

## Detection of Hypernasal Speech for Children with Cleft Palate

E. Akafi <sup>1</sup>, M. Vali <sup>2\*</sup>, N. Moradi <sup>3</sup>

<sup>1</sup> M.S.C, Department of Engineering, Shahed University, Tehran, Iran, ehsan\_ek12@yahoo.com

<sup>2</sup> Assistant professor, Department of Electrical & computer Engineering, K. N. Toosi University of Technology, Tehran, Iran

<sup>3</sup> Assistant professor, Speech Therapy Department, Jondishapour University of Medical sciences, Ahvaz, Iran,

neginmoradist@gmail.com

### Abstract

Hypernasality is a frequently occurring resonance disorder in children with cleft palate. Generally an operation is necessary to reduce the hypernasality and therefore an assessment of hypernasality is imperative to quantify the effect of the surgery and design the speech therapy sessions which are crucial after surgery. In this study, a new quantitative method is proposed to estimate hypernasality. The proposed method used the fact that an Autoregressive (AR) model for vocal tract system of a patient with hypernasal speech is not accurate; because of the zeros appear in the frequency response of vocal tract system due to existence of extra channel between oral and nasal cavity of these patients. Therefore in our method hypernasality was estimated by a quantity calculated from comparing the distance between the sequences of cepstrum coefficients extracted from AR model and Autoregressive Moving Average (ARMA) model. K-means and Bayes theorem were utilized for finding a threshold value for proposed index to classify the utterances of subjects. We achieved the balanced accuracy up to 82.18% on utterances and 97.72% on subjects. Since the proposed method needs only computer processing of speech data, compare to other clinical methods it is provides a simple evaluation of hypernasality.

**Key words:** Cleft palate, Hypernasality, Speech processing, Speech therapy, Cepstrum.

\*Corresponding author

Address: Department of Electrical & computer Engineering, K. N. Toosi University of Technology, Tehran, Iran, P.O.Box: 16315-1355, Postal code:16314

Tel: +982188462174

Fax: +982188462066

E-mail: mansour.vali@eetd.kntu.ac.ir

## شناسایی گفتار پرخیشومی در کودکان دارای شکاف کام

احسان عکافی<sup>۱</sup>، منصور ولی<sup>۲\*</sup>، نگین مرادی<sup>۳</sup>

<sup>۱</sup>دانش آموخته کارشناسی ارشد، گروه مهندسی پزشکی، دانشکده مهندسی، دانشگاه شاهد، تهران ehsan\_ek12@yahoo.com

<sup>۲</sup>استادیار، گروه مهندسی پزشکی، دانشکده مهندسی برق و کامپیوتر، دانشگاه خواجه نصیرالدین طوسی، تهران

<sup>۳</sup>استادیار، گروه گفتاردرمانی، دانشکده توانبخشی، دانشگاه علوم پزشکی جندی شاپور اهواز neginmoradist@gmail.com

### چکیده

پرخیشومی از رایج‌ترین اختلالات در کودکان دارای شکاف کام است. عموماً برای کاهش این نقیصه نیاز به جراحی است و بنابراین ارزیابی خیشومی بودن برای بررسی تأثیر جراحی و همچنین طراحی جلسات گفتار درمانی- که بعد از عمل‌های جراحی نیاز است- حیاتی است. استفاده از مدل‌های تمام قطب مانند *AR* برای مدل‌سازی سیستم لوله صوتی افراد سالم رایج و معتبر هستند؛ اما وجود کانال ارتباطی بین حفره دماغی و دهانی افراد دارای شکاف کام، منجر به اضافه شدن صفر به تابع تبدیل فیلتر لوله صوتی شده و در نتیجه مدل مذکور برای فیلتر لوله صوتی این افراد دقیق نیست. بر این اساس در این تحقیق روش کمی جدیدی برای تخمین میزان پرخیشومی بودن ارائه شده است. در روش ارائه شده میزان پرخیشومی بودن با کمیتی که از محاسبه فاصله بین بردار ضرایب کیستروم استخراج شده از ضرایب مدل *AR* و مدل *ARMA* بدست آمده، ارزیابی شد. روش *k-means* و روش بیز برای یافتن حد آستانه مناسب بمنظور طبقه‌بندی دادگان به کار رفت. با اجرای الگوریتم پیشنهادی برای مجموعه دادگان شامل واژه‌های /a/ استخراج شده از کلمه آزمون /pamap/ که ۱۳ فرد دارای شکاف کام و ۲۲ فرد سالم آنرا بیان کردند، صحت تراز شده ۸۲/۱۸ درصد برای طبقه‌بندی گویش‌ها و صحت تراز شده ۹۷/۷۲ درصد برای طبقه‌بندی افراد بدست آمد. از آنجایی که روش ارائه شده تنها به پردازش کامپیوتری دادگان نیاز دارد، در مقایسه با روش‌های بالینی دیگر، ساده و غیر تهاجمی است.

کلیدواژگان: شکاف کام، پرخیشومی، پردازش گفتار، گفتار درمانی، کیستروم.

\*عهده‌دار مکاتبات

نشانی: تهران، خیابان شریعی، نرسیده به پل سیدخندان، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده برق و کامپیوتر، کدپستی: ۱۶۳۱۴

تلفن: ۰۲۱-۸۸۴۶۲۱۷۴، دورنگار: ۰۲۱-۸۸۴۶۲۰۶۶، پیام‌نگار: mansour.vali@eedt.kntu.ac.ir

## ۱- مقدمه

متخصص، استفاده از تجهیزات غیر تهاجمی بالینی - که میزان خروجی از حفره‌های دهان و دماغ را می‌سنجند- و استفاده از روش‌های مبتنی بر پردازش سیگنال گفتار است.

## ۱-۱-۱- قضاوت‌های ادراکی

قضاوت‌های ادراکی شامل نمره‌دهی افراد متخصص یا تعیین میزان قابلیت فهم گفتار شنوندگان آموزش دیده می‌شود. قضاوت‌های ادراکی برای ارزیابی خیشومی بودن گفتار به افراد آموزش دیده و باتجربه نیاز دارد که ممکن است این توانایی در بین افراد تغییر کند. برای حل این مشکل نیاز به روش‌هایی بدون نظر افراد است.

## ۱-۱-۲- روش‌های بالینی

به دلیل اینکه پرخیشومی بودن گفتار از اتصال نامناسب حفره دهان و بینی نشأت می‌گیرد، محققان برای بدست آوردن یک نسبت خروجی دماغی به دهانی برای ارزیابی خیشومی بودن تلاش می‌کنند. محققان از فشار، لرزش و جریان دماغی به عنوان کمیت‌هایی برای گفتار پرخیشومی استفاده کرده‌اند [۲، ۳]. شاخص اتصال دهانی - دماغی هاری ( $HONC^3$ ) بر همین اساس معرفی شده است [۴]. همچنین با تغییراتی در این شاخص، شاخص لرزش شتاب‌سنجی دماغی ( $NAVI^4$ ) نیز معرفی شده است [۲، ۵]. این کمیت‌ها به غیر از میکروفون به حسگری بر روی دماغ نیاز دارند. وسیله دیگری که در این زمینه به کار می‌رود نیزومتر<sup>۵</sup> است که یک نسبت دماغی به دهانی برای تخمین خیشومی بودن ارائه می‌کند. این وسیله از دو میکروفون جداگانه استفاده می‌کند به طوری که یکی در مقابل دهان و یکی در مقابل مجاری بینی قرار می‌گیرد و به وسیله صفحه‌ای پلاستیکی از یکدیگر جدا می‌شوند. نیزومتر به صورت گسترده در موارد بالینی استفاده می‌شود، ولی گران‌قیمت است و استفاده از آن راحت نیست؛ همچنین آن را برای دادگان ضبط شده نمی‌توان به کار برد.

از شایع‌ترین دلایلی که روند کسب مهارت‌های گفتاری را مختل می‌کند، اختلالات ساختمانی سیستم تولید گفتار است. نمونه رایج این نوع اختلالات مؤثر بر گفتار، شکاف کام و لب است. بیان این حقیقت که در دنیا از هر هفتصد کودک یک نفر دارای شکاف کام یا لب<sup>۱</sup> است، نشانگر گستردگی و فراوانی این ناهنجاری در دنیا است. در شکاف کام و لب عموماً غیر از کام و لب‌ها اندام‌های دیگر نیز درگیرند که وجود مشکل در هر یک از این اندام‌ها پیچیدگی مشکل گفتاری را بیشتر می‌کند. کودکان دارای شکاف کام به درمان‌های جراحی و دندانی برای تثبیت مکانیسمی که کفایت لازم برای تولید گفتار طبیعی را داشته باشد؛ نیاز دارند.

پرخیشومی<sup>۲</sup> از رایج‌ترین اختلالات تشدید در کودکان دارای شکاف کام است و فراوانی آن در افراد دارای شکاف کام فارسی زبان ۷۱ درصد است [۱]. پرخیشومی به دلیل اینکه حفره دهانی و حفره دماغی به خوبی از هم جدا نشده‌اند ایجاد می‌شود و با تشدید اضافی در تولید واژه‌های طبیعی همراه است. هر چند جراحی‌های ثانویه در بیشتر کودکان منجر به حذف تشدید پرخیشومی قابل شنیدن می‌شود، خیشومی شدگی در برخی کودکان به دلایل مختلف می‌تواند باقی بماند. ارزیابی میزان پرخیشومی بودن برای تصمیم‌گیری در مورد انجام عمل‌های جراحی و یا طراحی فرایند گفتار درمانی اهمیت دارد.

## ۱-۱-۱- روش‌های ارزیابی گفتار پرخیشومی

روش‌های ارزیابی خیشومی بودن به دو دسته کلی تهاجمی و غیر تهاجمی تقسیم می‌شوند؛ روش‌های تهاجمی شامل مشاهده و بررسی عملکرد دریاچه کامی - حلقی با استفاده از تجهیزات تصویربرداری بالینی است. این روش‌ها به حضور فرد در کلینیک و استفاده از تجهیزات گران قیمت نیاز دارند. همچنین نتایج حاصل از این روش‌ها کمی نیست و وابسته به نظر افراد است. روش‌های غیر تهاجمی شامل ارزیابی‌های ادراکی توسط افراد

<sup>1</sup> Cleft lip and / or palate<sup>5</sup> Nasometer<sup>2</sup> Hyponasality<sup>3</sup> Horii Oral Nasal Coupling<sup>4</sup> Nasal Accelerometric Vibration Index

## ۱-۱-۳- روش‌های مبتنی بر پردازش سیگنال گفتار

محققان پرخیشومی بودن گفتار را با تحلیل گفتار افراد دچار مشکلات شنوایی، شکاف کام و لب و مشکلات ناشی از ناتوانی در بچه‌کامی - حلقی و یا پردازش گفتار پرخیشومی که به صورت مصنوعی ایجاد شده و همچنین با بررسی حروف صدادار پرخیشومی شده گفتار معمولی، شناسایی می‌کنند. محققان نشان داده‌اند که برای واج‌های واکنش پرخیشومی، در نتیجه اتصال نامناسب حفره دماغی و دهانی یک جفت صفر و قطب اضافی در طیف سیگنال ایجاد می‌شود [۶]. تحلیل‌های فرکانسی نشان می‌دهد که مهم‌ترین ویژگی گفتار پرخیشومی، تغییر در نواحی فرکانس پایین طیف گفتار است که با ایجاد پیک‌هایی با فرکانس‌های خیلی کم و با پهنای باند وسیع همراه با ظهور یک زوج صفر و قطب مشاهده می‌شود [۷،۸]. اما برخی معتقدند علاوه بر ایجاد تغییرات در ناحیه فرمنت اول، افزایش تغییرات در نواحی فرکانس بالای طیف در اثر پرخیشومی شدن نیز قابل مشاهده است [۷،۴].

مطالعات محققان منبع [۹،۲] بر این فرض استوار است که گفتار معمولی از تشدیدهای متعدد در فرکانس‌های مختلف تشکیل شده است، در حالی که گفتار پرخیشومی علاوه بر این تشدیدهای مربوط به حفره دهانی شامل تشدیدهای حفره دماغی و آنتی‌رزونانس‌ها<sup>۶</sup> است. بنابراین اگر سیگنال گفتار معمولی با یک فیلتر پایین‌گذر با فرکانس قطع بین فرمنت اول و دوم فیلتر شود، سیگنال بدست آمده فقط شامل فرمنت اول خواهد بود؛ در حالی که سیگنال گفتار پرخیشومی پس از عبور از فیلتر پایین‌گذر علاوه بر فرمنت اول، شامل آنتی‌فرمنت‌ها و فرمنت‌های خیشومی نیز خواهد بود. در مطالعات مذکور از حساسیت عملگر انرژی تیگر<sup>۷</sup> به سیگنال‌های چند جزئی استفاده شده است. وقتی این عملگر بر سیگنال گفتار پرخیشومی عبور داده شده از فیلتر پایین‌گذر اعمال شود، به دلیل خاصیت غیر خطی آن یک جمله شامل ضرب آنتی فرمنت و فرمنت پرخیشومی ظاهر می‌شود. ولی زمانی که سیگنال گفتار از فیلتر

میان‌گذر عبور داده شود، فقط شامل فرمنت اول خواهد بود و در نتیجه حاصل اعمال عملگر تیگر بر این سیگنال با حالت قبل، متفاوت خواهد بود. مدل پیش‌بینی کننده خطی<sup>(LPC)</sup> به طور گسترده برای تحلیل سیگنال گفتار جهت مدل سازی فیلتر لوله صوتی به کار می‌رود [۱۰]. در گفتار عادی، اغلب از مدل LPC درجه ۸ تا ۱۰ استفاده می‌شود. چون در طیف سیگنال گفتار پرخیشومی تعدادی قطب اضافی و صفر وجود دارد و حضور صفرها در طیف را می‌توان با تعداد نامحدودی قطب بیان کرد، مدل‌های LPC با درجه بالاتر گفتار پرخیشومی را بهتر مدل می‌کنند. بر این اساس با استفاده از فاصله بین ضرایب LPC مرتبه پایین و مرتبه بالا، معیاری برای میزان خیشومی بودن گفتار می‌توان معرفی کرد [۱۱].

طیف مدل LP با درجه پایین در بیان دو فرمنت بسیار نزدیک اغلب ناتوان است. از طرف دیگر تحلیل‌ها با مدل‌های LP درجه‌های بالاتر تعداد پیک‌های بیشتری را در طیف حاصل نمایان می‌کند؛ حتی در گویندگان عادی به دلیل هارمونیک‌های فرکانس اصلی گفتار، تعدادی پیک در نواحی کمتر از فرمنت اول مشاهده می‌شود. در یک طیف اگر پیک مربوط به یک فرمنت حقیقی حذف شود، سیگنال بازسازی شده از روی این سیگنال در حوزه زمان تغییرات زیادی می‌کند. با توجه به اینکه پیک‌های مربوط به فرمنت‌های حقیقی مقادیر بیشتری در مقایسه با پیک‌های مربوط به هارمونیک‌های پیچ دارند، تکنیک اصلاح قطب بر پایه مدل LP برای ارزیابی پرخیشومی بودن گفتار معرفی شده است [۱۲،۱۳]. روش به این صورت است که از مدل LP با درجه بالا برای گفتار استفاده کرده؛ با شناسایی قطب مربوط به قوی‌ترین پیک در ناحیه فرکانس پایین، قطب مذکور تضعیف می‌شود. سپس سیگنال گفتار از روی مدل LP اصلاح شده بازسازی شده است و بیشینه مقدار همبستگی متقابل بین سیگنال اصلی و سیگنال بازسازی شده به عنوان معیاری برای شناسایی پرخیشومی بودن ارائه می‌شود.

<sup>6</sup> AntiResonances<sup>7</sup> Teager Energy Operator<sup>8</sup> Linear Predictive Coefficient

می‌شود، از این جمله‌اند [۱۶، ۱۷]. قابلیت اطمینان روش‌های کلاسیک بر پایه تحلیل آکوستیک و اغتشاش در مواردی که سیگنال، رفتار غیر عادی نشان می‌دهد و حالت تناوبی کمی دارد، کاهش می‌یابد. برای غلبه بر این مشکل، روش‌های مرتبط با دینامیک غیر خطی گفتار مانند نگاشت پوانکاره<sup>۱۶</sup>، بعد فرکتال، آنتروپی کلموگورف<sup>۱۷</sup>، بعد همبستگی و نمای لیاپانوف<sup>۱۸</sup>، برای تحلیل فعالیت‌های غیرعادی یا آشوبی استفاده شده است [۱۸]. در این مقاله در ادامه ابتدا روش‌های ارزیابی گفتار پرخیشومی بیان شده است. سپس روشی ابتکاری با استفاده از تکنیک‌های پردازش گفتار، برای ارزیابی میزان خیشومی بودن گفتار معرفی شده است. با بررسی نتایج حاصل از طبقه‌بندی با روش ارائه شده، پارامترهای بهینه برای این روش مشخص شده است. در پایان با تحلیل نتایج بدست آمده پروتکلی برای ارزیابی گفتار پرخیشومی ارائه شده است.

## ۲- دادگان استفاده شده و روش پیشنهادی

### ۲-۱- دادگان

در این پژوهش افراد مبتلا به شکاف کامل و دوطرفه کام با یا بدون شکاف لب همراه با بدی عملکرد دریاچه کامی - حلقی بررسی شدند. نمونه‌های مورد مطالعه از کودکان ۴-۱۲ سال و از هر دو جنس دختر و پسر انتخاب شدند. هیچ یک از افراد مورد مطالعه انحراف تیغه بینی نداشتند. همچنین در تمام این افراد لب ساختار مناسب را برای بیان تکالیف مورد نظر داشت. در تمام کودکان گروه بیمار مورد مطالعه، جراحی ثانویه انجام شده بود. از جمله شرایط شرکت در مطالعه برای افراد دارای شکاف کام این بود که پرخیشومی با شدت متوسط به بالا داشته باشند. به این منظور افراد دارای شکاف کام توسط دو آسیب‌شناس گفتار و زبان به طور جداگانه مورد ارزیابی ادراکی تشدید قرار گرفتند.

کودکان دارای شکاف کام از کلینیک‌های گفتار درمانی، بیمارستان کودکان مفید و مرکز طبی کودکان در شهر تهران و

در روشی دیگر از اختلاف بین دامنه فرمنت اول و دامنه پیک اضافی به عنوان کمی متناسب با میزان پرخیشومی بودن استفاده شده است [۴]. همچنین شاخص کمی نسبت آهنگ پایین به آهنگ بالای صدا ( $VLHR^9$ ) برای ارزیابی خیشومی بودن گفتار معرفی شده است [۱۴]. دسته‌ای دیگر از مقالات با استخراج ویژگی‌های آکوستیک و یا ویژگی‌های مرتبط با دینامیک غیر خطی گفتار و استفاده از روش‌های طبقه‌بندی مختلف مانند  $SVM^{10}$ ، به شناسایی و ارزیابی گفتار پرخیشومی پرداخته‌اند.

ویژگی‌های آکوستیک به کار رفته برای شناسایی گفتار پرخیشومی به سه دسته تقسیم می‌شوند. دسته اول پرخیشومی شدن گفتار را به صورت نویز جمع‌شونده با گفتار سالم مدل می‌کنند. از این دسته، موارد زیر را می‌توان برشمرد: ۱- نسبت هارمونیک به نویز ( $HNR^{11}$ ) که مربوط به مقدار میانگین مؤلفه‌های نویزی گفتار تولیدی است، ۲- نویز تحریک چاکنایی ( $GNE^{12}$ ) که به عنوان تخمینی از نویز بر اساس این فرض که پالس‌های چاکنایی ناشی از عبور هوا از میان تارهای صوتی بوده است و منجر به تحریک هم‌زمان باندهای فرکانسی مختلف می‌شوند و ۳- پیش‌بینی خطای بهنجار شده ( $NEP^{13}$ ) که به عنوان رابطه بین میانگین حسابی و هندسی بیان می‌شود. دسته دوم ویژگی‌های آکوستیک مربوط به نویز مدولاسیون فرکانسی هستند. از این دسته می‌توان فرکانس پایه گفتار و جیتر (میانگین درصد تغییرات دو مقدار متوالی پیچ) را نام برد. دسته سوم ویژگی‌های مرتبط با مدل‌های پارامتریک تولید گفتار هستند و از این دسته ضرایب کپستروم استخراج شده از ضرایب  $LP$ ، ضرایب  $MFCC$  و ضرایب تبدیل طیفی نسبی ( $RASTA^{14}$ ) را می‌توان نام برد [۱۵]. ناتوانی یا نارسایی دریاچه کامی - حلقی که بیماران دارای شکاف کام یا لب از آن رنج می‌برند، آن‌ها را وادار به تولید حرکت‌های جبرانی در لوله صوتی می‌کند. تولید توقف چاکنایی<sup>۱۵</sup> و مشکلات عمومی مرتبط با تولید گفتار چاکنایی که باعث ایجاد مشخصه‌های غیرخطی در حرکات لوله صوتی

<sup>9</sup> Voice Low Tone to High Tone Ratio

<sup>10</sup> Support Vector Machine

<sup>13</sup> Normalized Error Prediction

<sup>14</sup> Relative Spectral Transform

<sup>17</sup> Kolmogorov Entropy

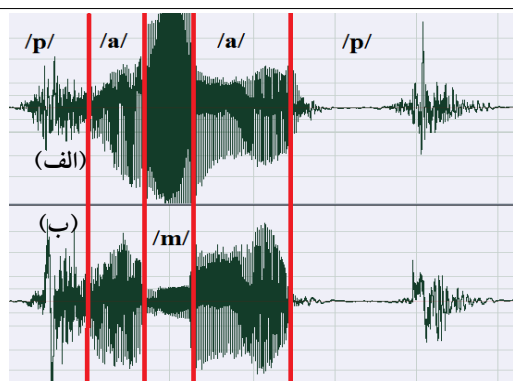
<sup>18</sup> Lyapunov Exponents

<sup>11</sup> Harmonic to Noise Ratio

<sup>12</sup> Glottal Noise excitation

<sup>15</sup> Glottal Stop

<sup>16</sup> Poincare



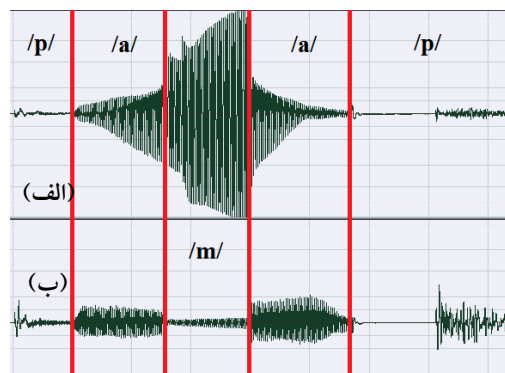
شکل (۲)- سیگنال گفتار یک فرد دارای شکاف کام  
(الف) سیگنال خروجی از دماغ (ب) سیگنال خروجی از دهان

همان‌طور که در شکل‌های (۱) و (۲) مشهود است، دامنه سیگنال خروجی دهان از دامنه سیگنال خروجی بینی کمتر است؛ چون میکروفون دهانی در فاصله ۳ سانتی‌متری دهان و میکروفون دیگر در تماس با حفره بینی واقع شده است. همان‌طور که انتظار می‌رود برای واکه‌های غیر خیشومی در گفتار فردی عادی، دامنه انرژی سیگنال ضبط شده با میکروفون دهانی از انرژی سیگنال ضبط شده از میکروفون بینی بیشتر است و بالعکس برای همخوان خیشومی /m/ انرژی و دامنه سیگنال ضبط شده از طریق میکروفون بینی بیشتر است. در حالی که در مورد سیگنال ضبط شده از یک فرد دارای شکاف کام در شکل (۲) این قضیه صادق نیست.

پس از ضبط دادگان، با استفاده از نرم افزار *Adobe Audition* از تمام تکرارهای هر فرد، هر دو واکه از سیگنال‌ها استخراج شد؛ در نهایت مجموعه دادگان شامل ۳۹۲ نمونه از افراد شامل ۲۴۶ گویش مربوط به ۲۲ فرد عادی و ۱۴۶ گویش از ۱۳ فرد دارای شکاف کام شد. به منظور به حداقل رساندن نویز محیط (کمتر از ۴۰ db) ضبط نمونه‌های گفتاری در محیط آکوستیک در کلینیک شنوایی شناسی دانشگاه تهران انجام شد. از دو میکروفون برای جمع‌آوری سیگنال‌های آکوستیکی از حفرات دهان و بینی به طور هم‌زمان، اما جداگانه استفاده شد. میکروفون دهان به حلقه‌ای وصل می‌شد و در فاصله ۳ سانتی‌متری سمت راست از دهان کودک قرار می‌گرفت. ابزار مورد

گروه طبیعی از مهدهای کودک و مدارس ابتدایی منطقه ۳ شهر تهران انتخاب شدند. افراد بدون شکاف کام از نظر سن و جنس با کودکان دارای شکاف مطابقت داده شدند. در مجموع ۳۵ نفر در این تحقیق شرکت کردند که ۱۳ نفر دارای شکاف کام بودند. در صورت وجود اثری از سرماخوردگی یا در صورت داشتن کم‌خیشومی و یا داشتن تشدید مختلط، افراد از مطالعه خارج شدند.

از تمام افراد شرکت‌کننده در مطالعه خواسته شد که بافت آوایی /pamap/ را چند بار ادا کنند. برای کنترل بلندی و سرعت گفتار مدل شنیداری از قبل ضبط شده بود و حداکثر سه مرتبه برای هر فرد پخش می‌شد. چون /p/ و /m/ قبل و بعد از واکه‌ها به حرکات باز و بسته شدن دریچه کامی-حلقی نیاز دارند، واکه‌های قرار گرفته بین این دو واج بهتر مشخصه‌های خیشومی شدگی را نشان می‌دهند [۱۹]. همچنین این بافت برای اندازه‌گیری ویژگی‌های زمانی خیشومی شدگی مفید است، زیرا هم صدای دهانی و هم صدای خیشومی در جایگاه لب‌ها تولید می‌شوند؛ بنابراین اثر تغییر جایگاه تولید بر تشدید خیشومی کنترل می‌شود [۲۰]. دو نمونه از سیگنال‌های صوتی ضبط شده مربوط به گروه کنترل و گروه بیمار در شکل‌های (۱) و (۲) نشان داده شده‌اند.



شکل (۱)- سیگنال گفتار یک فرد عادی  
(الف) سیگنال خروجی از دماغ (ب) سیگنال خروجی از دهان

هم‌پوشانی ۱۵ میلی‌ثانیه سیگنال پنجره‌گذاری شد. سپس با استفاده از یک الگوریتم جست‌وجو تکرارشونده -که خطای پیش‌بینی درجه دو را کمینه می‌کند- ضرایب  $ARMA^{22}$  و  $AR$  برای هر پنجره گفتار استخراج شدند. روش استفاده شده کمینه فاز بودن سیستم را تضمین می‌کند.  $d_i$  ها و  $p_i$  ها در رابطه (۲) ضرایب بدست آمده در این مرحله هستند؛  $n_p$  و  $n_z$  نیز به ترتیب تعداد صفرها و قطب‌ها در مدل هستند:

$$H(z) = \frac{1 + d_1 z^{-1} + \dots + d_{n_z} z^{-n_z}}{p_1 + p_2 z^{-1} + \dots + p_{n_p} z^{-n_p}} \quad (2)$$

برای داشتن دو دنباله هم‌طول و قابل مقایسه از ضرایب  $AR$  و  $ARMA$ ، از ضرایب کپستروم استفاده شد. برای بدست آوردن ضرایب کپستروم، ابتدا چند جمله‌ای حاصل شده در قسمت قبل را در قالب صفر و قطب رابطه (۳) می‌نویسیم؛

$$H(z) = \frac{K \prod_{i=1}^{n_z} (1 - a_i z^{-1})}{\prod_{i=1}^{n_p} (1 - b_i z^{-1})} \quad (3)$$

سپس با توجه به اینکه تمام صفرها و قطب‌ها درون دایره واحدند، ضرایب کپستروم از رابطه (۴)، یک بار با استفاده از ضرایب  $ARMA$  و یک بار با استفاده از ضرایب  $AR$  با در نظر گرفتن  $n_z$  برابر صفر بدست آمد.

$$c(m) = \begin{cases} \ln k & m = 0 \\ 0.5 \left( \sum_{i=1}^{n_p} \frac{b_i^m}{m} - \sum_{i=1}^{n_z} \frac{a_i^m}{m} \right) & m > 0 \\ 0 & m < 0 \end{cases} \quad (4)$$

برای حذف اثر پاسخ فرکانسی میکروفون، میانگین بردارهای ضرایب کپستروم بدست آمده از هر گویش از آنها کاسته می‌شود [۱۰]. حال اگر ضرایب کپستروم استخراج شده از ضرایب  $AR$  را  $c_{AR}(m)$  و ضرایب کپستروم استخراج شده از ضرایب  $ARMA$  را  $c_{ARMA}(m)$  بنامیم؛ شاخص  $DI$ ، با محاسبه فاصله هندسی

استفاده در این مطالعه، شامل میکروفون دهان و بینی مدل  $Shure-beta 54$  (ساخت آمریکا) و کارت صدای  $Onyx.Blackjack premium 2*2$  بودند. میکروفون‌های مورد استفاده، از نوع میکروفون فشرده<sup>۱۹</sup> الکتريکی بود که بر روی سر قرار می‌گرفت. علی‌رغم اندازه کوچک، فشرده‌گی میکروفون‌ها باعث ارائه طبیعی گفتار می‌شد. در ضمن این میکروفون‌ها به دلیل داشتن روکش خاص مانع ایجاد نویز حاصل از برخورد هوای خروجی حین گفتار می‌شدند.

## ۲-۲- روش

بیشتر تحقیقات قبلی در این حوزه، صرفاً به پردازش سیگنالهای دهانی پرداخته‌اند. واضح است اگر سیگنال دهانی به تنهایی در شناسایی خیشومی بودن گفتار نتایج رضایت بخشی داشته باشد کاربری این تحقیقات در آینده بسیار مناسب‌تر خواهد بود. لذا ما نیز در این تحقیق فقط از قسمت دهانی گویش‌های فوق‌الذکر استفاده کردیم و استفاده از سیگنالهای دماغی را در صورت نیاز به تحقیقات آتی موکول کردیم.

در اولین گام، میانگین کلیه سیگنالهای صوتی از آنها کسر شد تا  $dc$  سیگنال حذف شود و نیز برای خنثی کردن اثر شدت صدا و دوری و نزدیکی میکروفونها، کلیه سیگنالها نسبت به انحراف معیارشان به‌نجار شدند. سپس آن‌ها را از فیلتر پیش‌تأکید<sup>۲۰</sup> که تبدیل  $Z$  آن به صورت رابطه (۱) است، عبور دادیم. استفاده از این فیلتر باعث حذف اثر پراکندگی ناشی از عبور سیگنال از لب‌ها و پخش شدن در هوا می‌شود. همچنین این فیلتر اثر مؤلفه‌های طیفی چاکنایی را از بین می‌برد و کاهش مؤلفه‌های طیفی فرکانس بالا در اثر عبور از فیلتر لوله صوتی را جبران می‌کند [۲۱]:

$$p(z) = 1 - 0.98z^{-1} \quad (1)$$

سیگنال خروجی از مرحله قبل باید وارد مرحله پنجره‌گذاری شود. در این مرحله با استفاده از پنجره همینگ<sup>۲۱</sup> با طول ۳۰ میلی‌ثانیه که شرط ایستادن بودن سیگنال را تضمین می‌کند؛ و با

<sup>19</sup> Condensor

<sup>20</sup> Pre-emphasis

<sup>21</sup> Hamming

<sup>22</sup> Auto Regressive Moving Average

متمایزتر بودن نتایج بین دو گروه است. جدول (۱) به خوبی نشان می‌دهد که استفاده از الگوریتم دوم- که با هنجارسازی دنباله ضرایب کپستروم همراه بود- نتایج بهتری در پی داشته است. همچنین بهترین نتایج با استفاده از ۱۲۰ ضریب کپستروم و ۲ صفر در مدل  $ARMA$  بدست آمده است. پراکندگی مقادیر بدست آمده برای دو گروه افراد عادی و افراد دارای شکاف کام، با استفاده از نمودار جعبه‌ای<sup>۲۳</sup> برای دو حالت بهینه که در جدول (۱) پر رنگ شده‌اند، در شکل (۳) رسم شده است. همان‌طور که در این شکل نیز مشهود است، استفاده از ۲ صفر در مدل  $ARMA$  مقادیر متمایزتری برای دو گروه به همراه داشته است؛ از این رو برای ادامه تحلیل‌ها از روش مبتنی بر کپستروم بهنجار با به کارگیری مدل  $ARMA$  با ۲ صفر و پارامتر  $M$  برابر ۱۲۰، استفاده کردیم. چون برای هر فرد تعداد متفاوتی گویش داشتیم؛ برای بررسی عادی یا پرخیشومی بودن گفتار فرد، میانگین  $DI_{average}$  های مربوط به همه گویش‌های یک فرد را به عنوان معیار برای هر فرد در نظر گرفتیم. در این قسمت می‌توانستیم با توجه به اینکه اکثر گویش‌های فرد در کدام گروه طبقه‌بندی شده، در مورد او تصمیم بگیریم ولی چون هدف نهایی ما از این تحقیق رسیدن به معیاری کمی و متناسب با میزان پرخیشومی بودن گفتار بود، از میانگین  $DI_{average}$  های گویش‌های هر فرد استفاده شد.

بمنظور یافتن حد آستانه مناسب برای طبقه‌بندی دادگان از روش‌های  $k$ -means و بیز<sup>۲۴</sup> استفاده شد. طبقه‌بندی در دو سطح یعنی برای گویش‌ها و افراد به طور جداگانه انجام شد که نتایج آن در ادامه، گزارش شده است. روش  $k$ -means طبق الگوریتمی، دو خوشه برای دادگان می‌یابد و این خوشه‌ها را با مقادیر مراکزشان مشخص می‌کند. از آنجایی که داده‌ها تک‌بعدی هستند و این روش با محاسبه فاصله بین مراکز خوشه‌ها، تعلق دادگان به دو طبقه را مشخص می‌کند، با میانگین‌گیری از دو نقطه‌ای که در این روش به عنوان مرکز خوشه انتخاب شده‌اند، حد آستانه مناسب محاسبه می‌شود.

بین دنباله‌های کپسترال  $c_{AR}(m)$  و  $c_{ARMA}(m)$  از رابطه (۵) حاصل می‌شود:

$$DI = \sum_{m=0}^M [c_{ARMA}(m) - c_{AR}(m)]^2 \quad (5)$$

برای رسیدن به دقت مناسب، پارامتر  $M$  - که تعداد ضرایب کپستروم دخیل در محاسبه شاخص است- باید حداقل سه برابر مرتبه مدل  $AR$  در نظر گرفته شود [۲۲]. پس از محاسبه شاخص  $DI$  برای تمام پنجره‌های یک گویش، با محاسبه میانگین آن روی همه پنجره‌ها، شاخص  $DI_{average}$  محاسبه و معرفی شد. این شاخص برای تصمیم‌گیری در مورد عادی یا پرخیشومی بودن گویش‌های افراد به کار می‌رود. به نظر می‌رسد سنجش فاصله بین دو بردار ضرایب کپستروم باید مستقل از توان سیگنال صوت افراد باشد؛ لذا تصمیم گرفته شد که ضرایب کپستروم برای کل پنجره‌ها بهنجار شود و نتایج در این حالت با نتایج فاصله‌سنجی بدون هنجار سازی بردارهای کپستروم مقایسه شود.

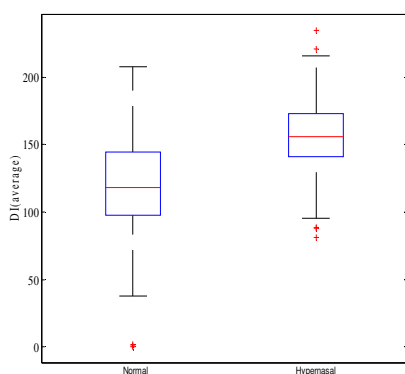
### ۳- بررسی و تحلیل نتایج

بمنظور یافتن تعداد صفر بهینه برای مدل  $ARMA$  و همچنین یافتن مقادیر مناسب برای تعداد ضرایب کپستروم (پارامتر  $M$  در رابطه (۵)) از آزمون  $t$  استفاده کردیم. برای این منظور الگوریتم پیشنهادی - که در قسمت قبل شرح داده شد- را برای مدل  $ARMA$  با ۱ تا ۵ صفر برای تمام گویش‌های عادی و پرخیشومی اجرا کردیم. چون مقدار  $M$  باید حداقل ۳ برابر تعداد قطب‌های در نظر گرفته شده برای مدل باشد [۲۲]، برای داشتن دقت بیشتر روش پیشنهادی را با در نظر گرفتن  $M$  برابر ۴۰ (۴ برابر تعداد قطب‌ها) و ۱۲۰ (با توجه به مطالعات قبلی [۱۱]) اجرا کردیم. سپس معنی‌دار بودن تفاوت بین نتایج بدست آمده برای هر دو گروه را با آزمون  $t$  سنجیدیم. مقادیر  $p$ -value حاصل از این آزمون در جدول (۱) آمده است. به طور معمول فرض صفر آزمون یعنی برابری میانگین دو گروه برای  $p$ -value های کمتر از ۰/۰۵ رد می‌شود و هر چه  $p$ -value کمتر باشد نشان‌دهنده

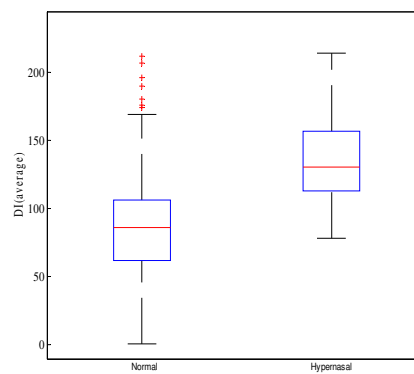


جدول (۱) - مقادیر  $p$ -value برای اجرای روش پیشنهادی با پارامترهای مختلف

پارامتر $M$	تعداد صفرهای مدل $ARMA$				
	۱	۲	۳	۴	۵
۴۰	۱	$2e-8.1$	۱	۱	$1e-9.7$
۱۲۰	۱	$3e-8.6$	$1e-8.6$	$4e-2.0$	$1e-9.7$
۴۰ (کیستروم بهنجار)	$1e-9.8$	$18e-1.3$	$9e-3.4$	$11e-7.6$	$10e-6.7$
۱۲۰ (کیستروم بهنجار)	$2e-8.0$	$38e-4.2$	$23e-1.3$	$21e-9.4$	$25e-4.8$



(ب)

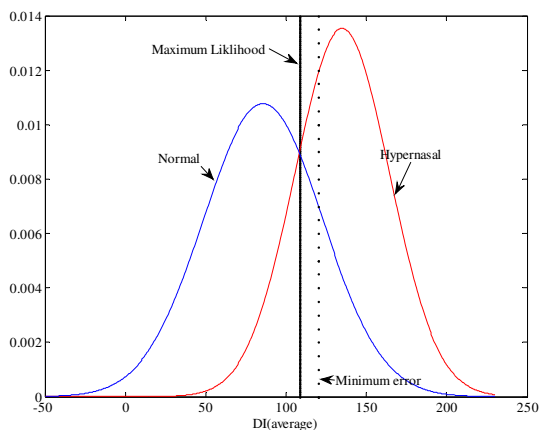


(الف)

شکل (۳) - پراکندگی شاخص استخراج شده از روش پیشنهادی با ۱۲۰ ضریب کیستروم بهنجار برای مدل  $ARMA$  با (الف) ۲ صفر (ب) ۵ صفر.

گوش‌های عادی، ۰/۶۲۷ و برای گوش‌های پرخیشومی، ۰/۳۷۲ است.

برای بدست آوردن حد آستانه مناسب با کمک روش بیز، هم از معیار درست‌نمایی ماکزیمم<sup>۲۵</sup> که احتمال حضور در دو گروه را برابر در نظر می‌گیرد و هم از رابطه بیز برای کمترین خطا که تفاوت تعداد دادگان دو گروه را در نظر می‌گیرد، استفاده شد. در واقع با تطبیق دو تابع گوسی بر دادگان آموزش دو گروه و یافتن نقطه تلاقی این دو تابع، حد آستانه مشخص و دادگان آزمون با این حد آستانه مقایسه شد. شکل (۴) حدود آستانه بدست آمده با به‌کارگیری دو روش مذکور را نشان می‌دهد. خط مرزی بدست آمده از رابطه بیز برای کمترین خطا - که در شکل (۴) با نقطه چین نشان داده شده است - نسبت به خط مرزی بدست آمده با معیار درست‌نمایی ماکزیمم، به سمت داده‌های مربوط به گوش‌های پرخیشومی که تعداد کمتری در مجموعه دادگان ما دارد؛ انتقال می‌یابد.



شکل (۴) - استفاده از روش بیز برای بدست آوردن حد آستانه مناسب برای مقادیر بدست آمده از روش پیشنهادی

با توجه به حجم دادگان دو گروه، احتمال قیاسی برای

<sup>25</sup> Maximum Likelihood

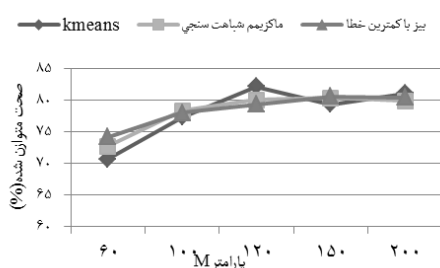
جدول (۲)- نتایج طبقه‌بندی گویش‌ها (بر حسب درصد)

صحت متوازن	صحت	اختصاصی بودن	حساسیت	روش
۸۲.۱۸	۸۱.۱۲	۷۸.۰۵	۸۶.۳۰	<i>K-means</i>
۸۰.۱۰	۷۹.۸۳	۸۰.۸۹	۷۸.۷۷	درست‌نمایی ماکزیمم
۷۹.۳۴	۷۷.۴۱	۸۴.۹۶	۶۹.۸۶	کمترین خطا

جدول (۳)- نتایج طبقه‌بندی افراد (بر حسب درصد)

صحت متوازن	صحت	اختصاصی بودن	حساسیت	روش
۹۵.۴۵	۹۴.۲۹	۹۰.۹۱	۱۰۰	<i>K-means</i>
۹۷.۷۲	۹۷.۱۴	۹۵.۵۴	۱۰۰	درست‌نمایی ماکزیمم
۹۷.۷۲	۹۷.۱۴	۹۵.۵۴	۱۰۰	کمترین خطا

برای بررسی بهینه بودن استفاده از ۱۲۰ ضریب کپستروم در روش ارائه شده، تغییرات صحت متوازن شده را برای هر بار اجرای روش با مقادیر مختلف پارامتر  $M$  در شکل‌های (۵) و (۶) برای هر سه روش یافتن حد آستانه، نمایش داده شده است. با توجه به هر دو نمودار- که برای طبقه‌بندی گویش‌ها و افراد رسم شده است- افزایش صحت متوازن برای مقادیر پارامتر  $M$  کمتر از ۱۲۰ قابل مشاهده است و برای مقادیر بزرگ‌تر از ۱۲۰ افزایش قابل توجهی دیده نمی‌شود حتی در برخی موارد کاهش صحت نیز اتفاق افتاده است. بنابراین با توجه به هزینه‌های محاسباتی استفاده از مقادیر بزرگتر پارامتر  $M$ ، استفاده از ۱۲۰ ضریب کپستروم در روش پیشنهادی تأیید می‌شود.



شکل (۵)- تغییرات صحت متوازن برای طبقه‌بندی گویش‌ها برای مقادیر مختلف پارامتر  $M$

بمنظور اعتبارسنجی روش‌های ارائه شده برای محاسبه حد آستانه و طبقه‌بندی گویش‌ها و افراد، از روش کنار گذاشتن یکی ( $LOO^{26}$ ) به صورت کنار گذاشتن یک فرد استفاده شد؛ به این صورت که در هر مرحله یک فرد و تمام گویش‌های کنار گذاشته شدند و حد آستانه با استفاده از بقیه دادگان بدست آمد. سپس داده‌های کنار گذاشته شده به عنوان داده‌های آزمون با حد آستانه بدست آمده مقایسه شدند و طبقه‌بندی انجام شد. این فرایند ۳۵ بار به تعداد کل افراد تکرار شد. مقادیر حساسیت<sup>۲۷</sup>، اختصاصی بودن<sup>۲۸</sup>، صحت و صحت تراز شده<sup>۲۹</sup>- که در واقع میانگین مجموع اختصاصی بودن و حساسیت است- در جداول (۲) و (۳) نمایش داده شده‌اند. شایان ذکر است که صحت تراز شده نابرابری دادگان را در محاسبه صحت لحاظ می‌کند. در صورت برابر بودن دادگان مربوط به دو گروه بیمار و سالم، مقادیر بدست آمده برای صحت و صحت تراز شده برابر می‌شود.

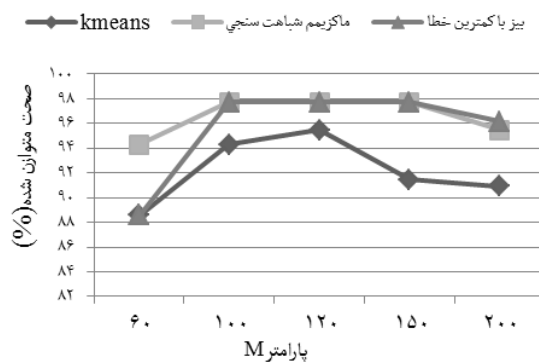
با بررسی جداول متوجه می‌شویم که روش بیز برای کمترین خطا با انتقال حد آستانه به سمت گویش‌های مربوط به افراد دارای شکاف کام، حساسیت را کاهش و اختصاصی بودن را افزایش داده است. البته این تغییر در سطح افراد مشاهده نمی‌شود. افزایش چشمگیر نتایج طبقه‌بندی در سطح افراد در مقایسه با نتایج در سطح گویش‌ها- که در جداول مشاهده می‌شود- می‌تواند برای طراحی پروتکلی برای ارزیابی میزان پریشومی بودن استفاده شود.

در بین مقادیر ارائه شده در جداول، میزان حساسیت و اختصاصی بودن، از آنجایی که هدف غربالگری و همانند آن نیست، نمی‌توانند ملاک تصمیم‌گیری باشند. چون حجم دادگان دو گروه برابر نیست، میزان صحت نیز می‌تواند در برخی موارد نادرست باشد. با توجه به اینکه هدف، بررسی تناسب معیار ارائه شده با میزان پریشومی بودن گفتار است و همچنین با توجه به نامتوازن بودن دادگان، صحت متوازن شده معیار مناسبی برای مقایسه است.

<sup>26</sup> Leave-One-Out<sup>27</sup> Sensitivity<sup>28</sup> Specificity<sup>29</sup> Balanced Accuracy

دارای شکاف کام و نارسایی دریاچه کامی - حلقی تأیید شد. با در نظر گرفتن حضور صفرها، مدل‌های تمام قطب متداول، نمی‌توانند به خوبی گفتار پرخیشومی را مدل کنند. بنابراین ما با در نظر گرفتن مدل *ARMA* برای تطبیق بهتر بر طیف گفتار پرخیشومی، با اندازه گیری فاصله بین بردار کپستروم استخراج شده از ضرایب *AR* و بردار کپستروم استخراج شده از ضرایب *ARMA* و مقایسه این فاصله در گفتار پرخیشومی و گفتار عادی، معیار کمی جدیدی معرفی کردیم. در واقع این فاصله برای گفتار پرخیشومی باید در مقایسه با گفتار عادی بیشتر باشد.

برای ارزیابی معیار ارائه شده، با قرار دادن حد آستانه مناسبی که با روش‌های خوشه‌بندی و بیز بدست آمد، به طبقه‌بندی گویش‌های افراد پرداختیم. همچنین مقایسه‌ای با روش مشابه مبتنی بر *LPC* ارائه شده در مطالعات گذشته، انجام شد. بهترین صحت تراز شده برای روش ما در سطح گویش‌ها (۸۲/۱۸ درصد) از بهترین صحت تراز شده روش مبتنی بر *LPC* (۷۴/۴۹ درصد) بیشتر بود. همچنین برای طبقه‌بندی افراد بهترین صحت تراز شده برای روش ما (۹۷/۷۲ درصد) از بهترین صحت تراز شده روش مبتنی بر *LPC* (۷۷/۱۰ درصد) بیشتر بود. با توجه به افزایش قابل توجه صحت در طبقه‌بندی افراد، می‌توان پروتکلی برای شناسایی گفتار پرخیشومی ارائه کرد؛ به این صورت که هر فرد باید کلمه آزمون (*pamap1*) را چندین بار تکرار کند، سپس شاخص معرفی شده در این تحقیق (*DIaverage*) برای هر یک از گویش‌های فرد محاسبه می‌شود و میانگین این شاخص بر روی کل گویش‌های فرد به عنوان معیاری کمی از گفتار او، با حد آستانه مناسب مقایسه می‌شود. در صورت بیشتر بودن این شاخص از حد آستانه، گفتار فرد مورد مطالعه، پرخیشومی طبقه‌بندی می‌شود.



شکل (۶) - تغییرات صحت متوازن برای طبقه‌بندی افراد برای مقادیر مختلف پارامتر M

برای ارزیابی کارایی روش پیشنهادی باید نتایج را با نتایج تحقیقات دیگر در این زمینه مقایسه کرد. از آنجایی که دسترسی به مجموعه دادگان استفاده شده در تحقیقات دیگران وجود ندارد و همچنین مقالاتی که میزان پرخیشومی بودن گفتار را بررسی کرده‌اند از روش محاسبه همبستگی بین معیار خود با نمرات دستگاه نیز کمتر یا آزمون‌های ادراکی بهره برده‌اند؛ لذا برای فراهم کردن معیار مناسبی جهت ارزیابی نتایج این تحقیق، الگوریتم ارائه شده در منبع [۱۱] اجرا شد. با توجه به مقاله مذکور پس از انجام پیش‌پردازش‌ها و پنجره‌گذاری ضرایب *LPC* با استفاده از الگوریتم بازگشتی لوینسون-دوربین برای مدل با تعداد قطب‌های ۱۰ و ۳۶، استخراج شدند. سپس فاصله هندسی بین ضرایب کپستروم استخراج شده از دنباله ضرایب *LPC* مرتبه بالا و پایین محاسبه شد. معیار کمی بدست آمده از این روش با انتخاب حد آستانه مناسب، برای طبقه‌بندی گویش‌ها و افراد به کار برده شد، که بالاترین صحت تراز شده ۷۴/۴۹ درصد برای گویش‌ها و ۷۷/۱۰ درصد برای افراد حاصل شد.

#### ۴- نتیجه گیری

در این تحقیق با بررسی ویژگی‌های آکوستیک سیگنال گفتار پرخیشومی کودکان دارای شکاف کام، بمنظور ارائه معیاری کمی برای میزان پرخیشومی بودن گفتار تلاش شد. در بررسی‌های انجام شده وجود صفر در تابع تبدیل فیلتر لوله صوتی افراد

- the analysis and detection of hypernasality; in TENCON 2009-2009 IEEE Region 10 Conference, 2009, pp. 1-5.
- [14] G. S. Lee, C. P. Wang, C. C. H. Yang, and T. B. J. Kuo, Voice low tone to high tone ratio: a potential quantitative index for vowel [a:] and its nasalization; *Biomed. Eng. IEEE Trans.*, 2006; 53(7): 1437-1439.
- [15] G. Castellanos, O. D. Castrillón, and E. Guijarro, Multivariate analysis techniques for effective feature selection in voice pathologies. *CASEIB*, 2004.
- [16] K. J. Golding-Kushner, *Therapy techniques for cleft palate speech and related disorders*. Singular San Diego, 2001.
- [17] A. Giovanni, M. Ouaknine, B. Guelfucci, P. Yu, M. Zanaret, and J. M. Triglia, Nonlinear behavior of vocal fold vibration: the role of coupling between the vocal folds; *J. Voice*, 1999; 13(4): 465-476.
- [18] J. J. Jiang, Y. Zhang, and C. McGilligan, Chaos in voice, from modeling to measurement; *J. Voice*, 2006; 20(1): 2-17.
- [۱۹] باغبان ک، ترابی نژاد ف، مرادی ن، بیگلریان الف، بررسی الگوی زمانی خیشومی شدگی در گفتار کودکان فارسی زبان ۴ تا ۱۲ ساله با و بدون شکاف کام، مجله پژوهش در علوم توانبخشی، دوره هشتم، شماره سوم، ۱۳۹۱.
- [20] S. Ha and D. P. Kuehn, Temporal Characteristics of Nasalization in Speakers with and Without Cleft Palate; *Cleft Palate. Craniofac. J.*, 2011; 48(2): 134-144.
- [21] A. Gray Jr and J. Markel, A spectral-flatness measure for studying the autocorrelation method of linear prediction of speech analysis; *Acoust. Speech Signal Process. Ieee Trans.*, 1974; 22(3): 207-217.
- [22] A. Gray Jr and J. Markel, Distance measures for speech processing; *Acoust. Speech Signal Process. Ieee Trans.*, 1976; 24(5): 380-391.
- از معیار و روش ارائه شده در این تحقیق می توان در طراحی سیستم های خودکار برای ارزیابی پرخیشومی بودن گفتار به منظور تصمیم گیری در مورد عمل های جراحی استفاده کرد. همچنین از این روش می توان در طراحی سیستم هایی خودکار یا نرم افزارهایی که امکان ارزیابی فرد از راه دور و خارج از محیط کلینیک را می دهند، برای کمک به فرایند گفتار درمانی بهره برد.
- ### ۵- مراجع
- [1] F. Derakhshandeh, M. Poorjavad, The Study of Speech Disorders and Middle Ear Diseases Following Primary Palatoplasty in Children with Cleft Palate; *J. Isfahan Med. Sch.*, 2011; 29 (130).
- [2] D. A. Cairns, J. H. L. Hansen, and J. E. Riski, A noninvasive technique for detecting hypernasal speech using a nonlinear operator; *Ieee Trans. Biomed. Eng.*, 1996; 43(1): 35-45.
- [3] "An accelerometric measure as a physical co... [J Speech Hear Res. 1983] - PubMed - NCBI."
- [4] M. Y. Chen, Acoustic parameters of nasalized vowels in hearing-impaired and normal-hearing speakers; *J. Acoust. Soc. Am.*, 1995.
- [5] M. A. Redenbaugh and A. R. Reich, Correspondence between an accelerometric nasal/voice amplitude ratio and listeners' direct magnitude estimations of hypernasality; *J. Speech Hear. Res.*, 1985; 28(2): 273.
- [6] G. Fant, *Acoustic theory of speech production*. Walter de Gruyter, 1970.
- [7] S. Hawkins and K. N. Stevens, Acoustic and perceptual correlates of the non-nasal-nasal distinction for vowels; *J Acoust Soc Am*, 1985; 77(4).
- [8] J. Glass and V. Zue, Detection of nasalized vowels in American English in Acoustics, Speech, and Signal Processing, *IEEE International Conference on ICASSP'85.*, 1985; 10: 1569-1572.
- [9] D. A. Cairns, J. H. L. Hansen, and J. F. Kaiser, Recent advances in hypernasal speech detection using the nonlinear teager energy operator in Spoken Language, 1996. *ICSLP 96.*, Fourth International Conference on, 1996; 2: 780-783.
- [10] L. R. Rabiner and R. W. Schafer, *Digital processing of speech signals*, Prentice-hall Englewood Cliffs; NJ, 1978; 100.
- [11] D. K. Rah, Y. I. KO, C. Lee, and D. W. Kim, A noninvasive estimation of hypernasality using a linear predictive model; *Ann. Biomed. Eng.*, 2001; 29(7): 587-594.
- [12] P. Vijayalakshmi, M. R. Reddy, and D. O'Shaughnessy, Acoustic analysis and detection of hypernasality using a group delay function; *Biomed. Eng. IEEE Trans.*, 2007; 54(4): 621-629.
- [13] P. Vijayalakshmi, T. Nagarajan, and V. Jayanthan Ra, Selective pole modification-based technique for