رده‌بندی سیگنال‌های مغزی EEG برای تشخیص بین دو واژه در گفتار خاموش

محمدرضا اصغری بجستانی*، غلامرضا محمدخانی، سعید گرگین، وحیدرضا نفیسی و غلامرضا فراهانی

پژوهشکده برق و فناوری اطلاعات، سازمان پژوهش‌های علمی و صنعتی ایران، تهران، ایران

چکیده

در این پژوهش، یک رابط مغز-رایانه در کاربرد مکالمهٔ خاموش برای شناسایی و تفکیک بین دو واژه پیاده‌سازی شده است. در طی آزمایش، بر اساس یک زمان‌بندی مشخص، افراد یکی از دو واژه یا سکوت را که به‌صورت تصادفی انتخاب شده است، بدون آن‌که برزبان آورند؛ در ذهن خود تکرار می‌کنند و سیگنال‌های مغزی آنان توسط یک دستگاه ثبت EEG آزمایشگاهی چهارده کاناله ثبت می‌شود. پس از پیش‌پردازش و حذف داده‌های مخدوش، ویژگی‌های مناسب از این سیگنال‌ها استخراج و برای شناسایی به یک رده‌بند داده می‌شود. دو ترکیب برای استخراج ویژگی و رده‌بندی انتخاب و بررسی شدند: استخراج ضرایب ویولت همراه با رده‌بند SVM و ویژگی حاصل از تحلیل مؤلفه‌های اساسی همراه با رده‌بند کمینه فاصله که ترکیب نخست عملکرد بهتری از خود نشان داد. تعداد کل رده‌ها در این آزمایش سه عدد بوده که شامل دو واژهٔ منتخب و سکوت می‌باشد. نتایج حاصل، نشان‌دهندهٔ امکان تفکیک واژگان با دقت متوسط ۵۶/۸ درصد (بیش از ۱/۷ برابر نرخ تصادف) است که در سازگاری با نتایج گزارش‌شده در فعالیت‌های مشابه است؛ اما هنوز دقت کافی برای کاربردهای واقعی ندارد.

واژگان کلیدی: مکالمه خاموش، رابط مغز-رایانه، تصور گفتار، سیگنال‌های مغزی.

Classification of EEG Signals for Discrimination of Two Imagined Words

Mohammad Reza Asghari Bejestani*, Gholam Reza Mohammadkhani, Saeed Gorgin, Vahid Reza Nafisi & Ghaolam Reza Farahani

Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

Abstract

In this study, a Brain-Computer Interface (BCI) in Silent-Talk application was implemented. The goal was an electroencephalograph (EEG) classifier for three different classes including two imagined words (Man and Red) and the silence. During the experiment, subjects were requested to silently repeat one of the two words or do nothing in a pre-selected random order. EEG signals were recorded by a 14 channel EMOTIV wireless headset. Two combinations of features and classifiers were used: Discrete Wavelet Transform (DWT) features with Support Vector Machine (SVM) classifier and Principle Component Analysis (PCA) features with a Minimum-Distance classifier. Both combinations were capable of discriminating between the three classes much better than the chance level (33.3%), none of them was reliable and accurate enough for a real application though. The first method (DWT+SVM) showed better results. In this case, feature set was D2, D3, D4 and A4 coefficients of 4-level DWT decomposition of the EEG signals, roughly corresponding to major frequency bands (Delta, Theta, Alpha and Beta) of these signals. Three binary SVM machines were used. Each machine was trained to classify between two of the three classes, namely Man/Red, Man/Silence or Red/Silence. Majority Selection Rule was used to determine final class. Once two of these classifiers presented the true class, a win (correct classification) was counted, otherwise a loss (false classification) was

* Corresponding author

*نویسندهٔ عهده‌دار مکاتبات

• تاریخ ارسال مقاله: ۱۳۹۷/۰۲/۱۶ • تاریخ پذیرش: ۱۳۹۹/۰۲/۲۴ • تاریخ انتشار: ۱۳۹۹/۰۶/۲۴ • نوع مطالعه: پژوهشی

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

considered. Finally, Monte-Carlo Cross Validation showed an overall performance of about 56.8% correct classification which is comparable with the results reported for similar experiments.

Keywords: Silent Talk, Imagined Speech, EEG signals, Classification, Brain-Computer interface

۱- مقدمه

مغز انسان از میلیاردها سلول عصبی(نرون) که به شکل شبکه‌های نورونی فعالیت می‌کنند، تشکیل شده است. انتقال و پردازش اطلاعات در مغز به‌وسیله پالس‌های الکتریکی درون و برون شبکه‌ای صورت می‌گیرد و این جریان‌ها موجب ایجاد میدان‌های الکتریکی و مغناطیسی در اطراف سر می‌شوند. دانشمندان بر این باورند که با اندازه‌گیری و ثبت تغییرات زمانی سیگنال‌ها و همچنین سیگنال‌های مکان‌های مختلف مغز و تفسیر آنها، می‌توان به فعالیت‌های مغزی صورت گرفته پی برد [1]. درمیان مهم‌ترین حوزه‌های پژوهشی در این زمینه می‌توان به توسعهٔ رابط‌های مغز-رایانه[1]، علوم شناختی[2]، عصب‌شناسی شناختی[3]، عصب‌شناسی محاسباتی[4] و... اشاره کرد.

رابط‌های مغز-رایانه به‌دنبال یافتن کانال ارتباطی مستقیم بین مغز و رایانه هستند تا از این طریق توانایی کنترل یک وسیله بیرونی را به‌طورمستقیم از طریق سیگنال‌های مغزی میسر سازند. این وسایل سیگنال‌های ناشی از فعالیت عصب‌ها را گرفته و آنها به فعالیت‌هایی که کاربر در نظر دارد ترجمه می‌کنند [2].

پردازش سیگنال‌های مغزی از سه مرحله اصلی تشکیل شده‌است: ۱- اخذ سیگنال‌های مغزی ۲- پردازش و استخراج ویژگی ۳- الگوریتم ترجمه. در مرحله اخذ سیگنال، میدان‌های الکتریکی یا مغناطیسی مغز ثبت و به سامانه پردازش منتقل می‌شوند. در مرحلهٔ پردازش به‌طورمعمول نوفه‌های موجود در سیگنال حذف شده، سیگنال‌های با باند فرکانسی مورد نظر از سیگنال اصلی استخراج و سپس ویژگی‌های مفید سیگنال محاسبه می‌شود. در مرحلهٔ آخر نیز با استفاده از الگوریتم ترجمه، فعالیت‌های مغزی مرتبط شناسایی و طبقه‌بندی می‌شوند [3-1].

مکالمه خاموش[5] یکی از پروژه‌های کاربردی در زمینهٔ پیاده‌سازی رابط‌های مغز-رایانه است که فعالیت‌های زیادی را در سال‌های اخیر برانگیخته است. هدف این پروژه ایجاد ارتباط نفر به نفر از طریق کدگشایی سیگنال‌های مغزی مرتبط با "گفتار اراده شده"[6] است [4].

در مکالمه خاموش، امواج مغزی حاصل از فعالیت عصبی پیش از تکلم، شناسایی و رمزگشایی و سپس بین دو نفر مبادله می‌شود.

کاربردی‌ترین فعالیت‌های انجام‌گرفته در این زمینه مربوط به استفاده از مؤلفه p300 امواج مغزی ERP[7] است [5, 6]. در نمونه‌ای از این کارها، ماتریسی از نویسه‌ها و علائم (به‌عنوان مثال در ابعاد ۶×۶) به کاربر نشان داده می‌شود و هر کدام از نویسه‌های موجود در ماتریس به‌صورت تصادفی پُرنور می‌شوند. درصورتی‌که نویسهٔ مورد نظر کاربر پُرنور شود، یک موج p300 در سیگنال‌های EEG فرد دیده می‌شود که از این طریق نویسهٔ مورد نظر او شناسایی می‌شود [7]. علاوه‌بر مسألهٔ استفاده از سیگنال‌های ERP به جای فعالیت عصبی مرتبط با گفتار اراده‌شده، مشکل این روش‌ها سرعت خیلی کم آنها است که سرعتی در حد چند نویسه در دقیقه را فراهم می‌کنند.

در گروهی دیگر از کارهای صورت‌گرفته در زمینهٔ مکالمهٔ خاموش، سیگنال‌های EEG به‌طور مستقیم برای شناسایی گفتار تصورشده استفاده شده‌اند. برخلاف روش‌های مبتنی بر مؤلفهٔ ERP-P300 که تشخیص بسیار کند است، این روش‌ها به‌دلیل استفاده مستقیم از سیگنال‌های EEG سرعت بالایی دارند؛ ولی میزان دقت دست‌یافته از آنها پایین و به همین دلیل تا کنون مطالعه روی آنها به موارد بسیار ساده محدود شده است. از این گروه، نتایجی برای تشخیص بین دو سیلاب [8, 9]، تشخیص بین حروف صدا دار انگلیسی [10, 11] یا تفکیک دو مفهوم [12] گزارش شده است. همچنین کاربردهایی مثل تشخیص گوینده [13] و آهنگ تکرار صداها [14, 15] نیز توسط روش‌های مشابه پیاده‌سازی شده‌اند. در [16] آزمایشی مربوط به تشخیص بین دو واژه left و right گزارش شده است.

۲- روش پژوهش

موضوع پژوهش، پیاده‌سازی یک BCI در کاربرد مکالمه خاموش برای تشخیص و تفکیک دو واژه در گفتار درونی است. دو واژهٔ Man و Red برای این منظور روی افراد

[1] Brain-Computer Interface
[2] Cognitive science
[3] Cognitive neuroscience
[4] Computational neuroscience
[5] Silent talk
[6] Intended/Imagined speech
[7] Event-Related Potential

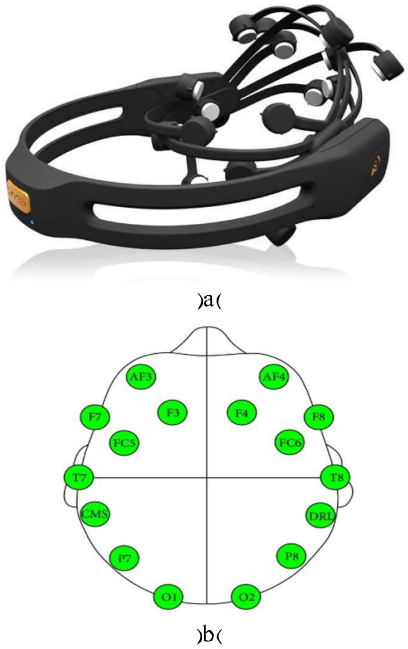
[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

فارسی‌زبان بررسی شده است. هر دوی این واژگان یک‌سیلابی، معنادار و دارای سه صدا هستند؛ اما از نظر اصوات و گروه معنایی متفاوتند. انتخاب این واژگان از جهاتی نزدیک به محتوای کارهای بررسی‌شدهٔ قبلی ولی از نظر کامل و معناداربودن واژگان مورد استفاده، گامی رو به جلوست. ترکیب روند، ابزار و مراحل پردازش نیز به همین ترتیب شباهت‌ها و تفاوت‌هایی با سایر پژوهش‌ها دارد.

۲-۱- ابزار ثبت

در این پژوهش از یک دستگاه Emotiv -که در شکل ۱-الف نشان داده شده است- برای ثبت سیگنال EEG استفاده شد. این دستگاه دارای چهارده کانال اصلی به‌علاوه دو کانال کنترلی با نرخ نمونه‌برداری ۱۲۸ هرتز است و سیگنال‌های گرفته‌شده را توسط یک رابط بلوتوث، به‌صورت بی‌سیم به رایانه منتقل می‌کند. محل اتصال الکترودهای این دستگاه به سر ثابت و توسط سازنده به‌صورت شکل ۱-ب در نظر گرفته شده است. آزمایش بر روی پنج نفر و هر کدام در سه جلسه صورت گرفته است. سیگنال‌ها ثبت و به‌صورت برون‌خط[1] پردازش شده‌اند. تمامی افراد مرد و در محدوده سنی بیست تا سی سال بوده‌اند.



(شکل-۱): دستگاه ثبت Emotiv (a) شکل ظاهری (b) محل قرارگیری الکترودها وکانال‌های ثبت

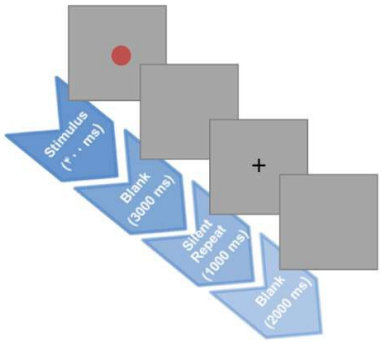
(Figure-1): EMOTIV EEG signal recorder. (a) The headset, (b) designed electrode positions

[1] off line

گفتنی است که در صورت به‌کارگیری دستگاه EEG با تعداد کانال بیشتر یا قابلیت جابه‌جایی و انتخاب مناسب‌تر محل الکترودها، ممکن است بتوان نتایج بهتری از آزمایش حاضر به‌دست آورد. ویژگی خاص این دستگاه دردسترس‌بودن، قیمت مناسب و به‌کارگیری آسان آن است.

۲-۲- شرح آزمایش

روند آزمایش در شکل (۲) نشان داده شده است. فرد مورد آزمایش در اتاقی آرام و با نور عادی در جلوی یک نمایش‌گر می‌نشیند. ابتدا یک علامت قرمز رنگ روی صفحهٔ نمایش‌گر نشان داده و هم‌زمان برای فرد یکی از واژگان Red، Man یا سکوت به‌صورت تصادفی پخش می‌شود. پخش این واژگان با فرکانس طبیعی برای انسان حدود چهارصد میلی ثانیه طول می‌کشد. بعد از پخش کلمه مورد نظر، یک بازه استراحت حدود سه ثانیه که صدایی پخش نمی‌شود و صفحه نمایش خاموش است وجود دارد. این فاصله به‌منظور سپری‌شدن هر گونه اثر حاصل از شنیدن و درک صدای محرک بر روی فعالیت‌های عصبی مغز در نظر گرفته شده است؛ سپس علامت "+" برای مدت یک ثانیه در صفحه نمایش نشان داده می‌شود. و درنهایت بعد از دو ثانیه فاصله که در آن نیز نمایش‌گر خاموش است، محرک بعدی پخش می‌شود. از کاربر خواسته شده‌بود که در زمان مشاهده علامت "+"، واژه‌ای را که شنیده بود، به‌صورت خاموش در ذهن خود- هر چند بار که می‌تواند- تکرار کند.



(شکل-۲): روند آزمایش

(Figure-2): Experiment timeline

در هر جلسه هر یک از محرک‌های Red، Man یا سکوت بیست بار با جایگشت تصادفی تکرار می‌شوند. هر فرد در سه جلسه جداگانه (هر جلسه درمجموع شصت تکرار) آزمایش را انجام می‌دهد تا از خستگی فرد به دلیل طولانی‌شدن جلسه آزمایش جلوگیری شود. برای کاهش

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]
[DOI: 10.29252/jsdp.17.2.120]

نوفه از کاربرها خواسته شده بود تا حد امکان، پلک‌زدن یا دیگر حرکت‌های خود را در فاصله‌ای که برای استراحت در نظر گرفته شده است، انجام دهند.

زمان دقیق شروع هر کدام از مراحل آزمایش و نوع محرک آن، هم‌زمان با انجام آزمایش و همراه با داده‌های EEG، توسط نرم‌افزار برچسب‌گذاری و ذخیره شده‌اند.

دلیل انتخاب این دو واژه معیارهایی مانند هم‌طول و هم‌آوابودن، رایج و آسان‌بودن و تشابه نسبی و نزدیک‌بودن به کارهای مشابه (مثل تک‌سیلابی‌بودن و تشابه نوع سیلاب‌ها) بوده است؛ به‌علاوه این دو واژه از نظر مفهوم در دو گروه معنایی متفاوت قرار دارند و در زبان فارسی نیز یکی معنادار و دیگری بی‌معناست.

۳-۲- پیش‌پردازش سیگنال‌ها

سیگنال‌های جمع‌آوری‌شده ابتدا با توجه به سازوکار برچسب‌گذاری برای محرک‌های مختلف تفکیک و دسته‌بندی می‌شوند. برای هر محرک دو بازه زمانی از داده‌ها اهمیت دارد: یکی از حدود ۱۲۰ میلی ثانیه تا ۶۷۰ میلی ثانیه از شروع پخش محرک صوتی که مربوط به سیگنال مغزی مرتبط با پردازش گفتار شنیده شده و دیگری از یکصد میلی‌ثانیه تا هزار میلی ثانیه بعد از شروع نمایش علامت + که نشان‌گر شروع و تکرار واژه شنیده شده است. این دو بازه زمانی به‌ترتیب معادل با درک گفتار شنیده شده و تولید گفتار خاموش در مغز هستند.

نرخ نمونه‌برداری برای دستگاه ثبت Emotiv، ۱۲۸ نمونه بر ثانیه است؛ بنابراین طول بردار داده برای سیگنال درک واژه شنیده‌شده هفتاد و برای سیگنال تولید گفتار خاموش ۱۱۵ خواهد بود.

به این ترتیب دادگان مورد استفاده شامل سه مجموعه شصت‌تایی (سه رده و هرکدام شصت تکرار، یا درمجموع ۱۸۰ آزمایش) از ۱۱۵ نمونه از چهارده کانال سیگنال‌های EEG برای هر کدام از آزمایش‌شوندگان بوده است.

با توجه به احتمال وجود داده‌های مخدوش در سیگنال ضبط‌شده، که بیش‌تر به‌خاطر پلک‌زدن یا حرکت‌های ناخواسته دیگر ایجاد می‌شوند، واریانس سیگنال هرکانال در هر آزمایش با متوسط واریانس سیگنال همان کانال در آزمایش‌های مختلف مقایسه و درصورتی‌که بیش از ۵۰٪ با آن اختلاف داشته باشد، تمام سیگنال آن آزمایش مخدوش فرض شده و از چرخهٔ پردازش حذف خواهد شد. این عمل بر این فرض استوار است که انرژی (واریانس) داده‌های ناشی از فعالیتی مثل پلک‌زدن بسیار بیش از داده‌های واقعی حاصل از تصور گفتار است. با توجه به کنترل مناسب شرایط آزمایش در زمان ثبت، تعداد آزمایش‌های مخدوش انگشت‌شمار بوده است. گفتنی است که به‌جز دسته‌بندی و حذف سیگنال‌های مخدوش، هیچ فیلتری روی داده‌ها اعمال نشده است. همچنین سیگنال‌های تمام کانال‌ها در رده‌بندی شرکت داده شده‌اند.

۴-۲- استخراج ویژگی و طبقه‌بندی

برای استخراج ویژگی روش‌های ویولت و تحلیل مؤلفه اساسی و برای طبقه‌بندی سه گروه رده‌بندهای SVM، KNN (یا فاصله از مرکز گروه) و شبکه‌های عصبی مختلف (ANN) مورد استفاده و آزمایش قرار گرفته و برای هر نوع ویژگی، مناسب‌ترین رده‌بند انتخاب شد. در ادامه این دو روش و رده‌بند استفاده شده با آنها توضیح داده شده است.

دو نوع ویژگی و روش‌های طبقه‌بندی انتخاب‌شده در اینجا از ساده‌ترین و رایج‌ترین روش‌ها در زمینهٔ پردازش سیگنال‌های EEG است [17] که در عین حال نتایج مطلوبی را هم نشان داده‌اند. واضح است که کار انتخاب ویژگی و روش رده‌بندی پیچیده‌ترین مرحله این نوع پژوهش است که روش‌های شناخته‌شده یا ناشناخته فراوانی را شامل می‌شود و موضوع بسیاری از فعالیت‌های پژوهشی حال حاضر در زمینه تحلیل خودکار سیگنال‌های مغزی را نیز تشکیل می‌دهند.

۱-۴-۲- روش ویولت گسسته (DWT)

برای پردازش سیگنال‌های زیستی، استفاده از ویولت یکی از روش‌های معمول است که کارایی خوبی از خود نشان داده است. در این پژوهش از موجک harr استفاده شده است هرچند استفاده از سایر موجک‌ها نیز تفاوت چندانی در نتایج ایجاد نمی‌کرد. تبدیل ویولت گسسته چهار مرحله‌ای برای انتخاب ویژگی استفاده و ضرایب ویولت D2، D3، D4 و A4 برای هر سیگنال ذخیره شد. با توجه به نرخ نمونه‌برداری ۱۲۸ هرتس، این ضرایب به‌طورتقریبی معادل باندهای فرکانسی مهم سیگنال EEG (باندهای ۱۶ تا ۳۲، ۸ تا ۱۶، ۴ تا ۸ و صفر تا چهار هرتس- دلتا، تتا، آلفا و بتا) هستند. برای استخراج ویژگی از تابع ویولت گسسته برنامه Matlab

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

استفاده شده است. بعد از اینکه بردار ویژگی با استفاده از این روش ایجاد شد، از رده‌بند SVM خطی برای رده‌بندی استفاده کردیم (طبقه‌بندی به‌وسیلهٔ SVM کارایی بهتری برای داده‌های ثبت‌شده فراهم می‌کرد). برای جداسازی سه رده رأی اکثریت را به‌کار بردیم. درواقع سه ماشین مختلف- هریک برای تفکیک دو رده از یگدیگر- آموزش داده شد و اگر دو ماشین رده را به‌درستی نشان می‌دادند، یک تشخیص درست و در غیر این صورت یک تشخیص نادرست محاسبه شده است.

۲-۴-۲- روش تحلیل مؤلفه اساسی (PCA)

برای این منظور، همهٔ سیگنال‌های مربوط به رده‌های مختلف محرک‌ها را به تابع تحلیل مؤلفه اساسی داده و درنهایت تعداد شانزده مؤلفه اساسی نخست را برای مرحله طبقه‌بندی (بر اساس حداقل فاصله) انتخاب کردیم. تعداد شانزده مؤلفه اساسی بهترین کارایی را فراهم می‌کرد.

۳- نتایج آزمایش

نتایج آزمایش برای هر یک از دو روش پردازش و به‌ازای هریک از سوژه ها در جدول (۱) آورده شده است.

با توجه به کم‌بودن تعداد نمونه‌ها، از روش ارزیابی متقابل مونته کارلو[1] استفاده کرده‌ایم. هر یک از مقادیر ذکرشده در جدول (۱)، متوسط دقت به‌دست‌آمده از چندین مرتبه آموزش و رده‌بندی در هرکدام از روش‌ها است. در هر مرتبه حدود ۷۰٪ از داده‌ها (حدود ۴۲ نمونه آزمایش از هر رده) با انتخاب تصادفی برای آموزش و ۳۰٪ بقیه (حدود ۱۸ نمونه) برای رده‌بندی به‌کار برده شده‌اند. مقدار انحراف معیار نتایج برای هر سوژه حدود ۱۰٪ بوده است.

(جدول-۱): دقت رده‌بندی به درصد بین تمام رده‌ها به‌ازای هر یک از پنج سوژه و دو روش پردازش

(Table-1): Percent of correct classifications for each subject and each processing method

روش پردازش	DWT	PCA
سوژه ۱	۴۴	۴۷
سوژه ۲	۵۷	۵۱
سوژه ۳	۵۴	۴۳
سوژه ۴	۶۸	۵۰
سوژه ۵	۶۱	۵۵
متوسط تمام سوژه‌ها	۵۶/۸	۴۹/۲

[1] Monte-Carlo Cross Validation

با توجه به وجود حالت سکوت علاوه‌بر دو واژهٔ مورد آزمایش، درواقع رده‌بندی بین سه رده انجام گرفته و نرخ دقت انتخاب تصادفی ٪۳۳ است و بنابراین با توجه به نتایج به‌دست‌آمده می‌توان گفت که در هر دو روش دقت عملکرد به‌طور معناداری بالای نرخ تصادف است. همچنین ملاحظه می‌شود که استفاده از روش نخست (ویژگی‌های DWT و طبقه‌بند SVM) نتایج بهتری نسبت به روش دوم ایجاد کرده‌است. البته دقت هیچکدام از دو روش در حد بالا، به‌طوری که بتوان از آن برای کاربردهای واقعی استفاده کرد، نیست.

۴- مقایسه و نتیجه‌گیری

همان‌طور که اشاره شد، آزمایش‌های گزارش‌شده مشابه در منابع در دسترس، به مقایسه بین دو یا چند صدا یا سیلاب بسنده کرده‌اند و گزارش کمّی از رده‌بندی واژگان وجود ندارد؛ به همین خاطر در اینجا نتایج گزارش‌شده در سه مقاله که روی سیلاب‌ها و صداها پژوهش کرده‌اند، برای مقایسه آورده می‌شود.

نتایج گزارش‌شده در منبع [9] در جدول (۲) آورده شده است. این آزمایش به‌منظور رده‌بندی دو سیلاب /ba/ و /ku/ در آزمایشی روی هفت نفر با استفاده از دستگاه ثبت EEGی ۱۲۸ کاناله و نرخ نمونه‌برداری ۱۰۲۴ هرتز، انجام شده است.

(جدول-۲): نتایج رده‌بندی بین دو آوای /ba/ و /ku/ برای هریک از هفت سوژهٔ مورد آزمایش در [9]

(Table-2): Average classification accuracy (/ba/ or /ku/) for each of the 7 subjects in [9].

	نرخ طبقه‌بندی	تعداد آزمایش
سوژه ۱	۵۶	۱۵
سوژه ۲	۸۸	۱۱
سوژه ۳	-	۲
سوژه ۴	۴۶	۱۳
سوژه ۵	۷۵	۲۴
سوژه ۶	۸۱	۱۲
سوژه ۷	۶۷	۸
متوسط تمام سوژه‌ها	۶۸/۸۴	۸۳

مقاله [18] به مسألهٔ تفکیک سیگنال‌های مربوط به تصور دو صدای /a/ و /u/ همراه با تصور تغییر شکل لب‌ها به‌هنگام تلفظ این صداها پرداخته است. آزمایش روی سه نفر انجام شده و سیگنال‌های مغزی با یک دستگاه ۶۴ کاناله و با نرخ نمونه‌برداری ۲۰۴۸ هرتز ثبت شده است. نتایج گزارش‌شده در جدول (۳) آورده شده است.

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

(جدول-۳): کارایی تشخیص واکه‌ها برای هر یک از افراد شرکت‌کننده و حالت‌های مختلف رده‌بند [18]

(Table-3): Classification accuracies and standard deviations (rounded %) for 3 subjects (S1, S2and S3) on 3 pairwise discriminations and overall.

	/a/:سکوت	/u/:سکوت:	/u/:/a/	متوسط
سوژه ۱	۳±۷۹	۴±۸۲	۳±۷۲	۵±۷۸
سوژه ۲	۵±۷۱	۴±۷۲	۵±۶۰	۷±۶۸
سوژه ۳	۴±۶۷	۳±۸۰	۴±۵۶	۱۲±۶۸
متوسط سه سوژه	۷۲/۳۳	۷۸	۶۲/۶۷	۷۱/۳۳

در [11] نتایج طبقه‌بندی دوبه‌دوی پنج رده آوایی منتشر شده است. پنج نفر در آزمایش شرکت کرده‌اند و از دو صدا برای هر رده استفاده شده است. جدول (۴)، رده‌ها و صداهای مربوط به هر یک را نشان می‌دهد.

(جدول-۴): ده نمونه صدای استفاده‌شده برای پنج رده آوایی در [11]

(Table-4): The ten indicated phonemes were chosen to contrast five patterns of vocal articulation (jaw, tongue, nasal, lips and fricative) [11]

Articulation class	*Phoneme*	*Examples*	*Manner of Production*
Jaw	-aa	saw, jaw	"ah"
	-ae	hat, cat	"aeh"
Tongue	-l	light, led	"el"
	-r	right, red	"are"
Nasal	-m	mat, mice	"em"
	-n	net, nice	"en"
Lips	-uu	who, drew	"oo"
	-ow	boat, over	"oh"
Fricative	-s	same, hiss	s - hissing
	-z	zoo, his	z - buzzing

در این آزمایش داده‌های ۵۲ کانال با نرخ هزار هرتز ثبت و چند ویژگی مبتنی بر دامنه طیفی فرکانس‌های چهار تا ۲۸ هرتز از آن استخراج شده است. یک طبقه‌بند LDA[1] برای طبقه‌بندی دوبه‌دوی پنج رده آوایی به‌علاوه یک رده "سکوت" به‌کار برده شد. نتایج حاصل در جدول (۵) دیده می‌شود. عدد نخست در هر سطر سلول‌های این جدول دقت رده‌بندی داده‌های یک روز و عدد دوم دقت رده‌بندی روی کل داده‌های ثبت شده است. متوسط دقت به‌دست‌آمده حدود هفتاد درصد برای داده‌های یک روز و شصت درصد برای کل دادگان است.

(جدول-۵): دقت عملکرد رده‌بند LDA در ۱۵ طبقه‌بندی دو به دوی پنج رده آوایی برای پنج سوژه [11]

(Table-5): LDA classification performance for 15 pairwise classifications of 5 articulation classes for 5 subjects. [11]

	Subject	Tongue	Nasal	Lips	Fricative	Relax
Jaw	A	68.0, *61.5*	*73.2*, 59.4	*72.4*, *62.0*	*73.7*, *63.0*	*71.9*, *68.8*
	B	66.9, 58.6	*70.1*, *62.2*	*69.6*, 60.7	*71.8*, 60.4	*73.7*, *68.8*
	C	*73.2*, *63.7*	71.3, *61.6*	*74.5*, 59.7	*75.2*, 60.5	*73.9*, *62.4*
	D	*71.7*, *60.7*	*69.8*, *60.7*	*73.0*, *60.7*	*74.8*, *61.2*	*70.8*, *61.9*
	E	*70.9*, *62.9*	*71.6*, *65.4*	*72.4*, *62.4*	*69.8*, *66.4*	*72.1*, *70.3*
Tongue	A		67.7, *65.6*	*75.3*, *62.2*	*74.0*, *66.4*	*72.1*, *69.5*
	B		71.2, *62.8*	*70.9*, 56.5	67.1, 61.3	*70.5*, *63.7*
	C	-	*74.0*, *64.7*	71.3, 58.9	*73.8*, *62.1*	*70.6*, *61.8*
	D		70.2, *62.6*	*72.2*, 57.7	*74.6*, 57.0	*69.1*, *60.5*
	E		*76.0*, *66.1*	*69.7*, *65.0*	*70.6*, *68.2*	*72.3*, *63.6*
Nasal	A			*72.1*, 58.3	*72.1*, *68.0*	*72.1*, *65.1*
	B			69.3, 58.6	*72.1*, *64.0*	*70.6*, *67.0*
	C	-	-	*72.6*, *61.1*	*73.6*, 58.4	*72.1*, *63.2*
	D			66.4, 59.1	*71.5*, 59.8	*70.3*, 59.1
	E			*73.7*, *66.1*	*71.0*, *67.8*	*72.9*, *68.7*
Lips	A				68.8, *62.8*	*75.5*, *69.0*
	B				*69.6*, *62.2*	68.1, *62.8*
	C	-	-	-	*74.8*, *64.5*	*69.0*, *60.0*
	D				*73.1*, *60.3*	69.5, 57.7
	E				*72.5*, *68.0*	*70.5*, *68.9*
Fricative	A					*75.5*, *69.0*
	B					*74.3*, *65.8*
	C	-	-	-	-	*71.0*, *64.5*
	D					*72.8*, *59.8*
	E					*71.5*, *66.6*

همان‌طور که ملاحظه می‌شود نتایج به‌دست‌آمده از پژوهش حاضر قابل مقایسه و به‌طورتقریبی معادل مراجع ذکرشده است. البته گفتنی است که پژوهش حاضر به‌خاطر استفاده از واژگان کامل و همچنین به‌کارگیری دستگاه ثبتِ ساده‌تر، نسبت به بسیاری از کارهای قبلی، برتری دارد. استفاده از سه واژه، هم در ثبت و هم در تفکیک داده‌ها وجه تمایز اصلی این پژوهش است.

نتایج حاصل، اگر چه دقت کافی برای به‌کارگیری جدی در یک سامانه کاربردی برخط را ندارند، اما نشان می‌دهند که هدف مکالمه خاموش، با صرف تلاش و انرژی بیشتر، دور از دسترس نخواهد بود و رسیدن به آن امکان‌پذیر است.

5- References ۵- مراجع

[1] N. Birbaumer and L. G. Cohen, "*Brain–computer interfaces: communication and restoration of movement in paralysis*", The Journal of physiology, vol. 579, no. 3, pp. 621–636, 2007.

[2] Ivan S. Kotchetkov, Brian Y. Hwang and et. al., "Brain computer Interfaces: military, neurosurgical and ethical prespective", *Neurosurg Focus*, vol. 28, no. 5, pp. 1-6, 2010.

[3] M. A. Lebedev and M. A. L. Nicolelis, "*Brain-Machine interfaces: past, present and future*", *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.

[4] Defense Advanced Research Projects Agency: Department of Defense Fiscal Year 2010 Budget Estimates Washington, DC, Department of Defense, 2009

[1] Linear Discriminant Analysis

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

[5] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, and T. M. Vaughan, "A P300-based brain–computer interface for people with amyotrophic lateral sclerosis", Clinical neurophysiology, vol. 119, no. 8, pp. 1909–1916, 2008.

[6] E. Donchin and Y. Arbel, "P300 based brain computer interfaces: a progress report", Foundations of Augmented Cognition, Neuroergonomics and Operational Neuroscience, pp. 724–731, 2009.

[7] L. A.Farwell, E.Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials", *Electroenceph clin Neurophysiol,* Vol. 70, pp. 510–523, 1988.

[8] S. Iqbal , Y.U. Khan , O. Farooq ," EEG based classification of imagined vowel sounds", *2nd International Conference on Computing for Sustainable Global Development (INDIACom)*,pp. 1591-1594, IEEE 2015.

[9] K. Brigham, B.V.K.V Kumar, " *Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy*", *4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE)*, pp. 1-4, IEEE 2010.

[10] R. Kamalakkannan, R. Rajkumar, R. M. Madan, D. S. Shenbaga, "Imagined Speech Classification using EEG", *Advances in Biomedical Science and Engineering*, Vol. 1, No. 2, pp.20-32, 2014.

[11] X. Chi, J. B. Hagedorna, D. Schoonovera, and M. D'Zmuraa, "EEG-based discrimination of imagined speech phonemes", *International Journal of Bioelectromagnetism,* vol. 13, no. 4, 2011.

[12] T. Kim, J. Lee, H. Choi, H. Lee et al., "Meaning based covert speech classification for brain-computer interface based on electroencephalography", *6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 53-56, 2013.

[13] K. Brigham , B.V.K.V. Kumar, "Subject identification from electroencephalogram (EEG) signals during imagined speech", *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)* ,pp. 1-8, 2010.

[14] M. D'Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, "Toward EEG sensing of imagined speech", *Human-Computer Interaction New Trends*, pp. 40-48, 2009.

[15] S. Deng, R. Srinivasan, T. Lappas, and M. D'Zmura, "EEG classification of imagined syllable rhythm using Hilbert spectrum methods", *Journal of neural engineering*, vol. 7, no. 4,pp. 046006, 2010.

[16] K. Yaser Arafat, S. S. Kanade, "Imagined Speech EEG Signal Processing For Brain Computer Interface", *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*, Vol. 3, No. 7, pp. 123, 2014.

[۱۷] ا. شاه بهرامی، ک. نجفی، ط. نجفی. "حوزه‌های مختلف کاربردی پردازش سیگنال مغزی در ایران". فصل‌نامه پردازش علائم و داده‌ها، جلد ۱۳، شماره ۳، صفحات ۱۲۹-۱۵۴، ۱۳۹۵.

[17] A. Shahbahrami, K. Nadjafi, T.Nadjafi, "Different Application Fields of Brain Signal Processing", *Quarterly Journal of Signal and Data Processing(JSDP)*, Vol.13, No.3, pp129-154, 2016.

[18] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns", *Neural Networks*, vol. 22, no. 9, pp. 1334–1339, 2009.

محمدرضا اصغری بجستانی تحصیلات خود در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق-مخابرات را به‌ترتیب از دانشگاه‌های شیراز و تربیت مدرس تهران و در سال‌های ۱۳۶۹ و ۱۳۷۳ به اتمام رسانده و هم اکنون نیز دانشجوی دکترای تخصصی همین رشته در سازمان پژوهش‌های علمی و صنعتی ایران (IROST) است. وی از سال ۱۳۷۰ تا کنون به فعالیت‌های آموزشی و پژوهشی در زمینه‌های مختلف مرتبط با رشته تحصیلی خود اشتغال داشته است. او به پژوهش و توسعه سامانه‌های نرم‌افزاری پردازش سیگنال‌های دیجیتال (صوت، تصویر، بیولوژیک و ...) علاقه‌مند است.

نشانی رایانامه ایشان عبارت است از:

bejestani@irost.ir

غلامرضا محمدخانی تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق از دانشگاه صنعتی شریف تهران به‌ترتیب در رشته‌های الکترونیک و مخابرات و در سال های ۱۳۷۲ و ۱۳۷۶ به اتمام رساند. وی از سال ۱۹۹۹ (۱۳۷۸) تا ۲۰۰۲ (۱۳۸۱) تحصیلات خود را در مقطع دکترا در مدرسه عالی مهندسی ENSIL از دانشگاه لیموژ (Ecole National Supérieur d'Ingénieur de Limoges) در کشور فرانسه ادامه داد. همچنین از سال ۲۰۰۲ (۱۳۸۱) تا ۲۰۰۶ (۱۳۸۵)

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]

به‌عنوان پژوهش‌گر پسادکترا در مرکز پژوهشی (LSS Laboratory of Signals and systems) از CNRS در Supelec پاریس مشغول به‌کار بوده است. وی از سال ۱۳۸۵تاکنون عضو هیأت علمی پژوهشکده برق سازمان پژوهش‌های علمی و صنعتی ایران است و مسئولیت آزمایشگاه رابط‌های مغز–رایانهٔ این پژوهشکده را نیز به‌عهده دارد. زمینه‌های مورد علاقه ایشان پردازش سیگنال در مخابرات دیجیتال و بی‌سیم است.

نشانی رایانامه ایشان عبارت است از:

mohammadkhani@irost.org

سعید گرگین مدرک کارشناسی و کارشناسی ارشد مهندسی رایانه خود را به‌ترتیب از دانشگاه آزاد اسلامی تهران جنوب و علوم و تحقیقات تهران و مدرک دکترا معماری رایانه را از دانشگاه شهید بهشتی، در سال ۱۳۸۹ دریافت کرد. وی هم‌اکنون استادیار مهندسی رایانه در گروه فناوری اطلاعات و سامانه‌های هوشمند سازمان پژوهش‌های علمی و صنعتی ایران (IROST) است. علایق پژوهشی ایشان شامل حساب رایانه‌ای و الگوریتم‌های محاسباتی، سامانه‌های با قابلیت پیکربندی مجدد، پردازش سریع و طراحی VLSI است.

نشانی رایانامه ایشان عبارت است از:

gorgin@irost.ir

وحیدرضا نفیسی تحصیلات خود را در مقطع دکترا رشته مهندسی پزشکی گرایش بیوالکتریک دانشگاه صنعتی امیرکبیر در سال ۱۳۸۴ به اتمام رسانده و از سال ۱۳۸۱ به‌عنوان عضو هیأت علمی در سازمان پژوهش‌های علمی و صنعتی ایران مشغول به کار است. حوزهٔ فعالیت وی ابزار دقیق (تجهیزات دیالیز و دستگاه‌های تنفس مصنوعی) و پردازش سیگنال‌های حیاتی می‌باشد.

نشانی رایانامه ایشان عبارت است از:

vr_nafisi@irost.org

غلامرضا فراهانی دانشیار پژوهشکده برق و فناوری اطلاعات سازمان پژوهش‌های علمی و صنعتی ایران است. ایشان در سال ۱۳۷۷ مدرک کارشناسی خود را از دانشگاه صنعتی شریف و در سال ۱۳۷۹ مدرک کارشناسی ارشد و سال ۱۳۸۵ مدرک دکترای خود را در رشته مهندسی برق گرایش الکترونیک دیجیتال از دانشگاه صنعتی امیرکبیر دریافت کرد. وی از سال ۱۳۸۷ تا کنون در پژوهشکده برق و فناوری اطلاعات به‌عنوان عضو هیأت علمی فعالیت دارد. یکی از زمینه‌های علمی مورد علاقه ایشان پردازش سیگنال (صوت و تصویر) است.

نشانی رایانامه ایشان عبارت است از:

farahani.gh@irost.org

[Downloaded from jsdp.rcisp.ac.ir on 2024-04-20]

[DOI: 10.29252/jsdp.17.2.120]