

شناسایی گوینده وابسته به متن با استفاده از الگوریتم خوشه‌بندی K-means و روش دسته‌بندی SVM

حسین خالقی بیزکی/مجتمع دانشگاهی برق و کامپیوتر/ دانشگاه صنعتی مالک اشتر / bizaki@gmail.com

مصطفی کاوانلویی/ کارشناس ارشد/ دانشکده مهندسی برق و الکترونیک/ دانشگاه آزاد اسلامی واحد دماوند

چکیده

هدف این مقاله توسعه یک سیستم شناسایی گوینده وابسته به متن با ارایه یک روش جدید در طبقه‌بندی گوینده می‌باشد. در این کار، یک سیستم شناسایی گوینده برای یک پایگاه داده ۱۰ نفره که در محیط عادی اتاق و با یک میکروفن معمولی هدست توسط نویسنده فراهم شده، با موفقیت آزمایش شده است. ویژگی مورد استفاده، ضرایب کپسترال مبتنی بر معیار مل (MFCC) بوده که از گفتار افراد استخراج شده و به همراه مشتقات اول و دوم بردار ویژگی ذخیره شده است. با استفاده از چندی‌سازی برداری (VQ) حجم داده‌ها کاهش یافته و از طبقه‌بندی‌کننده ماشین بردار پشتیبان (SVM) با هسته چند جمله‌ای و هسته توابع پایه شعاعی جهت طبقه‌بندی گوینده‌ها استفاده شده است. نتایج نشان می‌دهد که الگوریتم پیشنهادی در سیگنال‌های با قابلیت شناسایی قابل قبولی داشته اما با کاهش سیگنال به نویز، درصد شناسایی کاهش می‌یابد.

کلمات کلیدی: ضرایب کپسترال فرکانس مل، ماشین بردار پشتیبان، چندی‌سازی برداری، شناسایی گوینده وابسته به متن.

Text based speaker identification using K-means-clustering and SVM classification

H. Khaleghi Bizaki/ Department of Electrical and Computer Engineering/ Malek Ashtar University of Technology/ Tehran, Iran.

bizaki@gmail.com

M. Kavanloui/ MSc Student/ Science and Research Branch/ Islamic Azad University/ Damavand, Iran

ABSTRACT

The purpose of this paper is to develop a text-based speaker identification system by providing a new method for speaker classification. In this work, a speaker identification system has been successfully tested for a 10-person database provided by the authors in the normal room environment with a standard headset microphone. The Mel Frequency Cepstral Coefficient (MFCC) is used as a discriminating feature which is extracted from the speech of individual persons and stored with the first and second derivatives as a property vector. Using Vector Quantization (VQ), the data has been reduced and the Support Vector Machine (SVM) with the radial base functions and polynomial core to classify the speakers. The simulation results show that the proposed algorithm has acceptable detection capability in high Signal to Noise Ratio (SNR), but detection percentage decreases by decreasing the signal-to-noise ratio.

Keywords: Mel Frequency Cepstral Coefficient, backup vector machine, Support Vector machine, Vector Quantization, Text based Speaker Identification.

همچنین، یک سیستم تشخیص اتوماتیک گوینده مبتنی بر MFCC توسط ولیساوجویک [۴] ارائه شده است.

اکس یونگ در [۵] طراحی یک سیستم شناسایی گوینده مبتنی بر تکنیک MFCC با استفاده از ۱۲ ضریب را توصیف کرده است. چهار نسخه مختلف از سیستم با روش‌های مختلف استخراج ویژگی انجام شده‌اند. مولفان گزارش کرده‌اند که چهار روش آموزش مختلف کارایی‌های مشابه در دقت و درستی را نشان می‌دهند. این پیاده‌سازی شامل جمع‌آوری اطلاعات، حذف سکوت، پیش‌تاکیدسیگنال، محاسبه MFCC، شیفت سیگنال، آموزش، و ارزیابی می‌باشد.

در مرجع [۶] یک روش برای تشخیص گوینده بر اساس روش ترکیبی SVM و GMM ارائه شده است. در مرجع [۷] نیز کارایی روش‌های SVM و چند طبقه‌بندی‌کننده دیگر برای تشخیص گوینده بررسی شده و یک روش ترکیبی برای بهبود کیفیت تشخیص پیشنهاد شده است.

هدف از این مقاله، ارائه یک سیستم کارآمد برای شناسایی گوینده وابسته به متن با استفاده از روش‌های MFCC و Support Vector Machine (SVM) می‌باشد. پژوهش‌های گذشته نشان داده‌اند که MFCCها نمایانگر جزئیات خصوصیات فردی گوینده‌ها بوده و مقاوم هستند. از طرف دیگر، SVM یک طبقه‌بندی‌کننده بسیار قوی می‌باشد که می‌توان از آن برای تشخیص گوینده استفاده نمود. در ساختار پیشنهادی، مراحل زیر برای طراحی سیستم تشخیص گوینده در نظر گرفته شده است:

۱- اکتساب داده‌ها از طریق میکروفرن.

۲- استخراج ویژگی.

۳- فرسوده‌سازی داده‌ها با استفاده از الگوریتم Vector Quantization (VGLBG) Linde, Buzo and Gray

۴- تطبیق ویژگی.

با توجه به مراحل بالا، در بالاترین سطح، تمام سیستم‌های شناسایی گوینده دارای دو ماژول اصلی هستند: استخراج ویژگی و تطبیق ویژگی. در ادامه این مقاله، در ابتدا، استخراج ویژگی از صدای ورودی بر اساس روش MFCC بررسی شده سپس به ماشین بردار پشتیبان مورد استفاده برای تطبیق ویژگی پرداخته می‌شود.

۲- استخراج ویژگی و تطبیق ویژگی

۲-۱- استخراج ویژگی

واحد استخراج ویژگی، یکی از واحدهای مورد نیاز بیشتر کاربردهای بازشناسی الگو می‌باشد. هدف از ماژول استخراج ویژگی، کاهش حجم محاسبات و حذف افزونگی‌های موجود در سیگنال گفتار با استخراج تعداد محدودی پارامتر از آن و همچنین استخراج بردارهای ویژگی صوتی، که برای مشخص کردن خواص طیفی از زمان‌های مختلف سیگنال گفتار استفاده می‌شوند، می‌باشد. برخی از روش‌های پر کاربرد عبارتند از: کدگذاری پیش‌بینی خطی (Linear Prediction Coding (LPC)، ررایب کپسترال پیشگویی خطی (Linear Prediction Cepstral Coefficient (LPCC) و ضرایب کپسترال فرکانس-میل (MFCC) (مراجعه به جدول ۱). MFCCها بر اساس تغییرات شناخته شده‌ی پهنای باند بحرانی گوش انسان با فرکانس می‌باشد. تکنیک MFCC با استفاده از دو نوع فیلتر ایجاد می‌شود: فیلترهای به طول خطی و فیلترهای به طول لگاریتمی. همچنین MFCCها کمتر در معرض تغییرات شکل موج گفتار، با توجه با شرایط فیزیکی تارهای صوتی، بوده و بر همین اساس در این مقاله از این ویژگی استفاده شده است.

۲-۲- پردازنده MFCC

ضرایب کپسترال فرکانس-میل بهترین و شناخته شده‌ترین ویژگی

تمایل به ایجاد ارتباط گفتاری انسان با کامپیوتر و عدم امنیت شناسه‌های کاربری مختلف برای تعیین هویت افراد مانند: کلمات عبور و کارت‌های الکترونیکی و غیره در مقابل جعل، محققان را به یافتن شناسه‌های قابل اطمینان مانند شناسه‌های بیومتریک سوق داده است. شناسه‌های بیومتریک مانند: اثر انگشت، خصوصیات چهره، عنبیه چشم و صدا، کمتر قابل جعل هستند. بر این اساس، استخراج ویژگی‌های صوتی گوینده بر اساس متن موجود و سپس تشخیص فرد مورد نظر در بین افراد مختلف بر اساس ویژگی‌های استخراج شده از اهمیت ویژه‌ای برخوردار است. فرآیند شناسایی (تشخیص) یک گوینده به‌طور عموم به‌صورت اتوماتیک توسط ماشین و با استفاده از برخی خصوصیات صدای گوینده قابل انجام است.

تشخیص گوینده می‌تواند به دو قسمت: شناسایی گوینده و تصدیق هویت گوینده طبقه‌بندی شود. شناسایی گوینده فرآیند شناسایی گوینده از میان یک پایگاه داده بوده در حالی که تصدیق هویت گوینده فرآیند پذیرش یا رد کردن شناسایی یک گوینده می‌باشد. بسیاری از روش‌های تشخیص گوینده مانند تطبیق ساده الگو، رویکردهای پیچش زمانی پویا و رویکردهای تشخیص آماری الگو (از قبیل شبکه‌های عصبی و مدل‌های مخفی مارکوف)، مدل‌سازی مخلوط گوسی، پرسپترون‌های چند لایه، توابع شعاعی پایه و الگوریتم ژنتیک برای تشخیص گوینده استفاده شده‌اند.

از طرف دیگر، روش‌های تشخیص گوینده می‌توانند وابسته به متن و یا مستقل از متن باشند. در یک سیستم مستقل از متن، بدون در نظر گرفتن آنچه که قبلاً گفته شده، تشخیص گوینده انجام می‌شود در حالی که در سیستم وابسته به متن، گوینده‌ها بر اساس گفتن برخی عبارات خاص، مانند کلمات عبور، شماره کارت، پین کدها و غیره شناسایی می‌شوند. در مقاله حاضر، تشخیص گوینده به‌صورت یک تطبیق N:1 مدل می‌شود که در آن ویژگی‌های استخراج شده یک گوینده ناشناخته با تمام الگوهای موجود در مدل مرجع، برای پیدا کردن بیشترین شباهت، تطبیق داده می‌شوند. انتخاب ویژگی‌های گوینده در طراحی یک سیستم شناسایی گوینده مهم است. مجموعه ویژگی‌های گوینده باید تغییرپذیری میان-گوینده بالا و واریانس درون-گوینده پایینی داشته باشند. همچنین ویژگی‌های گوینده انتخاب شده به منظور کاهش افزونگی باید مستقل از یکدیگر باشند. پژوهش‌های زیادی در زمینه استخراج ویژگی‌های صوتی مناسب در خصوص سیستم‌های تشخیص گوینده انجام شده است. حسن و همکاران در [۱] یک سیستم شناسایی گوینده را با استفاده از ضرایب کپسترال فرکانس-میل (Mel Frequency Cepstral Coefficient (MFCC)) توصیف و توسعه دادند. مولفان برای آزمایش جهت نرخ شناسایی، از هر دو پنجره همینگ و مثلثی استفاده کرده و نشان دادند که نتایج با پنجره همینگ دارای کارایی بهتری می‌باشد. همچنین مولفان به‌منظور بدست آوردن نتایج رضایت‌بخش‌تر، افزایش تعداد مراکز ثقل متناسب با افزایش تعداد گویندگان را پیشنهاد کرده‌اند.

مودا و همکاران در [۲] یک سیستم تشخیص صوت با استفاده از MFCC و تکنیک پیچش زمانی پویا (Dynamic Time Warping (DTW)) توصیف کردند. نتایج نشان می‌دهد که هر دو تکنیک می‌توانند به‌طور موثر برای مقاصد تشخیص صوت استفاده شوند.

الیس در [۳] جزئیات یک پروژه با استفاده از متلب، که طراحی یک سیستم تشخیص گوینده می‌باشد، را ارائه داد. این سیستم بر تجزیه و تحلیل گام برای تشخیص یک گوینده مبتنی بوده و جزئیات مراحل لازم برای تجزیه و تحلیل گام را در حضور نویز ارائه می‌دهد.

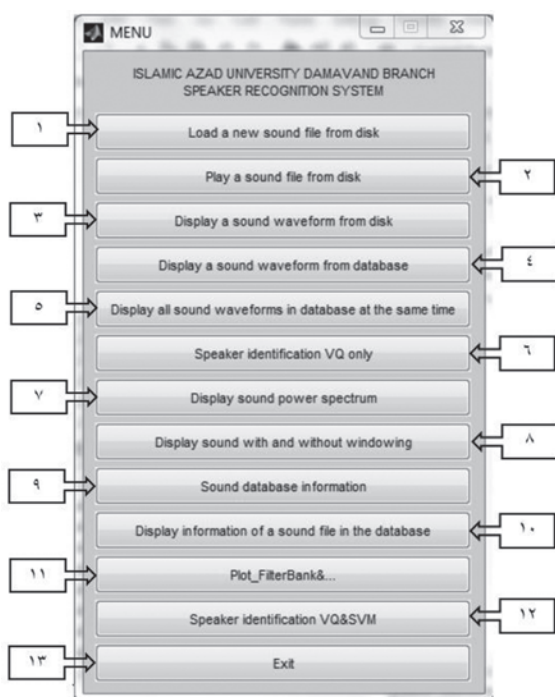
کردن یک فایل جدید صوتی به اندازه ۸ الی ۱۰ ثانیه را به کمک میکروفون به پایگاه داده و با در نظر گرفتن یک برچسب (ID) به عهده دارد.

گزینه ۲: پخش یک فایل صوتی از دیسک سخت که با کمک یک پنجره محاوره‌ای انجام می‌گیرد.

گزینه ۳: نمایش شکل موج صوتی از دیسک که به کاربر اجازه می‌دهد تا یک فایل صوتی را از جعبه محاوره انتخاب کرده و نمایش دهد.

گزینه ۴: نمایش شکل موج فایل صوتی از پایگاه داده که مشابه گزینه ۳ بوده با این تفاوت که فایل صوتی برای نمایش از صداهای پایگاه داده انتخاب می‌شود نه از روی دیسک سخت.

گزینه ۵: نمایش هم‌زمان تمام شکل موج‌های صوتی موجود در پایگاه داده.



شکل ۳: GUI سیستم شناسایی گوینده

گزینه ۶: با انتخاب این گزینه تشخیص گوینده به روش VQ انجام می‌شود. این گزینه اصلی‌ترین گزینه رابط کاربری می‌باشد که فرآیند پیاده‌سازی تشخیص گوینده را انجام می‌دهد. زمانی که کاربر این گزینه را انتخاب می‌کند یک کادر محاوره‌ای مانند شکل (۴) نمایش داده می‌شود، سیستم بعد از اجرای برنامه به کاربر دستور ورود یک فایل صوتی را برای شناسایی می‌دهد. الگوریتم MFCC در این گزینه اجرا شده و ویژگی‌های فایل صوتی را استخراج می‌کند. سپس این ویژگی‌ها توسط الگوریتم تطبیق ویژگی کوانتیزاسیون برداری (VQ) مقایسه شده و خروجی تطبیق و شناسایی گوینده در برنامه نشان داده می‌شود.

گزینه ۷: این گزینه طیف توان فایل صوتی که توسط کاربر انتخاب شده را به صورت خطی و لگاریتمی نمایش می‌دهد (شکل ۵).

گزینه ۸: نمایش صوت بدون پنجره‌گذاری و با پنجره‌گذاری را انجام می‌دهد که این کار به منظور جلوگیری از مساله ناشی از برش سیگنال انجام می‌شود.

گزینه ۹: جزئیات فایل‌های صوتی ذخیره شده در ساختمان داده را نمایش می‌دهد.

دلیل سهولت پیاده‌سازی و دقت بالا، استفاده شده است. روش VQ، فرآیند نقشه‌برداری بردارها از یک فضای برداری بزرگ به یک تعداد محدودی از نواحی را انجام می‌دهد. هر ناحیه با یک خوشه و مرکز آن یک کلمه کد نامیده شده است. جمع‌آوری تمام کلمات کد یک کتاب کد نامیده می‌شود [۸].

۳- تشخیص گوینده با استفاده از الگوریتم VQ

فرآیند تشخیص گوینده در دو مرحله اصلی انجام می‌شود:

- استخراج ویژگی

- تطبیق ویژگی.

۱-۳- استخراج ویژگی

هدف در استخراج ویژگی، استخراج پارامترهای اساسی برای شناسایی کردن یک سیگنال گفتار است. اگرچه چندین روش مختلف برای این کار وجود دارد، اما در این مقاله از روش MFCC استفاده شده است که خود شامل مراحل زیر می‌باشد:

الف- قاب‌بندی (تقسیم سیگنال به قاب‌ها).

ب- پنجره‌گذاری (بدست آوردن طیف دامنه برای هر قاب).

ج- تبدیل فوریه سریع (استخراج مولفه‌های فرکانسی).

د- پیچش فرکانسی میل (تبدیل به طیف مل).

ه- کپستروم (انجام تبدیل کسینوسی گسسته برای بدست آوردن ضرایب کپستروم).

۲-۳- تطبیق ویژگی

به منظور تطبیق ویژگی از کوانتیزاسیون برداری (VQ) استفاده شده است که خود شامل مراحل زیر می‌باشد:

الف- کوانتیزاسیون برداری

در این روش، در ابتدا، به منظور مشخص کردن ویژگی‌های خاص گوینده، به استخراج تعداد کمی از بردارهای ویژگی پرداخته و سپس بر اساس ویژگی‌های استخراج شده به کمک خوشه‌بندی، یک کتاب کد برای هر گوینده تشکیل می‌شود. در مرحله تشخیص، ویژگی‌های داده‌های گوینده‌ی مورد آزمایش با کتاب کد مقایسه شده و از روی میزان این اختلاف فاصله، به منظور تشخیص گوینده استفاده می‌شود. برای خوشه‌بندی از الگوریتم k-means، بر اساس کمینه کردن واریانس درون خوشه‌ها، استفاده شده است.

ب- اندازه‌گیری فاصله

به منظور اندازه‌گیری اختلاف ویژگی صوتی کاربر مورد نظر با کتاب کد از معیار کمینه فاصله اقلیدسی استفاده شده است.

۴- تشخیص گوینده با استفاده از الگوریتم SVM-VQ

در روش VQ، دسته‌بندی با استفاده از فواصل مراکز ثقل بردار ویژگی هر فرد و مقایسه آن با کل پایگاه داده، عمل شناسایی گوینده انجام می‌گیرد. به منظور بهبود الگوریتم از روش چندی‌سازی برداری جهت کاهش محاسبات استفاده شده و عملیات شناسایی گوینده در دو بخش آموزش و آزمایش به عهده ماشین بردار پشتیبان انجام شده است.

۵- شبیه‌سازی و نتایج آن

به منظور تشکیل پایگاه داده و اجرای سریع برنامه، یک رابط کاربری Graphical User Interface (GUI) به صورت شکل (۳) در نظر گرفته شده است. این رابط کاربری دارای ۱۳ گزینه است که عملکرد آنها به اختصار عبارتند از:

گزینه ۱: خواندن یک فایل صوتی جدید از حافظه که وظیفه اضافه

جدول ۲: اثر انداز کتاب کد و نویز بر روی نرخ شناسایی گوینده

تعداد مراکز ثقل (k)	نرخ شناسایی (بدون نویز)	نرخ شناسایی (واریانس نویز ۰.۰۰۱)	نرخ شناسایی (واریانس نویز ۰.۱)
۱	۸۰٪	۸۰٪	۶۰٪
۲	۸۰٪	۸۰٪	۳۰٪
۴	۱۰۰٪	۱۰۰٪	۵۰٪
۸	۱۰۰٪	۱۰۰٪	۷۰٪
۱۶	۱۰۰٪	۱۰۰٪	۱۰۰٪
۳۲	۱۰۰٪	۱۰۰٪	۱۰۰٪
۶۴	۱۰۰٪	۱۰۰٪	۱۰۰٪

۲-۵- آزمایش شناسایی گوینده با SVM&VQ

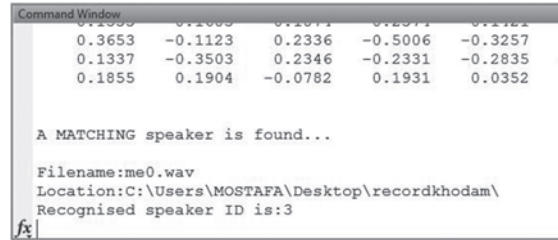
علاوه بر روش دسته‌بندی VQ، از روش دسته‌بندی ماشین بردار پشتیبان نیز شناسایی گوینده انجام شده است با این تفاوت که در مرحله قبل فقط با استفاده از فواصل مراکز ثقل بردار ویژگی هر فرد و مقایسه آن با کل پایگاه داده، فرد مورد نظر شناسایی می‌شد در حالی که در این قسمت، از روش چندی‌سازی برداری جهت کاهش حجم اطلاعات و همچنین کاهش محاسبات استفاده شده و عملیات شناسایی گوینده در دو بخش آموزش و آزمایش به عهده ماشین بردار پشتیبان انجام می‌گیرد.

ماشین بردار پشتیبان دارای سه نوع هسته متداول می‌باشد که با توجه به انتخاب هسته‌ها، طریقه دسته‌بندی متفاوتی داشته و در نتیجه نتایج حاصل از آن متفاوت خواهد بود. در این مقاله با استفاده از دو هسته MLP (Multilayer Perceptron) و RBF (Radial Basis Function)، نرخ شناسایی گوینده و میزان تطبیق داده‌های گوینده هدف با دیگر گوینده‌ها بررسی شده است. نرخ شناسایی و میزان تطبیق داده‌ها با انتخاب حجم کامل ضرایب کپسترال، هسته RBF و نوع خوشه‌بندی k-means و k-meanLBG با SNR=100 (تقریباً بدون نویز) و با مراکز ثقل ۴، ۸ و ۱۶ (به‌منظور شناسایی گوینده اول) در جدول (۳) نشان داده شده است. از این جدول مشاهده می‌شود که با افزایش تعداد مراکز ثقل، نرخ شناسایی کاهش می‌یابد.

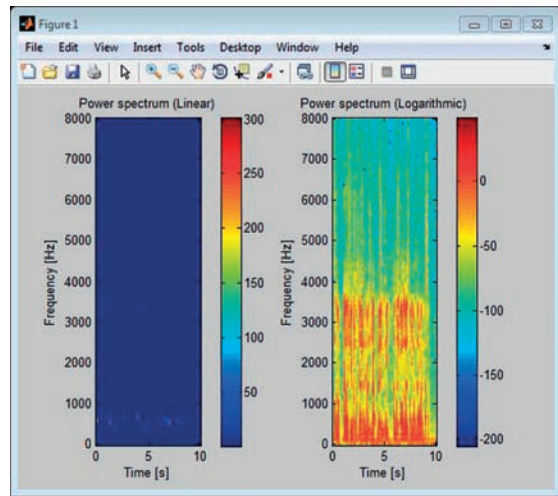
همچنین، اثر تغییرات میزان ضرایب کپسترال بر نرخ شناسایی، به ازای هسته RBF و خوشه‌بندی k-meanLBG مرکز ثقل ۸ و SNR=100 بررسی شده و در جدول (۴) نشان داده شده است. از این جدول مشاهده می‌شود که با کاهش تعداد ضرایب، نرخ شناسایی افزایش می‌یابد. و بالاخره، شناسایی گوینده در حضور نویز گوسی با هسته RBF، خوشه‌بندی k-meanLBG با مرکز ثقل ۸، تعداد ضرایب ۲۰۰۰ و میزان سیگنال به نویز (SNR) متفاوت در جدول (۵) نشان داده شده است. نتایج نشان می‌دهد که با کاهش SNR، نرخ شناسایی کاهش می‌یابد.

۶- نتیجه‌گیری

در این مقاله، یک سیستم مبتنی بر متلب برای تشخیص گوینده توسعه داده شد به گونه‌ای که قادر به تشخیص گوینده مبتنی بر پایگاه داده‌اش می‌باشد. نتایج شبیه‌سازی نشان می‌دهد که الگوریتم پیشنهادی در سیگنال به نویز های بالا دارای قابلیت شناسایی قابل قبولی می‌باشد اما با کاهش سیگنال به نویز، درصد شناسایی کاهش می‌یابد.



شکل ۴: کادر محاوره‌ای تشخیص گوینده



شکل ۵: نمایش طیف توان یک سیگنال صوتی

گزینه ۱۰: وظیفه نمایش اطلاعات یک فایل صوتی در پایگاه داده را به عهده دارد.

گزینه ۱۱: رسم فیلتر بانک مل و سیگنال گفتار و ضرایب کپسترال و طیف توان را انجام می‌دهد (مشابه شکل ۲).

گزینه ۱۲: شناسایی گوینده به روش SVM-VQ و رسم نتایج برای هر گوینده را به صورت یک نمودار میله‌ای انجام می‌دهد.

گزینه ۱۳: وظیفه خروج از برنامه را برعهده دارد.

به‌منظور بررسی سیستم تشخیص گوینده، در ابتدا، به ضبط گفتار و جمع‌آوری فایل صوتی از گوینده‌ها اقدام کرده و با یک میکروفن ساده و منویی که در GUI متلب جهت ضبط صدا بیان شد، پایگاه داده را فراهم نمودیم. برای تشکیل پایگاه، مدت زمان گفتار هر فرد به طور تقریبی ۸ تا ۱۰ ثانیه، تعداد افراد برابر با ۱۰ نفر شامل ۵ زن و ۵ مرد در نظر گرفته شده است. به دلیل اینکه صدای افراد و طریقه بیان جملات در شرایط مختلف کمی تفاوت دارد سعی شده جهت بررسی سیستم در شرایط مختلف، گفتار افراد در ده نسخه مختلف ضبط شده است (تشکیل یک پایگاه ۱۰۰ تایی). با این کار پایداری سیستم نسبت به شرایط، لهجه و تن صدای افراد، محک زده می‌شود.

۱-۵- آزمایش سیستم شناسایی گوینده با VQ

اثر تعداد مراکز ثقل (اندازه کتاب کد) مورد استفاده در الگوریتم و اثر آن بر نرخ شناسایی گوینده (درصدی از تعداد گوینده‌های شناسایی شده به کل گوینده‌های موجود در پایگاه داده) در جدول (۲) نشان داده شده است. نتایج نشان می‌دهد که زمانی که اگر تعداد مراکز ثقل بیش از ۴ و نویز کمتر از ۰.۰۰۱ باشد، نرخ شناسایی ۱۰۰ درصد حاصل می‌شود، یعنی تمام گوینده‌های پایگاه داده به درستی شناسایی می‌شوند در حالی که با افزایش نویز باید تعداد مراکز ثقل را زیاد کرد تا نرخ شناسایی ۱۰۰٪ باقی بماند.

کنترل مقاوم امیدانس بازوی ربات تک رابط انعطاف پذیر در محیطی با سختی نامشخص

علی فیاضی
استاتید راهنما: دکتر ناصر پریز، دکتر علی کریم پور
دانشگاه فردوسی مشهد

ربات‌های انعطاف پذیر سبک به علت پاسخ سریع و کار در سرعت بالا در مقایسه با ربات‌های صلب مرسوم به طور گسترده‌ای در کاربردهای مختلف از جمله هوافضا، پزشکی، صنایع و غیره مورد استفاده قرار می‌گیرند. اما به علت قابلیت انعطاف ساختاری گسترده، مدل‌سازی و کنترل ربات‌های انعطاف پذیر پیچیده‌تر است. به علاوه برخی از پیچیدگی‌های کنترلی در این دست از ربات‌ها وجود دارد که از آن جمله می‌توان به مشخصه غیرحداقل فاز، رفتار زیر تحریک، مشخصه غیرخطی تغییر پذیر با زمان و پارامتر گسترده اشاره کرد.

رهیافت‌های کنترلی مختلفی در گذشته معرفی شده است، اما بیشتر این استراتژی‌های کنترلی به کنترل موقعیت بازوی ربات تک رابط انعطاف پذیر پرداخته شده است و به اثر متقابل و کنترل نیرو در برخورد با محیط توجه چندانی نشده است. علاوه بر این در تمام مطالعات انجام شده در کنترل موقعیت و نیرو تک رابط انعطاف پذیر، نقطه برخورد با محیط نقطه انتهایی رابط در نظر گرفته شده است و تا آنجایی که نویسنده اطلاع دارد، مفهوم کنترل مقاوم امیدانس ربات تک رابط انعطاف پذیر در محیط نامشخص با برخورد در نقطه نامعلوم میانی، به نسبت جدید و برای نخستین بار به آن پرداخته شده است. کنترل امیدانس دارای یک ویژگی جامع است که آن را در کنترل حرکات تعاملی کارآمدتر می‌کند. در حقیقت، کنترل امیدانس با کنترل رفتار دینامیکی بین بازو و محیط به طور غیرمستقیم وظیفه کنترل نیرو را انجام می‌دهد. هدف نهایی این پژوهش اعمال این روش کنترلی به بازوی ربات تک رابط انعطاف پذیر در تعامل با محیط نامعلوم با برخورد در نقطه انتهایی و میانی بازو است. در رویکرد جدید پیشنهادی، کنترل کننده به صورت خودکار از مود کاری حرکت آزاد به مود کاری مقید تغییر وضعیت می‌دهد. بنابراین نیازی به الگوریتمی جداگانه برای تشخیص برخورد رابط با محیط نمی‌باشد. در این راستا، کنترل کننده امیدانس با حلقه‌ی درونی موقعیت پیشنهاد شده است. بدین معنی که در حرکت آزاد نیروی اعمالی به محیط صفر است و زاویه‌ی مرجع برای حلقه‌ی درونی موقعیت همان موقعیت مطلوب است. در حرکت مقید، مسیر مرجع برای حلقه‌ی درونی با استفاده از دینامیک امیدانس مطلوب تعیین می‌شود. استراتژی کنترل پیشنهادی، در برابر تغییرات پارامترهای محیط (نظیر سختی و ضریب میرایی)، اغتشاش نامعلوم اصطکاک کولمبی، تغییرات جرم رابط و اصطکاک ویسکوز موتور مقاوم است. همچنین، روش پیشنهادی برای هر دو حرکت آزاد و مقید معتبر می‌باشد. بخش اصلی طراحی کنترل کننده در استراتژی کنترل امیدانس پیشنهادی، کنترل کننده موقعیت حلقه‌ی داخلی است. استراتژی کنترل موقعیت حلقه‌ی داخلی بر اساس یک کنترل کننده مود لغزشی با سطح لغزش مرتبه کسری تعمیم یافته مجهز به رویکرد اغتشاش نامعلوم می‌باشد. کارایی و اثربخشی طرح کنترل پیشنهادی با استفاده از شبیه‌سازی عددی نشان داده شده است.

کلمات کلیدی:

کنترل مقاوم امیدانس، کنترل مود لغزشی مرتبه کسری، رویکرد اغتشاش نامعلوم، بازوی ربات تک رابط انعطاف پذیر، محیط نامشخص.

جدول ۳: نرخ شناسایی گوینده به ازای خوشه بندی های مختلف

خوشه بندی	مرکز ثقل	درصد تشخیص
k-means	۴	۷۱.۳۸
	۸	۶۱.۳۵
	۱۶	۳۷.۱۲
k-means LBG	۴	۷۰.۳۹
	۸	۶۱.۹۵
	۱۶	۳۷.۱۹

جدول ۴: نرخ شناسایی گوینده به ازای تعداد ضرایب مختلف

تعداد ضرایب	۶۰۰	۸۰۰	۱۰۰۰	۲۰۰۰	۴۰۰۰	۶۰۰۰
نرخ شناسایی	۶۳.۰۹	۶۴.۵۸	۶۸.۷۴	۷۳	۷۴.۸۶	۷۷.۴۵
تعداد ضرایب	۴۰	۲۰	۱۰	۵	۲	۱
نرخ شناسایی	۸۴.۸۸	۸۹.۸۲	۹۶.۷۹	۹۸.۴۶	۹۸.۷۸	۹۸.۷۸

جدول ۵: نرخ شناسایی گوینده اول به ازای SNRهای مختلف

SNR (dB)	۳۰	۴۰	۶۰	۸۰	۱۰۰
درصد تشخیص	۵۶.۲۶	۶۳.۵۲	۶۹.۸۵	۷۲.۱۹	۷۳
SNR (dB)	۲۰	۲۱	۲۲	۲۵	۲۵
درصد تشخیص	۴۹.۶۹	۵۲.۱۲	۵۴.۹۵	۵۷.۹۹	۵۷.۹۹

سپاسگزاری

بر خود لازم می‌دانیم که از زحمات جناب آقای دکتر مجتبی لطفی‌زاد کمال تشکر را داشته باشیم.

مراجع

- [1] Hasan, R., Jamil, M., Rabbani, G., and Rahman, S., "Speaker Identification Using Mel Frequency Cepstral Coefficients", 3rd International Conference on electrical & Computer Engineering, ICECE 2004, Dhaka, Bangladesh.
- [2] Muda L., Begam M and Elamvazuthi I., "Voice Recognition Algorithm using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", Journal of Computing, vol. 2, Issue 3, March 2010.
- [3] Ellis., E. D. "Design of a Speaker recognition Code using MATLAB", Department of Computer and Electrical Engineering, University of Tennessee, Knoxville, 2001.
- [4] Velisavljevic., V. "An automatic speaker recognition system", Mini-Project, Ecole Polytechnique Federale De Lausanne, 2003.
- [5] Xiong, Y. "Speaker Identification System using HMM and Mel Frequency Cepstral Coefficient. Pattern Recognition and Clustering (May 2006)
- [6] H. Bourouba, C.A. Korba, and Rafik Djemili, "Novel Approach in Speaker Identification using SVM and GMM", Control Engineering and Applied Informatics, Vol.15, No.3 pp. 87-95, 2013.
- [7] Yussouf Nahayo, Seçkin Arı, "Performance of svm, k-nn and nbc classifiers for text-independent speaker identification with and without modelling through merging models", SAÜ Fen Bil Der 20. Cilt, 1. Sayı, s. 1-6, 2016, DOI - 10.16984/saufenbilder.00295.
- [8] Nazia Hossain, Rifat Jahan, and Tanjila Tabasum Tunka, "Classified Frame-Energy Signal Using K-means Clustering", International Journal of Software Engineering & Applications (IJSEA), Vol.9, No.4, July 2018
- [9] John D. Markel and Jr. Augustine H. Gray, 1976, "Linear Prediction of Speech", Springer Verlag, Berlin, first edition
- [10] He, J., Liu, L. and Palm, G., 1999, "A Discriminative Training Algorithm for VQ-Based Speaker Identification", IEEE Trans on Speech and Audio Processing
- [11] 14. Fang, C. (2009). "From time wrapping (DTW) to Hidden markov Model (HMM)", University of Cincinnati.
- [12] Linde, Y., Buzo, A., and Gray, R. (1980). "An algorithm for vector quantization design", IEEE Transactions on Communications, Vol. 28, pp. 84-95.
