها در صورتی مجاز است که به اثر اصلی به نحو مقتضی استناد و ارجاع داده شده باشد.

www.SID.ir

Arshive of SID

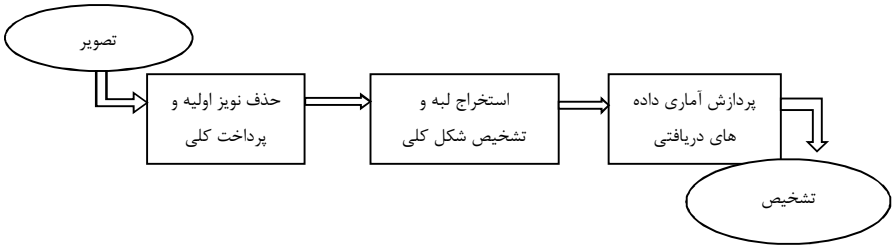
مقدمه

در فرآیند ارتباط کلامی صوت ابزاری است که نه تنها پیام را انتقال می دهد بلکه به معنای آن نیز می افزاید. احساس گوینده را بیان می کند و می تواند پیام را قوی‌تر به شنونده انتقال دهد. صوت علاوه بر آنکه بیانگر حالت های عاطفی، هیجانی و روانی گوینده است، می تواند حالت‌های جسمی گوینده به ویژه وضعیت حنجره را مشخص نماید. صوت نابهنجار می‌تواند به دلیل وجود انواع بیماری‌ها اعم از عضوی یا عملکردی باشد (۱). صوت نابهنجار به هرگونه انحراف در زیر و بمی بلندی کیفیت یا سایر ویژگی‌های پایه صوت اطلاق می‌شود که به طور پیوسته در برقراری ارتباط اختلال ایجاد کرده، توجه سایرین را جلب نماید و بر گوینده و شنونده تأثیر نامطلوب بگذارد (۲). در منابع مختلف وضوح گفتار تعاریف متفاوتی دارد، اما به طور کلی وضوح گفتار به وسعت سیگنال‌های آکوستیکی اشاره دارد که گوینده تولید کرده و شنونده به طور درست دریافت می‌کند (۳). به طور کلی تعامل ویژگی‌های زنجیری و زبر زنجیری وضوح گفتار را تحت تأثیر قرار می‌دهند. به این معنی که تولید، سرعت، روانی، کیفیت و شدت صوت عوامل موثر بر وضوح گفتار هستند (۴). سنجش وضوح گفتار کاربردهای زیادی دارد از جمله مبنای مقایسه به منظور استناد و بررسی تغییر در عملکرد گفتار در پاسخ به درمان، تعیین محدودیت عملکردی گفتار، سنجش شدت اختلال گفتاری و وضوح گفتار به عنوان معیار آسیب، ناتوانی یا معلولیت نیز به کار می‌رود. با این وجود معمولاً در گزارشات بالینی و پژوهش‌ها، مستند کردن سطوح وضوح گفتار رایج نیست (۵). ارزیابی وضوح گفتار باید شامل تک کلمه، تکرار جمله و گفتار خود انگیخته باشد. تخمین وضوح تک کلمه پیش‌بینی خوبی برای وضوح جمله یا گفتار پیوسته نخواهد بود (۶). در ادامه با توجه به پایه‌های ریاضی و منطقی مدلی ارایه می شود که بر اساس آن، با پردازش سیگنال های صوتی، صوت افراد دارای اختلالات صوتی از افراد سالم تفکیک می‌شود. در بخش ۲ ساختار کلی مدل پیشنهادی ارایه می‌شود و در آن عملگرهای منطقی مختصاتی که پایه اصلی مدل پیشنهادی است توضیح داده می‌شود. در بخش ۳، دو مرحله اصلی حذف نویز و استخراج لبه‌های سیگنال به صورت زیر مجموعه روش بررسی ارایه می‌شود و در بخش ۴ نتایج مدل پیشنهادی شرح داده می‌شود.

روش‌کار

در این بخش الگوریتم و ساختار کلی استفاده شده در این مقاله مورد بررسی قرار می‌گیرد. در روش پیشنهادی، ابتدا نویزهای موجود در تصویر سیگنال صوتی دریافتی با استفاده از فیلترهای پیشنهادی حذف نویز، از بین می‌روند. باید توجه داشت که عمل حذف نویز در تحلیل تصاویر پزشکی بسیار حیاتی و مهم است به دلیل اینکه در صورتی که نویزهای موجود در یک تصویر به طور کامل حذف نشوند، عملیات و پردازش‌های بعدی علاوه بر تصویر هدف، بر روی نویزها نیز انجام می‌گیرند و بنابراین نتیجه‌ی به دست آمده از تحلیل و بررسی تصاویر نمی‌تواند مفید باشد. پس از انجام عملیات حذف نویز، مرحله‌ی تشخیص و جداسازی لبه‌های تصویر صورت می‌گیرد. در این مرحله، لبه ها و حاشیه‌های موجود در تصویر استخراج شده و مورد بررسی قرار می‌گیرد. در مدل پیشنهادی تمرکز بر روی وجود فاصله های ناهمگون در میان ادای واژه‌ها و همچنین شیب ادای واژه ها در هنگام صعود و نزول ادای آنهاست. بنابراین می توان تنها با استفاده از داده های موجود در لبه های تصاویر به تحلیل سیگنالها پرداخت. شکل ۱ مراحل و شمای کلی الگوریتم مدل پیشنهادی را به صورت فلوچارت تصویری نمایش می‌دهد.

عملگرهای منطقی مختصاتی (Coordinate Logic Operation)، یکی از ابزارهای مهم برای انجام عملیات هایشناسایی الگو و پردازش تصویر هستند. زمینه ها و کاربردهایی مانند کاهش و حذف نویز تصاویر باینری و سطح خاکستری (Gray Level)، پر کردن نواحی و اشیا موجود در تصاویر، شناسایی و تشخیص لبه ها (۷و۸) و همچنین کاربردهای تشخیص پزشکی از این دسته اند (۹و۱۰). یکی از روش‌های استفاده از این عملگرها ایجاد فیلترهای مختلف برای کاربردهای خاص است. فیلترهای منطقی مختصاتی (Coordinate Logic Filters) دسته‌ای از فیلترهای دیجیتالی غیر خطی اند که با استفاده از عملگرهای منطقی ساخته می شوند.



شکل (۱): شمای کلی الگوریتم مدل پیشنهادی.

www.SID.ir

Arshive of SID

عملکرد این فیلترها برای سیگنال‌ها و تصاویر باینری همانند فیلترهای مورفولوژیکال (Morphological) ولی برای تصاویر سطح خاکستری سریع تر عمل می‌کنند (۱۱). عملگرهای منطقی مختصاتی مورد استفاده برای فیلترها، AND، OR، XOR و ترکیبات آن‌ها هستند. اساس کار مدل های پیشنهادی ایجاد فیلترهای منطقی مختصاتی است. فیلترها بر اساس روابط و خواص حاکم بر عملگرهای منطقی ساخته می‌شوند. خواص جبر بولی حاکم بر روابط دیجیتالی برای این عملگرها نیز مورد استفاده قرار می‌گیرند. به عنوان نمونه با فرض دو مقدار ۱۰۰۱۰۰۱۱ و ۱۰۱۰۱۰۱۰ نتیجه‌ی حاصل از عمل AND مختصاتی آن‌ها برابر با ۱۰۰۰۰۰۱۰ خواهد بود. بنابراین عملیات منطقی مختصاتی برای دو عدد دهدهی W و V در سیستم باینری به صورت زیر تعریف می شود (۱۲):

$$X = V \bullet W \quad (۱)$$

که در آن • می‌تواند عملگر منطقی مختصاتی AND، OR، NOT، XOR و یا ترکیبی از آن‌ها باشد. برای تصویر مفروض I با تعداد M×N پیکسل که در آن N و M تعداد سطرها و ستون‌های تصویر است، درایه های ماتریس تصویر I به صورت زیر تعریف می‌شوند (۱۱):

$$S(ij)=\sum_{p=0}^{n-1}2^p(S_p(i,j)) \qquad i=1,2,..M,\ j=1,2,..N \quad (۲)$$

یافته ها

در این مطالعه ۸۰ نفر از افراد سالم و دارای اختلالات گفتاری، شامل ۳۲ مرد و ۱۸ زن سالم در محدوده سنی ۱۸تا ۲۵سال، با میانگین سنی ۲۳/۱۹ سال و ۱۲ مرد و ۱۸ زن دارای اختلالات گفتاری در محدوده سنی ۸ تا ۱۵سال، با میانگین سنی ۱۲/۲۴ سال بررسی شدند. تمام افراد منتخب در حال تحصیل و مجرد بوده اند. در مطالعه حاضر نمونه‌ها به دلیل انتخاب از دوره تحصیلی مختلف (دانشگاه و مدارس) همسان سازی نشده‌اند که این مورد از جمله محدودیت‌های موجود در مطالعه حاضر است. روش نمونه‌گیری و انتخاب افراد به صورت نمونه‌گیری منظم (سیستماتیک) انجام شده است. نمونه سالم ازدانشجویان دانشگاه حکیم سبزواری انتخاب شده‌اند و نمونه‌هایی که دچار اختلال گفتاری بودند از مراکز گفتار درمانی و مدارس استثنایی انتخاب شدند. با مشورت با صاحب نظران گفتار درمانگر (آسیب شناس گفتار و زبان) جمله‌ای انتخاب شد که محاوره ای باشد. این جمله شامل ۵ کلمه و ۸ هجا بود. آزمودنی‌ها در حین نمونه‌گیری در وضعیت نشسته و کاملاً صاف قرار می‌گرفتند و فاصله آزمودنی تا میکروفن ۳۰ سانتی‌متر بود. قبل از نمونه‌گیری از آزمودنی ها خواسته شد جمله ها را با بلندی عادی و حالت طبیعی و بدون تأکید بر روی واژه‌ای خاص و همچنین بدون مکث بیان کنند. تمامی شرکت‌کنندگان در آزمون فرم رضایت آگاهانه را مطالعه و تایید نموده‌اند و تمامی اصول مشخص شده در بیانیه هلسینکی برای انجام آزمون و انتشار نتایج رعایت شده است. نرم افزار مورد استفاده برای تحلیل داده ها و اجرای عملیات، نرم افزار MATLAB است که بر روی کامپیوتر با مشخصات Intel® Core™2Duo CPU T6400 @ 2.00GHz processor with 2GB of installed memory (RAM) under Microsoft Windows 7 operating نصب شده است و میکروفن مورد استفاده از مدل Rode Video Micro Camera Microphone با فرکانس پاسخ گویی ۱۰۰ تا ۲۰۰۰۰ هرتز و حساسیت ۳۳دسی بل می‌باشد. در مرحله اول در روش پیشنهادی با استفاده از تبدیل سیگنالهای صوتی به سیگنالهای تصویری، تصویر مربوط به صوت هر کدام از نمونه ها به دست می‌آید. باید توجه داشت که پردازش تصاویر با دقت بالاتر و سرعت بالاتری نسبت به پردازش صوت انجام می شود. برای پردازش تصاویر در ابتدا باید نویز مربوط به انتقال از صوت به تصویر را حذف کرد. این نویزهای موجود در تصاویر با استفاده از متغیرهای مختلف قابل جداسازی هستند. فیلتر مورد استفاده در روش پیشنهادی، فیلتر گاوسی است. از آنجا که صوت اشخاص سالم را به صورت نسبی تابعی از قانون گاوس هستند بنابراین می توان با اعمال یک فیلتر گاوسی بر روی تصاویر، حاشیه‌ها و نقاط اضافی را در صورت افراد در تصاویر حذف کرد. نویزهای موجود در تصاویر معمولاً ناشی از نویز مربوط به محیط در هنگام تهیه پایگاه داده می‌باشد. همچنین در هنگام ضبط صدا میکروفن نیز دارای نویزی است که در این مرحله تمام این نویزها به صورت قابل قبولی حذف می‌شوند. با استفاده از فیلتر گاوسی می توان نویز موجود در تصاویر را تا حدود قابل قبولی حذف کرد. از آنجا که این فیلتر از دو تابع میانگین و واریانس برای حذف نویز استفاده می‌کند بنابراین می توان با تکیه بر آن تا حدود زیادی به هدف حذف کامل نویزهای تصویر نزدیک شد. معادله مربوط به فیلتر گاوسی در زیر آمده است:

$$G(x) = exp\left(\frac{-0.5\,(x-\mu)^2}{\sigma^2}\right) \quad (۳)$$

که در این معادله μ بیانگر میانگین و σ واریانس داده ها است. این فیلتر به طور قابل قبولی کارایی لازم برای عملیات حذف نویز را دارد. با استفاده از فیلترهای منطقی مختصاتی می‌توان شناسایی الگوی لبه‌ها و استخراج و آشکارسازی لبه‌های تصاویر را انجام داد. مدل پیشنهادی برای انجام این کار به صورت فیلتری شامل عملگر منطقی مختصاتی XOR ارایه می‌شود. در این مدل با استفاده از تکنیک پهن شدگی (Dilation) بر روی تصویر، پیکسل های تشکیل دهنده‌ی لبه‌ها شناسایی و استخراج می‌شوند. این فیلتر با استفاده از تحلیل پیکسل‌های موجود در تصویر و به کمک

www.SID.ir

Arshive of SID

خواص عملگرهای منطقی به سادگی و با سرعت بالا آشکارسازی لبه‌ها را انجام می‌دهد. مدل فیلتر پیشنهادی به صورت زیر بیان می شود:

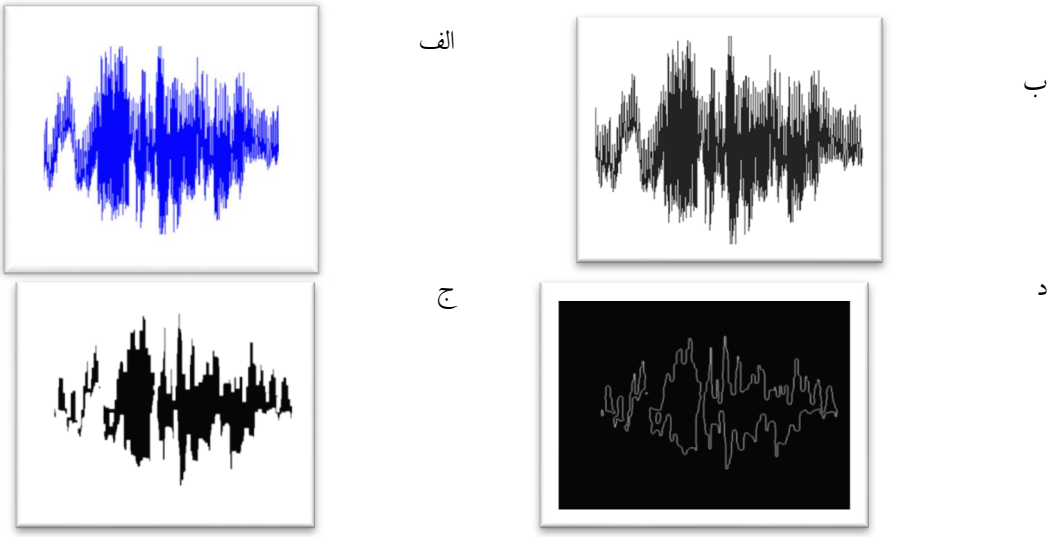
$$\text{Edge}(I_\gamma(i,j)) = I_\gamma(i-1,j) \bullet I(i+1,j) \bullet I_\gamma(i,j+1) \bullet I_\gamma(i,j-1) \quad (۴)$$

در معادله (۴)، • بیانگر عملگر منطقی مختصاتی XOR است. پس از این مرحله، داده‌های به دست آمده مورد پردازش آماری قرار می‌گیرند. از آنجا که بسیاری از داده های اضافی سیگنال اصلی در مراحل قبلی حذف گردیده‌اند بنابراین در این مرحله تنها کافیست که دو مشخصه ی اصلی داده‌ها یعنی میانگین و واریانس مورد بررسی و مقایسه قرار گیرند. باید توجه داشت که در صورتی که مراحل پردازشی قبل انجام نمی‌گرفتند، در این مرحله، این دو داده آماری نمی‌توانستند به تنهایی معیار مناسبی برای تشخیص اختلالات گفتاری باشند. در زیر روابط مربوط به میانگین و واریانس داده ها آورده شده است:

$$\text{Mean } \mu = \frac{1}{N}\sum_{j=1}^{N} P_j \quad (۵)$$

$$\text{Variance } \sigma^2 = \frac{1}{N-1}\sum_{j=1}^{N}(p_j - \mu)^2 \quad (۶)$$

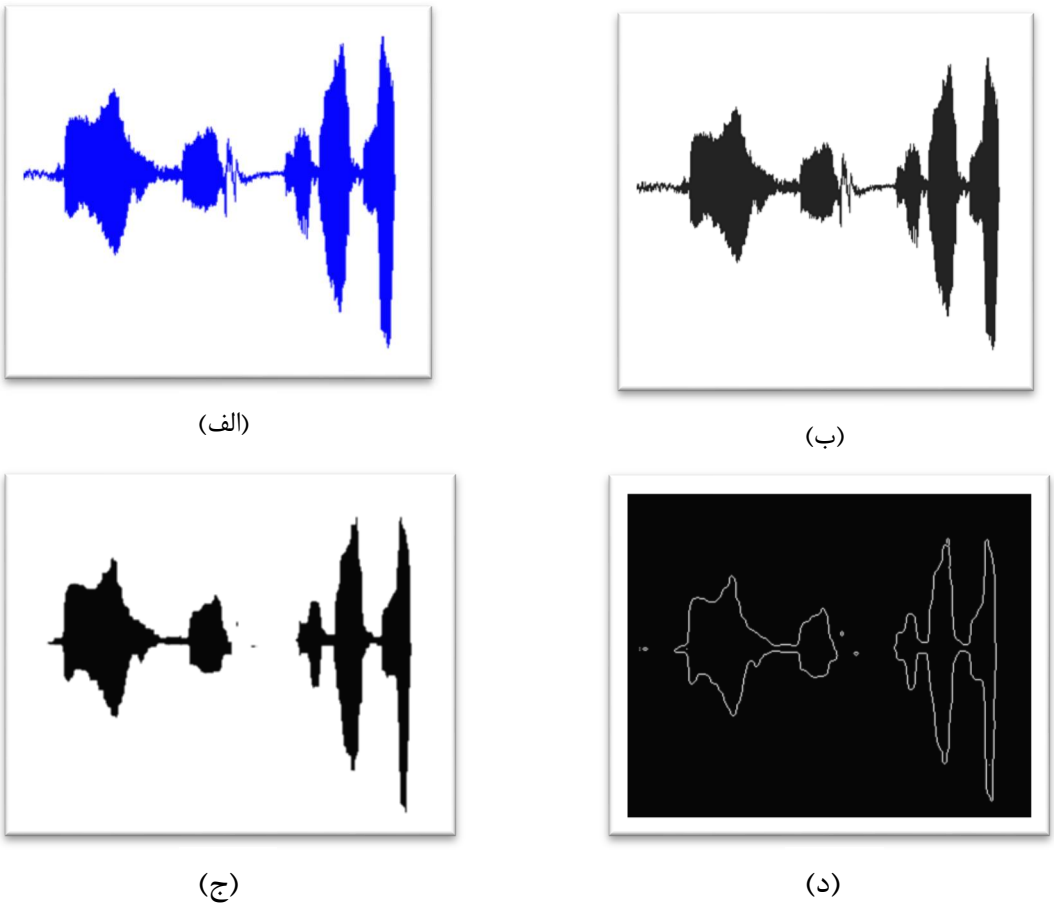
تصویر الف شکل ۲ مربوط به سیگنالهای یک شخص سالم است؛ همانطور که در این تصویر مشخص است فاصله خاصی بین ادای واژه ها وجود ندارد و فراز و فرودها به صورت پیوسته ومشخص اتفاق می افتد. در تصویر ب شکل تبدیل شده این سیگنال به صورت خاکستری(Gray) مشاهده می‌شود. دلیل خاکستری کردن تصویر اصلی این است که پردازش برروی تصاویر خاکستری با سرعت بیشتری انجام می‌شود. پس از این مرحله عملیات حذف نویز با استفاده از فیلتر پیشنهادی برروی تصویر انجام می‌شود و نتیجه حاصل در شکل ج قابل مشاهده است. همانطور که در این شکل مشخص است داده‌های اضافی تا حدود زیادی حذف شده اند و در نهایت عملیات تشخیص لبه بر روی تصویر اعمال می شود و تصویر حاصل بیانگر تمام اطلاعات مفید لازم برای پردازش داده ها می‌باشد و داده های اضافی به طور کامل حذف گردیده‌اند تا عملیات پردازش با سرعت و دقت بیشتری انجام شود.



شکل ۲: تصاویر مربوط به سیگنال صوتی فرد سالم. (الف): تصویر اصلی. (ب): تبدیل تصویر اصلی به تصویر خاکستری (ج): پس از انجام حذف نویز (د): تصویر نهایی پس از انجام عملیات جداسازی و استخراج لبه‌ها. در شکل(۳) تصاویر مربوط به شخصی با اختلال گفتاری دیده می شود. همانطور که در قسمت الف مشاهده می شود در سیگنالهای مربوط به این شخص فاصله مشهودی دیده می‌شود و فراز و فرودها بر خلاف شکل۲ به صورت پیوسته اتفاق نمی‌افتد بنابراین داده‌های آماری در این تصویر متفاوت از شخص سالم است. تمام مراحل پردازش که برای شکل ۲ بیان شد برای این سیگنال هم تکرار می‌شود.

www.SID.ir

Arshive of SID



(الف) (ب)

(ج) (د)

شکل ۳: تصاویر مربوط به سیگنال صوتی فرد دارای اختلال گفتاری. (الف): تصویر اصلی. (ب): تبدیل تصویر اصلی به تصویر خاکستری (ج): پس از انجام حذف نویز (د): تصویر نهایی پس از انجام عملیات جداسازی و استخراج لبه‌ها.

جدول ۱، نتایج مربوط به انجام مدل پیشنهادی برروی ۸۰ نفر از افراد سالم و دارای اختلالات گفتاری را نشان می‌دهد. همان‌طور که در این جدول مشخص است محدوده آماری افراد سالم و دارای اختلالات گفتاری اختلاف مشهودی دارند و بنابراین می‌توان با استفاده از مدل پیشنهادی به تفکیک افراد پرداخت. باید توجه داشت که در صورتی که محدوده های نزدیک به یکدیگر باشند، آزمایش تکرار می‌شود تا از نتایج اطمینان حاصل شود. تجزیه و تحلیل داده‌ها با استفاده از آزمون تحلیل واریانس انجام شد.

جدول ۱: نتایج آماری اعمال سیستم هوشمند پیشنهادی بر روی افراد سالم و دارای اختلال گفتاری.

پردازش آماری	محدوده افراد سالم	محدوده افراد دارای اختلال	موارد تکرار آزمایش
میانگین	$\mu>2100$	$\mu<2000$	$<\mu<2100$
واریانس	$\sigma>50$	$\sigma<42$	$42<\sigma<50$
تعداد افراد	۴۶	۲۹	۱۲
درصد صحت تشخیص	٪۹۲	٪۹۶.۶۶	٪۱۵

در نهایت با استفاده از روش پیشنهادی و انجام تست بر روی افراد مختلف، اطلاعات آماری مربوط به تست ۷۵ مورد سالم و دارای اختلال به صورت زیر نتیجه شده است:

$$\text{True positive= Sensitivity=}\frac{29}{30}=0.96$$
$$\text{True negative = Specificity=}\frac{46}{50}=0.92=92\%$$
$$\text{False positive=}\frac{12}{80}=0.15=15\%$$
$$\text{False negative=}0.4$$

کمیت های فوق در ادبیات رشته آمار و اپیدمیولوژی به ترتیب حساسیت یا مثبت حقیقی، ویژگی یا منفی حقیقی، مثبت کاذب و منفی کاذب ترجمه می‌شوند. از آنجا که منفی کاذب در پژوهش حاضر بسیار پایین می‌باشد نشان‌دهنده این موضوع است که کمتر شخص دارای اختلالی به اشتباه سالم تشخیص داده شده است و به دلیل اهمیت این موضوع می‌توان انتظار داشت که افراد بیمار توسط سیستم شناسایی شوند.

www.SID.ir

Arshive of SID

بحث

صوت یکی از جنبه‌های گفتار است که در روند برقراری ارتباط نقش مهمی ایفا می‌کند و بخشی از مراجعه‌کنندگان به کلینیک‌های گفتار درمانی را مبتلایان به اختلال صوت تشکیل می‌دهند که آسیب شناس گفتار و زبان باید با بررسی ابعاد مختلف مشکل برنامه درمانی مناسب را برای آن‌ها طراحی کند با توجه به ماهیت پیچیده و چند بعدی صوت فرآیند ارزیابی اختلالات صوت باید به گونه‌ای باشد که ابعاد مختلف آن را در برگیرد (۱۳). اختلالات گفتاری می‌تواند ناشی از اختلالات صوتی باشد که اختلال صوتی نیز به دلیل وجود انواع بیماری‌ها اعم از عضوی یا عملکردی است. تاکنون درمانگران برای ارزیابی وضوح گفتار و تشخیص اختلال گفتاری متکی به روش‌های کیفی بوده اند. در صورتی که ارزیابی مبتنی بر داده های حاصله از اندازه‌گیری ابزاری می‌تواند معیار دقیق و مشخصی را در اختیار درمانگر و بیمار قرار دهد و تاکنون روش ها و تکالیف گفتاری متفاوتی برای ارزیابی گفتار مطرح شده است که این روش‌ها برای پایه آزمون است. فرد چندین جمله را تکرار کرده و سپس صدای ضبط شده وی توسط خبره و گفتار درمان مورد بررسی قرار می گیرد که این نیز درنهایت به صورت کیفی ارزیابی می شود(۱۴). Tian و همکاران (۱۵) تکنیک هوشمندی برای افزایش کیفیت سیگنال‌های صوتی در ارایه داده اند که الگوریتم کاهش نویز برای دریافت بهتر سیگنال می‌باشد که این الگوریتم‌ها می‌تواند در جهت بهبود تحلیل سیگنال صوتی مفید باشند اما به تنهایی نمی‌تواند سیستمی جامع برای تشخیص اختلال گفتاری باشد. Hu and Raj (۱۶) نیز در زمینه تحلیل هوشمند سیگنال های صوتی، مدلی برای تشخیص گفتار از روی حرکت لب ها ارایه دادند که این مدل با استفاده از دوربین پردازش لرزش لب ها را بررسی می‌کند ولی به دلیل اینکه در افراد دارای اختلال گفتاری، ممکن است سیگنال صوتی بدون حرکت لب ها ارسال شود این مدل نمی‌تواند به طور صحیح افراد دارای اختلالات گفتاری را شناسایی کند.

سرعت تشخیص در روش ارایه شده بسیار بالاتر از روش‌های پیشنهادی در مقالات معتبر علمی و نیز روش تشخیص انسانی توسط فرد است. ارتباط میان میزان تشخیص انسانی با تشخیص توسط سیستم هوشمند ($p = ۰/۰۰۱$) معنی‌دار بود.

نتیجه‌گیری

اختلال گفتاری یکی از شایع‌ترین اختلالات ظاهری در افراد جامعه است. از آنجا که تشخیص وجود اختلالات گفتاری، معمولا تنها بر اساس تشخیص متخصصان گفتار درمانی و بر اساس اصول شنیداری انجام می‌شود، لذا احتمال بروز خطای انسانی در تشخیص افراد بسیار بالاست. در این پژوهش مدلی ارایه شده است که به صورت هوشمند قادر به تفکیک افراد سالم و افراد دارای اختلال گفتاری استیک جمله شامل ۵ کلمه و ۸ هجا تعیین شد و از ۸۰ نفر نمونه‌گیری شد که به صورت تصادفی در میان افراد سالم و از گروه های سنی و جنسیت مختلفی انتخاب شده اند. سیستم پیشنهادی بر روی سیگنال‌های صوتی افراد مختلف تست شد و داده‌های آماری استخراج شدند. نتایج حاصل نشان دهنده تفکیک افراد با دقت بیش از ۹۲٪ بوده است که در مقایسه با نتایج سایر گزارشات نتیجه بسیار قابل قبولی می باشد. در آینده و در راستای این پژوهش می‌توان با تحلیل داده‌های آماری به تفکیک درجه‌بندی های مختلف اختلال گفتاری افراد نیز پرداخت.

ملاحظات اخلاقی

این مقاله شامل ملاحظات اخلاقی نمی شود.

منابع مالی

این مقاله منابع مالی ندارد.

منافع متقابل

مؤلف اظهار می دارد که منافع متقابلی از تالیف و یا انتشار این مقاله ندارد.

مشارکت مؤلفان

ا ع، ح م و همکاران، طراحی، اجرا و تحلیل نتایج مطالعه را بر عهده داشته اند و همچنین مقاله را تالیف نموده و نسخه نهایی آن را خوانده و تایید کرده اند.

References

1. Colton RH, Casper JK, Leonard R. *Understanding voice problem: a physiological perspective for diagnosis and treatment*. 3rd ed. Baltimore, Lippincott Williams &Wilkins, 2006; PP: 341-375.
2. Nicolosi L, Harryman E, Keresheck J. *Terminology of communication disorders speech-language-hearing*. 5th ed. Baltimore, Lippincott Williams & Wilkins, 2004; PP: 145-176.
3. Weismer G. *Motor Speech Disorders*. 1st ed. San Diego, Plural Pub, 2007; PP: 34-49.
4. Brookshire RH. Introduction to Neurogenic Communication Disorders. 6th ed. St Louis, Mosby–Year Book, Inc, 2003; PP: 54-86.
5. Love RJ. *Childhood motor speech disability*. 2nd ed. Boston, Allyn and Bacon, 2000; PP: 123-142.

www.SID.ir

Arshive of SID

6. Pascoe M, Stackhouse J, Wells B. *Persisting speech difficulties in children*. 3rd ed. Canada, John Wiley, 2006; **8**(3): 231-244.
7. Danahy E, Panetta K, Agaian S. Coordinate logic transforms and their use in the detection of edges within binary and grayscale images. *IEEE International Conference on Image Processing* 2007; **3**: 53-56. doi: 10.1109/icip.2007.4379244
8. Tsirikolias K, Mertzios B. Logic filters in image processing, *Proceedings of the International Conference on Digital Signal Processing* 1991; 285-287.
9. Papadopoulos A, Fotiadis D, Costaridou L. Improvement of micro calcification cluster detection in mammography utilizing image enhancement techniques. *Computers in Biology and Medicine* 2008; **3**8(10): 1045-1055. doi: 10.1016/j.compbiomed. 2008.07.006
10. Cheng H, Xiaopeng C, Xiaowei C, Liming H, Xueling L. Computer-aided detection and classification of micro calcifications in mammograms: a survey. *Pattern Recognition* 2003; **36**(12): 2967-2991. doi: 10.1016/S0031-3203(03)00192-4
11. Mertzios B, Tsirikolias K. Applications of coordinate logic filters in image analysis and pattern recognition. *The 2nd International Symposium on Image and Signal Processing and Analysis* 2001; 125-130.
12. Tsirilolias K, Mertzios B. Edge extraction and enhancement using coordinate logic filters Proceedings of the International Conference on Image Processing. *Theory and Applications* 1993; 251-254. doi: 10.1109/icdsp.1997.62806
13. Green M, Mathieson L. *The voice and its disorders*. London, Whurr Pub, 1991; PP: 319-332.
14. Heydari S, Torabi Nezhad F, Agha Rasouli Z, Hoseyni F. Development of speech intelligibility measurement test for 3 to 5 years old normal children. *Audiol* 2011; **20**(1): 47-53. (Persian).
15. Tian Y , Li S, Lv H, Wang J, Jing X. Smart radar sensor for speech detection and enhancement. *Sensors and Actuators A: Physical* 2013; **191**: 99-104.
16. Hu R, Raj B. A robust voice activity detector using an acoustic Doppler radar. *IEEE Workshop on Automatic Speech Recognition and Understanding* 2005; 319-324. doi: 10.1109/ASRU.2005.1566497

www.SID.ir