

استفاده از ماشین بردار پشتیبان ابربردار مدل مخلوط گوسی برای تشخیص خودکار نوع مدولاسیون دیجیتال

محمد هادی ولی پور^۱، محمد مهدی همایون پور^۱، محمد امین مهرعلیان^۱

^۱ آزمایشگاه پردازش هوشمند سیگنال‌های صوتی و گفتاری، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، تهران، ایران
{valipour, homayoun, mehralian}@aut.ac.ir

چکیده

تشخیص خودکار مدولاسیون دیجیتال در سیستم‌های هوشمند مخابراتی یکی از موضوعات مطرح در رادیوی نرم‌افزاری و رادیوی شناختی است. در این مقاله روشی برای دسته‌بندی خودکار نوع مدولاسیون با استفاده از ابربردار مدل مخلوط گوسی و ترکیب با دسته‌بندی کننده ماشین بردار پشتیبان ارائه می‌شود. در این روش مجموعه‌ای از سه نوع ویژگی متفاوت مبتنی بر طیف، آماری و مبتنی بر تبدیل موجک از فریم‌های یک قطعه سیگنال استخراج خواهند شد و با استفاده از تطبیق با مدل زمینه ابربردار مورد نظر برای آن قطعه سیگنال ایجاد می‌شود. استفاده از چند ماشین بردار پشتیبان به صورت سلسله مراتبی برای دسته‌بندی کننده‌های برون کلاسی و درون کلاسی و نیز انتخاب ویژگی مناسب نتایج خوبی را حاصل کرده است، به طوری که در نویز کم شناسایی با خطای صفر صورت می‌گیرد. همچنین بر اساس نتایج آزمایش‌ها، روش ارائه شده مقاومت مناسبی را در مقابل نویز نشان می‌دهد.

کلمات کلیدی

مدولاسیون دیجیتال، رادیو نرم‌افزاری، ابربردار، مدل مخلوط گوسی، ماشین بردار پشتیبان.

۱- مقدمه

مروری بر کارهای انجام شده در این زمینه نشان می‌دهد که تشخیص خودکار نوع مدولاسیون به دو شیوه اصلی انجام می‌پذیرد. در شیوه اول با استفاده از روش‌های مبتنی بر درست‌نمایی^۲ تلاش می‌شود سیگنال ورودی به دسته‌ای منتسب شود که بر اساس مدل آن دسته بیشترین شباهت را با آن داشته باشد. شیوه دوم که سابقه کمتری نسبت به شیوه قبلی دارد بر اساس ویژگی‌های استخراج شده از سیگنال ورودی عمل می‌کند. از آنجا که در این مقاله نیز بر اساس شیوه دوم عمل شده است در ادامه به چند نمونه از کارهای صورت گرفته در این شیوه اشاره خواهیم کرد.

در [1] مدولاسیون‌های کلاس 2ASK, 4ASK, 8ASK, 4PSK, 8PSK و 16QAM مورد بررسی قرار گرفته است. انباشتک‌های^۳ مرتبه ۴ و ۶ ویژگی‌های اصلی استفاده شده در این مقاله است و در نهایت مجموع تمام ویژگی‌ها توسط یک ماشین بردار پشتیبان با ساختار درخت دودویی دسته‌بندی می‌شوند. در [2] ۵ کلاسه پایه‌ی مدولاسیون ASK, FSK, PSK, QAM, GMSK مورد تفکیک قرار گرفته است و ویژگی‌های اصلی مورد استفاده فاز، دامنه و فرکانس

تشخیص نوع مدولاسیون با گسترش روز افزون گیرنده‌های هوشمند که قادر به شناسایی کور سیگنال دریافتی باشند تبدیل به یک موضوع چالش برانگیز در مهندسی ارتباطات، مخابرات و کامپیوتر شده است.

چنین سیستم‌هایی دارای کاربردهای نظامی وسیعی از جمله در جنگ الکترونیک هستند. علاوه بر این تشخیص کور نوع مدولاسیون در سیگنال‌های دریافتی مساله مهمی در سیستم‌های تجاری خصوصاً در سیستم‌های رادیویی نرم‌افزاری^۱ که درگیر دامنه وسیعی از سیستم‌های ارتباطی هستند، می‌باشد. در این سیستم‌ها معمولاً اطلاعات تکمیلی برای پیکربندی سیستم ارسال می‌شود اما روش‌های کور می‌توانند با استفاده از گیرنده‌های هوشمند با کاهش سربارهای اطلاعاتی باعث افزایش راندمان انتقال شوند. که رسیدن به چنین هدفی جز با آنالیز نوع و مشخصات مدولاسیون داده‌های دیجیتال امکان پذیر نیست.

² Likelihood
³ Cumulant

¹ Software Defined Radio (SDR)

سعی بر این بوده تا دسته‌بندی‌ها با دقت بالایی صورت گیرد در حالی که در دسته‌ی دیگری از مقالات تلاش نویسندگان بر این بوده است که سیستم ارائه شده قادر به دسته‌بندی طیف وسیعی از مدولاسیون‌های مختلف باشد که با افزایش تعداد این کلاس‌ها دقت دسته‌بندی کاهش پیدا کرده است.

در این مقاله روشی جدید برای دسته‌بندی نوع مدولاسیون با استفاده از ابربردارها و مبتنی بر قطعه ارائه می‌شود. می‌توان یک قطعه سیگنالی که توسط یک مدولاسیون خاص تولید شده است را توسط یک ابربردار واحد نمایش داد. در اینجا با الهام از همین روش، برای هر قطعه یک ابربردار گوسی واحد در فضای ابربردار به دست می‌آید. در ابتدا قطعه سیگنال مورد نظر فریم بندی می‌شود و یک انتقال از فضای سیگنال به فضای ویژگی انجام خواهد شد، که ویژگی‌های مورد نظر در بخش استخراج ویژگی آمده است. با توجه به اینکه از یک ساختار سلسله مراتبی برای دسته‌بندی استفاده شده است، دسته‌بندی کننده‌های مختلفی خواهیم داشت که برای هر کدام انتخاب ویژگی انجام خواهد شد. در مرحله بعد مجموعه‌ای از بردارهای ویژگی خواهیم داشت که از فریم‌های متوالی یک قطعه سیگنال استخراج شده‌اند. از مجموعه‌ی این بردارهای ویژگی برای تمام قطعات یک مدل زمینه ساخته خواهد شد و از خروجی تطبیق مدل زمینه با بردارهای ویژگی مربوط به یک قطعه سیگنال به عنوان ابربردار استفاده خواهد شد. همچنین دو کرنل مختلف برای دسته‌بندی کننده ماشین بردار پشتیبان به کار گرفته شده و مورد ارزیابی قرار گرفته‌اند.

ادامه مقاله در شش بخش سازمان‌دهی شده است. در بخش ۲ توضیح مختصری از مفهوم ابربردار ارائه می‌شود. بخش ۳ مربوط به استخراج و انتخاب ویژگی از فریم‌های یک قطعه سیگنال است. روش ساخت ابربردار در بخش ۴ ارائه می‌شود. در بخش ۵ به مسایل مربوط توسعه دسته‌بندی‌کننده پرداخته می‌شود. در بخش ۶ نتایج حاصل ارائه و بررسی می‌گردند. بخش ۷ نیز به خلاصه‌ی مطالب، بحث و نتیجه‌گیری اختصاص داده شده است.

۲- ابربردار

ابربردارها یکی از جدیدترین روش‌ها در ارائه نمونه‌های سیگنال‌های گفتاری هستند و عمده تحقیقات در زمینه آن مربوط به ۱۰ سال گذشته است و بیشتر مقالات منتشر شده از ۲۰۰۴ تا کنون بوده است، لذا می‌توان این مبحث و تکنیک‌ها و کاربردهای آن در پردازش هوشمند را نسبتاً نوظهور دانست. نقطه عطف کاربرد ابربردارها استفاده از پارامترهای مدل‌های مخلوط گوسی و ترکیب آن با دسته‌بندی کننده‌هایی همچون ماشین بردار پشتیبان بود که نتایج آن توسط کمپل و همکارانش در ۲۰۰۶ ارائه شده است [6]. در سال‌های اخیر استفاده از ابربردارها به منظور بهسازی و همچنین افزایش مقاومت در برابر نویز گسترش پیدا کرده است و از این قابلیت در

لحظه‌ای است. نویسنده در این مقاله از روش پرودوگرام برای تخمین فرکانس مرکزی استفاده کرده است. در [3] نویسنده از گشتاورهای مرتبه بالای تبدیل موجک پیوسته به عنوان ویژگی استفاده کرده است، ویژگی‌های مذکور توسط یک دسته‌بندی کننده مبتنی بر شبکه عصبی BP دسته‌بندی می‌شوند. علاوه بر این در فرآیند پیش‌پردازش برای افزایش راندمان، نویسنده از روش PCA^۱ برای انتخاب ویژگی استفاده کرده است و نتایج این مقاله نشان داده است روش مذکور مقاومت مناسبی را در برابر نویز از خود نشان می‌دهد. در [4] با استفاده از روش fuzzy c-means نمودار فلکی خوشه‌بندی شده است و از آنجا که دسته‌بندی کلاس‌های مختلف مدولاسیون بر مبنای نمودار فلکی آنها صورت گرفته است بنابراین دسته‌بندی تنها محدود به سیگنال‌های کلاس PSK و QAM خواهد بود.

یکی از روش‌های مقاوم جهت ارائه نمونه‌های گفتاری به صورت یک بردار واحد، ابربردارها^۲ هستند که به تازگی مورد استفاده قرار گرفته‌اند. روش‌های موجود برای ارائه نمونه‌های سیگنالی به صورت یک ابربردار واحد متعدد نیستند، و این مسئله بیشتر به دلیل معطوف بودن به حوزه بازشناسی گوینده به صورت خاص بوده است. اما می‌توان با در نظر گرفتن مفهوم ابربردار به صورت یک نمایش کلی از یک نمونه سیگنالی در یک بازه‌ی زمانی (برای مثال سیگنال‌های صوتی) یا مکانی (برای مثال سیگنال‌های تصویری) خاص، روش‌های مختلفی را در این دسته قرار داد. حتی برخی از روش‌های قدیمی مانند میانگین زمانی ویژگی‌های استخراج شده از فریم‌های متوالی به عنوان نوعی روش نمایش یک نمونه سیگنالی با طول زیاد در فضای ابربردار دانست. پس لزومی به بالا بودن ابعاد در فضای ابربردار نیست. آنچه در تولید و ساخت یک ابربردار مهم است، انتقال یک مجموعه با اندازه‌های مختلف از بردارهای ویژگی به یک فضا با ابعاد مشخص و ثابت است [5].

غالب روش‌هایی که برای ساخت ابربردار وجود دارد، روش‌هایی هستند که از خروجی بخشی از فرآیندهای تشخیص و یا تصدیق و نیز مدل‌سازی نمونه‌های سیگنالی (عمدتاً گفتاری و بعضاً تصویری) به دست آمده‌اند. برای مثال استفاده از میانگین زمانی به عنوان ابربردار، روشی برای حذف برخی از نویزهای جمع شونده با توزیع مشخص بوده است، همچنین در روش‌های مبتنی بر مدل مخلوط گوسی، از خروجی بخش تطبیق گوینده و مدل زمینه جهانی برای تولید ابربردار استفاده می‌شود. با توجه به اینکه بسیاری از روش‌های مبتنی بر ابربردارها به صورت مشخص برای یک حوزه‌ی خاص و یک کاربرد مشخص توسعه داده شده‌اند (برای مثال حوزه‌ی پردازش گفتار، کاربرد در بازشناسی گوینده)، ممکن است بعضاً قابل گسترش برای سایر حوزه‌ها و کاربردها نباشند. طی بررسی‌های انجام شده، تا کنون هیچ تحقیق یا کاربرد عملی در زمینه تشخیص خودکار مدولاسیون مبتنی بر استفاده از ابربردارها گزارش نشده است.

در مجموع با مروری بر کارهای انجام شده در زمینه تشخیص خودکار نوع مدولاسیون می‌توان دریافت که در دسته‌ای از مقالات

$$\sigma_{aa} = \sqrt{\frac{1}{N_s} \left[\sum_{i=1}^{N_s} a_{cn}^2(i) \right] - \left[\frac{1}{C} \sum_{i=1}^{N_s} |a_{cn}(i)| \right]^2} \quad (4)$$

۵- انحراف معیار مقدار مطلق فرکانس لحظه‌ای نرمال شده و مرکزی

$$\sigma_{af} = \sqrt{\frac{1}{C} \left[\sum_{a_n(t) > a_t} f_N(i)^2 \right] - \left[\frac{1}{C} \sum_{a_n(t) > a_t} |f_N(i)| \right]^2} \quad (\Delta)$$

در رابطه فوق f_N فرکانس لحظه‌ای است که از معادله (۶) به دست می‌آید.

$$f(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \quad (6)$$

در محاسبه فرکانس لحظه‌ای، $\emptyset(t)$ تقریب خطی مولفه‌ی غیرخطی فاز لحظه‌ای است.

۳-۲- ویژگی‌های آماری

تحقیقات صورت گرفته در [10] نشان می‌دهد بکارگیری انباشتک‌های مرتبه‌ی بالا نیز می‌تواند اثربخشی خوبی بر نتایج نهایی داشته باشد. بنابراین از انباشتک‌ها به عنوان ویژگی‌های بعدی استفاده خواهیم کرد.

برای یک فرایند تصادفی مختلط، گشتاور مرکب مرتبه p با استفاده از رابطه (۷) محاسبه می‌شود.

$$M_{pq} = E[X(t)^{p-q} X^*(t)^q] \quad (\mathbf{Y})$$

در این تساوی * نشان‌دهنده تابع مختلط مزدوج است. انباشتک-های مرتبه بالا یک سیگنال با استفاده از تساوی‌های زیر محاسبه می-شود.

$$C_{X_1, X_2} = \langle X_1, X_2 \rangle = \frac{1}{N} \sum_{n=1}^N x_1^n x_2^n \quad (\text{A})$$

$$C_{X_1, X_2, X_3} = \langle X_1, X_2, X_3 \rangle = \frac{1}{N} \sum_{n=1}^N x_1^n x_2^n x_3^n \quad (9)$$

$$C_{X_1, X_2, X_3, X_4} = \frac{1}{N} \sum_{n=1}^N x_1^n x_2^n x_3^n x_4^n - C_{X_1, X_2} C_{X_3, X_4} \quad (9)$$

$$-C_{X_1, X_3} C_{X_2, X_4} - C_{X_1, X_4} C_{X_2, X_3}$$

مجموعه ۱۲ ویژگی استخراج شده از سیگنال‌ها $C_{R,R}$, $C_{R,I}$, $C_{I,I}$, $C_{R,R,R}$, $C_{R,R,I}$, $C_{R,I,I}$, $C_{I,I,I}$, $C_{R,R,R,R}$, $C_{R,R,R,I}$, $C_{R,R,I,I}$, $C_{R,I,I,I}$, $C_{I,I,I,I}$ هستند، که R و I به ترتیب بیانگر بخش حقیقی و موهومی H_y هستند که از رابطه (۱۱) به دست می‌آید.

$$H_y = [y(t) + j\hat{y}(t)]\exp(-j2\pi f_c t) \quad (11)$$

در فرمول فوق $\hat{y}(t)$ تبدیل هیلبرت $y(t)$ و f_c فرکانس مرکزی است.

۳-۳- ویژگی‌های مبتنی بر تبدیل موجک

ویژگی‌های دیگری که در [3] مورد استفاده قرار گرفته است
ممنت‌های مرتبه بالای تبدیل موجک پیوسته تا مرتبه ۶ است. برای
تبدیل موجک پیوسته از رابطه زیر استفاده می‌شود [7].

پردازش سیگنال‌های مختلف از جمله سیگنال‌های تصویری هم استفاده شده است [7] [8].

بعضی اوقات “ابربردار” به ترکیب تعداد زیادی بردارهای با ابعاد کوچک با هم و حاصل شدن یک بردار با ابعاد بالاتر اطلاق می‌شود؛ برای مثال با پشته‌سازی بردارهای میانگین d بعدی از یک مدل مخلوط گوسی با K مؤلفه و تبدیل به یک اربردار گوسی Kd بعدی [9]. هر چند می‌توان اربردار را به صورت وسیع‌تری، شامل هر گونه ارائه از نمونه‌های گفتاری با ابعاد بالا و ثابت دانست. چیزی که مهم است، سیستم مختصاتی مشترک برای تمامی اربردارهای نمونه‌های گفتاری مختلف است و یا اینکه همگی براساس یک چند جمله‌ای ثابت تولید شده باشند [6]. در این صورت عنصرهای اربردارها در زمان محاسبات مربوط به شباهت در فضای اربردار، به طور معناداری قابل مقایسه خواهند بود.

۳- استخراج و انتخاب ویژگی از فریم

در این مقاله دسته‌بندی نوع مدولاسیون بر اساس ویژگی‌های استخراج شده از سیگنال باند پایه صورت می‌گیرد بنابراین بررسی مسائل مربوط به همزمانی گیرنده و فرستنده در آن مطرح نخواهد شد بلکه با فرض در اختیار داشتن سیگنال در باند پایه به استخراج ویژگی‌ها خواهیم پرداخت. برای هر فریم ۲۷ ویژگی از سیگنال استخراج شده است که شامل سه دسته می‌باشند؛ ویژگی‌های مبتنی بر طیف، ویژگی‌های آماری سیگنال و ویژگی‌های مبتنی بر تبدیل موجک، که در ادامه تشریح خواهند شد.

۳-۱- ویژگی‌های مبتنی بر طیف

ویژگی‌های استفاده شده به شکل زیر انتخاب شده‌اند [10].

۱- ماکزیمم مقدار طیف توان دامنه لحظه‌ای نرمال شده و مرکزی

$$\gamma_{max} = \frac{\max |FFT(a_{cn}(i))|^2}{N_s} \quad (1)$$

که در آن $a_{cn}(i)$ مقدار نرمال شده و مرکزی دامنه لحظه‌ای در زمان $t = i/f_s$ بوده و N_s تعداد نمونه‌هاست.

۲- انحراف معیار، مقدار، مطلق، مولفه غیر خطی، فاز، لحظه‌ای مرکزی

$$\sigma_{ap} = \sqrt{\frac{1}{C} \left[\sum_{a_n(t) > a_t} \phi_{NL}(i)^2 \right] - \left[\frac{1}{C} \sum_{a_n(t) > a_t} |\phi_{NL}(i)| \right]^2} \quad (2)$$

که C تعداد نمونه‌های $\emptyset_{NL}$ و $\emptyset_{NL}$ مولفه‌ی غیرخطی فاز لحظه‌ای است.

۳- انحراف معیار مقدار مستقیم مولفه غیر خطی، فاز لحظه‌ای مرکزی

$$\sigma_{dp} = \sqrt{\frac{1}{C} \left[\sum_{a_n(t) > a_t} \phi_{NL}(i)^2 \right] - \left[\frac{1}{C} \sum_{a_n(t) > a_t} \phi_{NL}(i) \right]^2} \quad (3)$$

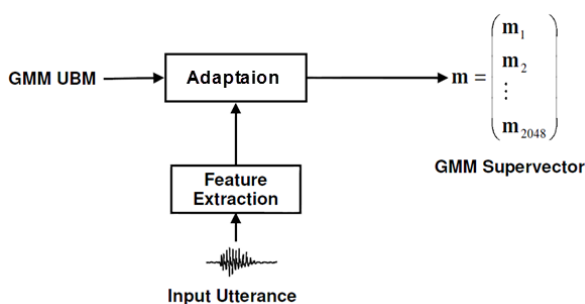
۴- انحراف معیار، مقدار مطلق دامنه لحظه‌ای نرمال شده و مرکزی

داد. در اینجا با الهام از همین روش، برای هر قطعه یک ابربردار گوسی واحد در فضای ابربردار به دست آورد. در روش‌های استفاده شده برای بازشناسی گوینده نمونه‌های گفتاری متعدد برای یک گوینده به یک ابربردار واحد انتقال می‌یابند که در واقع ابربردار آن گوینده خواهد بود، در اینجا هدف ارائه یک ابربردار برای یک قطعه سیگنالی از روی فریم‌های مختلف سیگنال است.

در تحقیق مورد بررسی برای پیاده‌سازی، روش ساخت مرسوم برای مدل‌های گوینده، تطبیق MAP روی میانگین‌های مدل زمینه جهانی معرفی شده است. در این روش فرض می‌شود که یک مدل زمینه جهانی مدل مخلوط گوسی به صورت زیر داریم:

$$g(x) = \sum_{i=1}^N \lambda_i N(x; m_i, \Sigma_i) \quad (13)$$

که λ_i وزن‌های گوسی، $N()$ گوسی‌ها و m_i, Σ_i میانگین و کواریانس گوسی‌ها هستند. سپس تطبیق MAP روی میانگین‌ها را به صورت شمای نشان داده شده در شکل (۱)، در نظر می‌گیرد که نتیجه آن ساخت یک ابربردار مدل مخلوط گوسی است. همان‌طور که در شکل مشاهده می‌شود، ابربردار خروجی، توالی بردارهای میانگین مربوط به هر مؤلفه گوسی در مدل مخلوط گوسی است، که نتیجه تطبیق مدل زمینه با داده جدید است.



شکل (۱): استفاده مدل مخلوط گوسی تطبیق یافته با روش پیشینه‌سازی پسینه به عنوان ابربردار [9]

۴-۲- استفاده مدل مخلوط گوسی تطبیق یافته با روش رگرسیون خطی پیشینه‌سازی شباهت

این روش کاملاً شبیه روش قبل است، با این تفاوت که به جای تطبیق با روش پیشینه‌سازی پسینه، از تطبیق به وسیله رگرسیون خطی پیشینه‌سازی شباهت MLLR^۶ استفاده شده است [12] که در ادامه توضیح داده خواهد شد. دو نوع ابربردار می‌توان از این روش استخراج کرد، که در اینجا از نوع اول یعنی ابربردار مدل مخلوط گوسی استفاده شده است. نوع دوم، ابربرداری در فضای بردار انتقال MLLR است که با توجه به ماهیت استفاده از آن در مقالات مرتبط، مناسب کاربردهای پردازش گفتار است [12].

در تحقیق انجام شده در [12]، از پارامترهای تبدیل MLLR به عنوان ورودی ماشین بردار پشتیبان استفاده شده است. MLLR

$$CWT(a, \tau) = \int_{-\infty}^{+\infty} s(t) \psi_{a, \tau}^*(t) dt \quad (12)$$

که در آن $a > 0$ متغیر مقیاس و $\tau \in R$ متغیر تبدیل است و از موجک‌ها به مقیاس‌های ۱ تا ۸ به عنوان موجک مادر استفاده شده است. ۲ مجموعه ۵ تایی ویژگی در این قسمت استخراج می‌شود، مجموعه اول شامل گشتاورهای مرتبه ۲ تا ۶ ضرایب به دست آمده از تبدیل موجک به طور مستقیم و مجموعه دوم شامل ممتهای مرتبه ۱ تا ۶ ضرایب به دست آمده از تبدیل موجک که یک فیلتر میانه بر روی آن‌ها اعمال شده است. استفاده از مجموعه دوم منجر به افزایش کارایی در حضور نویز خواهد شد. این مجموعه به ترتیب با نمادهای $HOM_{order}(CWT(a, \tau))$ و $HOM_{order}(CWT(a, \tau))_{filtered}$ نمایش داده خواهد شد.

۳-۴- انتخاب ویژگی‌های مناسب

همان‌طور که اشاره شد مجموعه ویژگی‌های استخراج شده، شامل ۲۷ ویژگی مختلف است. با توجه به اینکه زیاد بودن ابعاد علاوه بر افزایش پیچیدگی و زمان اجرا، منجر به کاهش قابلیت تعمیم نیز می‌شود، لذا نیاز به رویکردی داریم تا ابعاد فضای ابربردار را تقلیل دهیم و از بین آن‌ها ویژگی‌هایی را انتخاب کنیم که منجر به افزایش دقت دسته‌بندی شوند و ویژگی‌هایی که اطلاعات مفیدی ارائه نمی‌دهند را حذف نماییم. روش استفاده شده در اینجا روش بکار رفته در [11] است. معیار مورد نظر در انتخاب ویژگی نرخ بهره اطلاعات^۳ است. در این حالت یک رتبه‌بندی بین ویژگی‌ها بر اساس نرخ بهره اطلاعات انجام می‌شود؛ ویژگی‌های با رتبه‌های بهتر اولویت بیشتری برای انتخاب دارند.

۴- ساخت ابربردار مبتنی بر مدل مخلوط گوسی و مدل زمینه جهانی

همان‌طور که پیش‌تر گفته شد، در این تحقیق از روش‌های مبتنی بر تطبیق مدل زمینه مخلوط گوسی برای ساخت ابربردار هر قطعه از سیگنال استفاده شده است. در ادامه به تشریح دو روش به کار رفته برای این منظور خواهیم پرداخت.

۴-۱- استفاده از مدل مخلوط گوسی تطبیق یافته با روش پیشینه‌سازی پسینه

نظر به اینکه مدل زمینه جهانی^۴ به عنوان بخشی از اکثر سیستم‌های بازشناسی گوینده به کار می‌رود، یک راه طبیعی برای ساخت ابربردارها را به وجود می‌آورد [9] و این مسأله منجر به تولید یک دسته‌بندی کننده ترکیبی می‌شود، که مدل زمینه تولید شده براساس مدل مخلوط گوسی برای ساخت بردارهای ویژگی ماشین بردار پشتیبان استفاده می‌شود. می‌توان یک قطع سیگنالی که توسط یک مدولاسیون خاص تولید شده است را توسط یک ابربردار واحد نمایش

حال کرنل واگرایی KL به صورت زیر تعریف می‌شود:

$$K(S_a, S_b) = \sum_{i=1}^N \lambda_i (m_i^a)^t \Sigma_i^{-1} m_i^b \quad (18)$$

$$= \sum_{i=1}^N \left(\sqrt{\lambda_i \Sigma_i^{-\frac{1}{2}}} m_i^a \right)^t \left(\sqrt{\lambda_i \Sigma_i^{-\frac{1}{2}}} m_i^b \right)$$

S_b و S_a نمونه‌های سیگنالی مورد نظر هستند.

۵-۲- کرنل ضرب داخلی L^2 مدل مخلوط گوسی^۱

دومین کرنل ارائه شده با الهام از استفاده از ضرب داخلی فضای تابع است. فرض کنید g_a و g_b مدل‌های مخلوط گوسی باشند که در اختیار داریم و در واقع به وسیله یکی از فرایندهای تطبیق ذکر شده در بخش‌های قبل از دو نمونه سیگنالی S_a و S_b ایجاد شده‌اند. یک ضرب داخلی استاندارد در فضای تابع به صورت زیر است [9]:

$$K(S_a, S_b) = \int_{R^n} g_a(x) g_b(x) dx \quad (19)$$

می‌توان فرم بسته راه حل انتگرال موجود در رابطه قبل را به دست آورد. با استفاده از جایگذاری مولفه‌های گوسی در این رابطه خواهیم داشت:

$$K(S_a, S_b) = \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \int_{R^n} N(x; m_i^a, \Sigma_i) N(x; m_j^b, \Sigma_j) dx \quad (20)$$

که در آن 0 برداری با مقادیر صفر است. از آنجا که هر ترم موجود در رابطه بالا یک کرنل است و جمع چند کرنل، خود یک کرنل خواهد بود، بنابراین رابطه (۲۰) یک کرنل خواهد بود.

یک تخمین مناسب برای رابطه (۲۰) این است که فرض کنیم میانگین‌های مولفه‌های مخلوط متفاوت از یکدیگر دور هستند. این فرض در زمانی که $j \neq i$ برقرار باشد، ترم‌های رابطه فوق را کوچک خواهد کرد. در نتیجه برای کرنل خواهیم داشت:

$$K(S_a, S_b) = \sum_{i=1}^N \lambda_i^2 N(m_i^a - m_i^b, 0, 2\Sigma_i) \quad (21)$$

شکل (۲) شمایی کلی از فرآیند انتقال از فضای سیگنال به ابربردار مدل مخلوط گوسی و آموزش دسته‌بندی کننده را نشان می‌دهد.

۶- دادگان و ارزیابی

در آزمایش‌های انجام شده، استخراج ویژگی از سیگنال‌ها با این پیش‌فرض خواهند شد که فرکانس مرکزی در آن‌ها توسط روش پریودوگرام محاسبه شود، که در اینجا ۱۸۰۰ هرتز در نظر گرفته شده است. برای ایجاد تنوع در داده‌ها، پارامترهای متفاوتی در نظر گرفته

بردارهای میانگین یک مدل غیر وابسته به گوینده را به صورت $\mu'_k = A\mu_k + b$ تبدیل می‌کند، که μ'_k بردار میانگین سازگار شده، μ_k بردار میانگین مدل جهانی و پارامترهای A و b مشخص کننده تبدیل خطی هستند. پارامترهای A و b با بیشینه‌سازی احتمال داده آموزشی به وسیله یک الگوریتم بیشینه‌سازی امید ریاضی EM^۲ تغییر یافته، تخمین زده می‌شوند.

۵- ماشین بردار پشتیبان ابربردار مدل مخلوط

گوسی

آنچه از روش اول ساخت ابربردار مبتنی بر مدل مخلوط گوسی (استفاده از مدل مخلوط گوسی تطبیق یافته با روش بیشینه‌سازی پسینه به عنوان ابربردار) و همچنین روش دوم (استفاده از مدل مخلوط گوسی تطبیق یافته با روش رگرسیون خطی بیشینه‌سازی شباهت) می‌توان برداشت کرد این است که مدل‌ها، تنها در میانگین‌شان با هم اختلاف دارند. در این بخش دو کرنل مناسب برای دسته‌بندی این ابربردارها به کمک ماشین بردار پشتیبان مطرح و بررسی می‌شوند.

۵-۱- کرنل خطی ابربردار مدل مخلوط گوسی^۳

فاصله طبیعی بین دو مدل مخلوط گوسی با توجه به معیار واگرایی KL به صورت زیر است [9]:

$$D(g_a || g_b) = \int_{R^n} g_a(x) \log \left(\frac{g_a(x)}{g_b(x)} \right) dx \quad (14)$$

متأسفانه این معیار واگرایی شرط مرسر را ارضا نمی‌کند، بنابراین استفاده از SVM سخت خواهد شد. حال به جای استفاده مستقیم از این معیار واگرایی یک تخمین را در نظر می‌گیریم [9]. ایده اصلی برای این کار ارائه محدوده برای واگرایی مورد نظر است که به وسیله نامساوی جمع-لگاریتم به صورت زیر تعریف می‌شود:

$$D(g_a || g_b) \leq \sum_{i=1}^N \lambda_i D(N(\cdot; m_i^a, \Sigma_i) || N(\cdot; m_i^b, \Sigma_i)) \quad (15)$$

که در این فرمول ما ابربردارهای میانگین تطبیق یافته مربوط به هر نمونه a و b را با m^a و m^b به ترتیب نشان می‌دهیم. با فرض در نظر گرفتن کواریانس قطری، تخمین موجود در فرمول قبل به صورت فرم بسته زیر می‌تواند محاسبه شود:

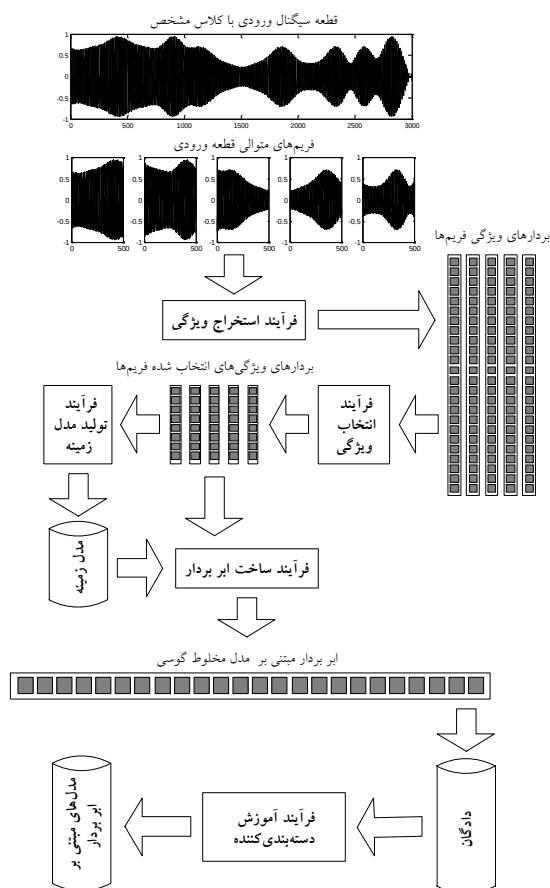
$$d(m^a, m^b) = \frac{1}{2} \sum_{i=1}^N \lambda_i (m_i^a - m_i^b)^t \Sigma_i^{-1} (m_i^a - m_i^b) \quad (16)$$

و در نهایت نامساوی به صورت زیر خواهد آمد:

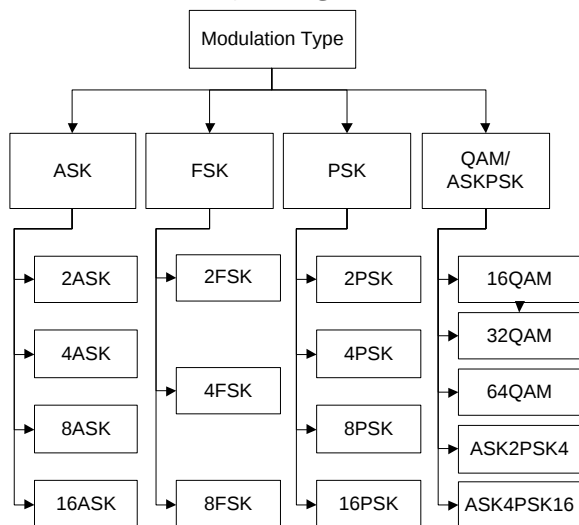
$$0 \leq D(g_a || g_b) \leq d(m^a, m^b) \quad (17)$$

هرچند این معیار فاصله در کاربردهای خوشه‌بندی گوینده مناسب بوده است، ولی در اینجا نیز قابل استفاده است.

دسته‌بندی برون کلاسی و پس از آن یک دسته‌بندی درون کلاسی انجام می‌شود.



شکل (۲): شمایی کلی از فرآیند انتقال از فضای سیگنال به ابر بردار مدل مخلوط گوسی و آموزش دسته‌بندی کننده



شکل (۳): کلاس‌های متنوع مدولاسیون مورد بررسی و ساختار سلسله مراتبی دسته‌بندی کننده

شده است. برای این منظور از نرخ‌های سمبل ۸۰۰، ۶۰۰، ۲۰۰، ۱۰۰ و ۱۸۰۰ باود استفاده شده است. همچنین در مدولاسیون‌های PSK و QAM پارامتر همگرایی طیفی^۴ با مقادیر ۰.۵، ۰.۳۵ و ۰.۱ تنظیم شده است. نرخ‌های سیگنال به نویز مورد بررسی بی‌نهایت، ۲۰، ۱۰، ۵ و صفر دسی‌بل است. آموزش و آزمایش هم با دادگان با نویز مشابه و هم غیر مشابه انجام شده است. طول هر قطعه سیگنال به صورت تصادفی بین ۲ تا ۱۵ ثانیه و طول هر فریم ۰.۱ ثانیه است. همچنین داده مدوله شده نیز از تنوع کافی برخوردار است و به روش‌های متنوع تولید شده است: رشته‌ی دودویی تصادفی، رشته‌های دودویی صفر، رشته‌های دودویی یک، رشته‌های دودویی با بلوک‌های صفر و یک متوالی، رشته‌های دودویی تولید شده از متون (متن‌ها از سایت ویکی‌پدیا استخراج شده‌اند) و رشته‌های دودویی تولید شده از فایل‌های صوتی.

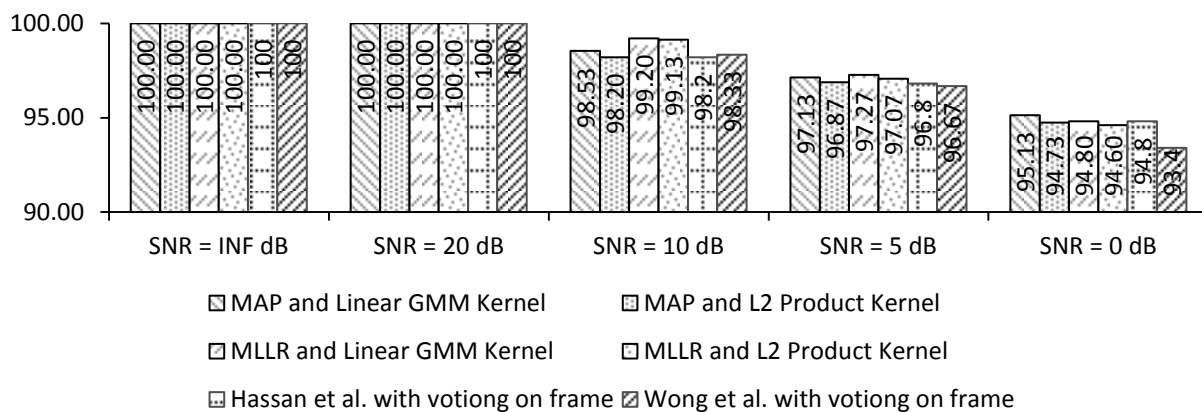
برای هر کلاس مدولاسیون مورد بررسی ۲۵۰ نمونه آموزشی و ۱۵۰ نمونه آزمایشی تولید شده است. شکل (۳) کلاس‌های متنوع مدولاسیون مورد بررسی و ساختار سلسله مراتبی دسته‌بندی کننده را نشان می‌دهد.

شکل (۴) نتایج به دست آمده از این روش را نشان می‌دهد. همان‌طور که مشاهده می‌شود عملکرد این روش در نرخ سیگنال به نویز بالا بسیار مناسب است و همچنین در مقابل نرخ‌های پایین مقاومت مناسبی نشان می‌دهد. در زمان حضور نویز، ویژگی‌های استخراج شده از بعضی از فریم‌ها به دچار خطا می‌شوند و این مسئله در زمان تطبیق با مدل زمینه باعث ایجاد خطای جزئی می‌گردد. در صورتی که طول قطعه مورد نظر کوتاه باشد و ویژگی‌ها هم دچار خطا شده باشند، در زمان تطبیق نه تنها دارای خطا هستیم، بلکه تغییرات چندانی هم در مدل زمینه پس از فرآیند تطبیق صورت نمی‌گیرد. از آنجا که این مسئله در حضور نویز بیشتر نمود پیدا می‌کند، در نرخ‌های پایین سیگنال به نویز شاهد کاهش دقت دسته‌بندی کننده هستیم.

۷- خلاصه و نتیجه‌گیری

در این مقاله روشی برای دسته‌بندی خودکار مدولاسیون دیجیتال مبتنی بر استفاده از ماشین بردار پشتیبان ابر بردار مدل مخلوط گوسی ارائه شد. ایده استفاده از ابر بردارها از کاربرد موفق آن‌ها در بازشناسی گوینده وام گرفته شده است، و در این حوزه از مخابرات هوشمند مورد استفاده قرار گرفته است. تفاوت این روش با سایر روش‌های تشخیص پیشین در این است که تصمیم‌گیری مبتنی بر توالی‌ای از فریم‌هاست و بر روی یک فریم تصمیم اتخاذ نمی‌شود. ساختار دسته‌بندی کننده ارائه شده به صورت سلسله مراتبی است، بدین صورت که یک

^۴ Roll off



شکل (۴): نتایج حاصل از به کارگیری ماشین بردار پشتیبان با کرنل‌های ابربردار مدل مخلوط گوسی مختلف در حضور نویز

- [7] Xiaodan Zhuang, Xi Zhou, Mark A. Hasegawa-Johnson, and Thomas S. Huang, "Real-world acoustic event detection," Pattern Recognition Letters, vol. 31, no. 12, pp. 1543-1551, Sept. 2010.
- [8] Xiaodan Zhuang, Xi Zhou, Mark Hasegawa-Johnson, and Thomas Huang, "Face age estimation using patch-based hidden Markov model supervectors," in 19th International Conference on Pattern Recognition (ICPR 2008), Tampa, FL, 2008, pp. 1-4.
- [9] W.M. Campbell, D.E. Sturim, and D.A. Reynolds, "Support vector machines using GMM supervectors for speaker verification," IEEE Signal Processing Letters, vol. 13, no. 5, pp. 308-311, 2006b.
- [10] MLD Wong and AK Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," Signal Processing, vol. 84, pp. 351-365, 2004.
- [11] Chotirat Ann Ratanamahatana and Dimitrios Gunopulos, "Scaling up the Naive Bayesian Classifier: Using Decision Trees for Feature Selection," in proceedings of IEEE International Conference on Data Mining (ICDM 2002), Maebashi, Japan, 2002.
- [12] Z.N. Karam and W.M. Campbell, "A new kernel for SVM MLLR based speaker recognition," in Proc. Interspeech 2007 (ICSLP), Antwerp, Belgium, 2007, pp. 290-293.

۱۶ کلاس از انواع مدولاسیون‌ها در بررسی‌ها استفاده شده است. همچنین سعی شده علاوه بر تنوع در کلاس‌های مورد بررسی نیز با تغییر پارامترهای مدولاسیون تنوع حفظ شود. نگاهی بر نتایج و تعداد کلاس‌های قابل تفکیک در این مقاله نشان می‌دهد که در این کار تعادلی بین دقت در بازشناسی و افزایش در تعداد کلاس‌ها ایجاد شده است. علاوه بر این کارایی بالای این روش در شرایط حضور نویز مقاومت روش را تصدیق می‌کند.

مراجع

- [1] Lan-Xun Wang, Yu-Jing Ren, and Rui-Hua Zhang, "Algorithm of digital modulation recognition based on support vector machines," vol. 2, 2009, pp. 980-983.
- [2] Syed M Baarrij, Fahad Nasir, and Saad Masood, "A Robust Hierarchical Digital Modulation Classification Technique: using Linear Approximations," 2006, pp. 545-550.
- [3] K. Hassan, I. Dayoub, W. Hamouda, and M. Berbineau, "Automatic Modulation Recognition Using Wavelet Transform and Neural Networks in Wireless Systems," EURASIP Journal on Advances in Signal Processing, vol. 2010, 2010.
- [4] Bijan G. Mobasseri, "Digital modulation classification using constellation shape," Signal Processing, vol. 80, no. 2, pp. 251-277, 2000.
- [5] Tomi Kinnunen and Haizhou Li, "An overview of text-independent speaker recognition: From features to supervectors," Speech Communication, vol. 52, pp. 12-40, 2010.
- [6] W.M. Campbell, J.P. Campbell, D.A. Reynolds, E. Singer, and P.A. Torres-Carrasquillo, "Support vector machines for speaker and languagerecognition," Computer Speech and Language, vol. 20, no. 2-3 SPEC. ISS., pp. 210-229, 2006.

زیرنویس‌ها

- ¹ Principal Component Analysis
- ² Supervector
- ³ Information Gain Ratio
- ⁴ Universal Background Model (UBM)
- ⁵ Maximum A Posteriori (MAP)
- ⁶ Maximum Likelihood Linear Regression (MLLR)
- ⁷ Expectation Maximization (EM)
- ⁸ GMM Supervector Linear Kernel
- ⁹ GMM L^2 Inner Product Kernel