

روش پیشنهادی برای شناسایی خودکار گوشه‌ها در ردیف موسیقی سنتی ایرانی با رویکرد میکرو*

امیر وفائیان^۱

کیوان برنا^۲

حامد ساجدی^۳

داریوش علیمحمدی^۴

پویا سرایی^۵

چکیده

طراحی سامانه‌ای که بتواند به طور خودکار دستگاه‌ها و گوشه‌های موسیقی سنتی ایرانی را تشخیص دهد، مدتی است که توجه پژوهشگران ایرانی را در حوزه بازیابی اطلاعات موسیقایی به خود جلب کرده است. از آنجا که در این پژوهش‌ها، به طور عمده به دسته‌بندی خودکار پنج دستگاه اصلی پرداخته‌اند و نه تفکیک گوشه‌ها از یکدیگر، پژوهش حاضر با هدف ارائه الگوریتم و روشی برای شناسایی خودکار گوشه‌ها در ردیف دستگاهی موسیقی سنتی ایرانی انجام شده است. از این رو، پس از معرفی دو رویکرد اساسی در شناسایی خودکار دستگاه و گوشه، روش شناسایی و الگوریتم‌های به کار رفته در برخی از پژوهش‌های مرتبط آمده است. سپس، الگوریتم و روشی برای شناسایی خودکار گوشه‌های ردیف دستگاهی موسیقی سنتی ایرانی بر اساس رویکرد میکرو ارائه شده است. در روش پیشنهادی، در درجه نخست تلاش برای شناسایی نغمه‌های یک قطعه موسیقی سنتی ایرانی، به عنوان مرحله اولیه و مهم در شناسایی گوشه‌ها و بهبود پارامترهای آن بوده است. پس از این مرحله، با تعریف سیر ملودی جمله معرف گوشه‌ها و بررسی میزان شباهت آن به کل قطعه، می‌توان گوشه موردنظر را شناسایی کرد.

واژه‌های کلیدی: الگوریتم شناسایی گوشه، بازیابی اطلاعات موسیقایی، ردیف دستگاهی موسیقی سنتی ایرانی، شناسایی خودکار موسیقی، گوشه.

* تاریخ دریافت: ۹۷/۹/۱؛ تاریخ پذیرش: ۹۷/۱۰/۲۵.

amirvafa@gmail.com

^۱ دکتری علم اطلاعات و دانش‌شناسی، دانشگاه خوارزمی، تهران، ایران (نویسنده مسئول)

borna@khu.ac.ir

^۲ استادیار، گروه علوم کامپیوتر، دانشکده علوم ریاضی و کامپیوتر، دانشگاه خوارزمی، تهران، ایران

sadjedi@shahed.ac.ir

^۳ استادیار، گروه برق و الکترونیک، دانشکده فنی و مهندسی، دانشگاه شاهد، تهران، ایران

^۴ استادیار، گروه علم اطلاعات و دانش‌شناسی، دانشکده روانشناسی و علوم تربیتی، دانشگاه خوارزمی، تهران، ایران

webliographer@gmail.com

pooya_saraee@yahoo.com

^۵ استادیار، گروه موسیقی، دانشکده هنر، دانشگاه آزاد واحد تهران مرکزی، تهران، ایران

مقدمه

تقریباً همه استادان و اهل فن، ردیف موسیقی سنتی ایرانی را الفبای آهنگسازی ایرانی دانسته و آن را به هنرجویان و نوازندگانی که قصد دارند این نوع موسیقی را به طور جدی دنبال کنند، آموزش می‌دهند. از این رو، افرادی که قصد دارند به صورت حرفه‌ای موسیقی سنتی ایرانی را فراگیرند، لازم است به طور دقیق و عمیق با دستگاه‌ها و گوشه‌های ردیف موسیقی سنتی ایرانی آشنا شوند. از این رو، سیستمی که بتواند دستگاه‌ها و گوشه‌ها را بر اساس ویژگی‌های اساسی موسیقایی ایرانی تشخیص دهد، انگیزه اصلی برای انجام چنین پژوهش‌هایی بوده است.

هفت دستگاه موسیقی ایرانی در کنار پنج آواز، اساس و پایه موسیقی در ایران بوده و می‌توان آن را موسیقی رسمی کشور دانست. امروزه، چکیده و برآیند فرهنگی این نوع موسیقی، در قالب ردیف موسیقی دستگاهی، به معنای رپرتوار مجموعه^۱ تجلی یافته است (بهنام‌نیا، ۱۳۸۹: ۵). به عبارت دیگر، در کنار هفت دستگاه، ردیف موسیقی کنونی ایران دارای پنج آواز نیز هست که از آن‌ها با عنوان ملحقات یا متعلقات نیز یاد شده است. از دستگاه شور، چهار شعبه یا آواز به نام‌های ابوعطا، افشاری، بیات ترک و دشتی منشعب می‌شود. از دستگاه همایون هم آواز بیات اصفهان منشعب می‌شود (فخرالدینی، ۱۳۹۲: ۹۹). لازم به ذکر است که در تقسیم‌بندی دستگاه‌ها و آوازها از نظر تعداد و مرز بین آن‌ها، بین نظریه پردازان و موسیقی‌دان‌ها اتفاق نظر وجود ندارد.^۲

^۱ Repertoire

^۲ در تقسیم‌بندی تعداد دستگاه‌ها و آوازها اتفاق نظر وجود ندارد. در این بین می‌توان حداقل سه دیدگاه مختلف را مطرح کرد. افرادی چون روح‌الله خالقی و مرتضی حنانه (حنانه، ۱۳۶۷، ص ۱۲ و ۴۳) قائل به دسته‌بندی هفت دستگاه و پنج آواز در ردیف موسیقی سنتی ایرانی هستند و این دیدگاه امروزه نیز متداول و رایج شده و در ردیف اساتیدی چون حسینقلی، میرزا عبدالله و درویش‌خان نیز به کار رفته است. با این حال، اساتیدی نیز دسته‌بندی غیر از دسته‌بندی فوق قائل هستند. مثلاً علی‌نقی وزیری، موسیقی سنتی ایرانی را در پنج دستگاه خلاصه می‌کند. عده‌ای دیگر (از جمله دکتر هرمز فرحت) نیز، قائل به آواز و شعبه‌های فرعی دستگاه‌ها نبوده و هر یک از مقام‌های ابوعطا، بیات ترک، افشاری، دشتی و بیات اصفهان را دستگاهی مستقل به‌شمار می‌آوردند. از این رو، تعداد دستگاه‌ها را دوازده ذکر کرده‌اند.

اصطلاح دستگاه در موسیقی ایرانی را می‌توان با واژه گام دیاتونیک یا بالفعل^۱ در موسیقی غربی، راگا یا راگ^۲ در موسیقی هندی و نیز واژه مقام (مقام^۳ در موسیقی ترکی و مکام^۴ در موسیقی عربی) مترادف دانست. در تعریف موسیقی غربی، توالی هشت نغمه (نت) پیایی که دارای نسبت فواصل معینی از یکدیگر هستند را گام می‌گویند. برای ساده کردن مفهوم دستگاه در موسیقی ایرانی، همین تعریف را برای هفت دستگاه اصلی موسیقی ایرانی (ماهور، شور، سه‌گاه، چهارگاه، همایون، نوا و راست‌پنجگاه) نیز به کار برده‌اند، با این تفاوت که در موسیقی سنتی ایرانی، در اغلب گوشه‌های ردیف (ملودی‌های کوتاه)، به جای هشت نغمه، توالی ۴ یا ۵ نغمه در نظر گرفته می‌شود؛ اما در تعریف دقیق‌تر، دستگاه عبارت است از یک سیکل یا چرخه چندمدی که از یک مد (مقام) شروع و به همان مد ختم می‌شود و در مسیر خود از مدهای (مقام‌های) متعددی عبور می‌کند (اسعدی، ۱۳۸۲: ۵۲؛ مبانی نظری موسیقی ایرانی، ۱۳۸۸: ۳۴).

امروزه در موسیقی ایرانی، اصطلاح مقام معادل مفهوم مد استفاده می‌شود. مدها طی قرن‌ها و بر اساس زیبایی‌شناختی جمعی فرهنگ خاصی شکل گرفته‌اند. وقتی به یک یا چند نغمه در یک گام بالفعل، اهمیت بیشتری نسبت به دیگر نغمه‌ها در همان گام می‌دهیم، می‌توان گفت که یک مد ساخته‌ایم. به عبارت دیگر، با داشتن الف. یک گام بالفعل یا اشل صوتی (متشکل از یک یا دو دانگ^۵)، ب. یک نظام سلسله‌مراتب درجه‌ها (یا کارکرد درجات) و ج. یک یا چند سرمشق یا الگوی لحنی یا ملودیک خاص (ملودی مدل) که در قطعات و گوشه‌های مختلف تکرار می‌شوند، می‌توان یک مد به وجود آورد. بعضی موسیقی‌های دنیا فقط یکی از این سه عنصر را برجسته می‌کنند؛ درحالی‌که در موسیقی ایرانی، برای به وجود آوردن مقام، تقریباً هر یک از این سه عنصر مهم هستند (مبانی نظری موسیقی ایرانی، ۱۳۸۸: ۳۰، ۳۱ و ۳۶ و اسعدی، ۱۳۸۲: ۴۷).

1 Diatonic Scale
2 Raga/Raag
3 Mugam
4 Maqam
5 Tetra chord

شکل گیری موسیقی ایرانی از خرد به کلان است. بدین معنی که مجموع چند نغمه، یک انگاره (فیگور)^۱ را می‌سازند و چند انگاره یک جمله^۲ را، جمله‌ها گوشه‌ها را و گوشه‌ها دستگاه را. از این رو، با کنار هم و در ارتباط با هم قرار دادن چند جمله، گوشه پدید می‌آید (طلایی، ۱۳۹۴: ۱۲-۱۳). هر یک از دوازده دستگاه و آواز موسیقی ایرانی، از گوشه‌های متعددی تشکیل شده است که با نظمی خاص در کنار یکدیگر قرار گرفته‌اند. مثلاً برخی از گوشه‌های مهم دستگاه شور عبارت‌اند از: شهناز، قرچه، حسینی و رضوی. در واقع، گوشه‌ها کوچک‌ترین و در عین حال، اساسی‌ترین جزء هویت‌دار ثبت شده از ردیف موسیقی ایرانی به حساب می‌آیند و می‌توان آن‌ها را مدهای نسبتاً کوتاه یا مترادف با ملودی مدل^۳ دانست. هر گوشه، آهنگ مستقلی است که به گرد یک نغمه معین شکل می‌گیرد که شاهد آن گوشه نامیده می‌شود. چنانچه تعدادی از این گوشه‌ها که از لحاظ اشل و ترتیب فواصل صوتی با یکدیگر مشابه‌اند، بر اساس نظم خاص پشت سر هم قرار گیرند، مجموعه آن‌ها یک دستگاه را به وجود می‌آورد (صفوت و کارن، ۱۳۹۱: ۱۳۵).

بنابراین، این آهنگ‌ها یا گوشه‌ها معمولاً به ترتیب از پائین به بالا، پشت سر هم چیده می‌شوند. بدین معنا که نغمه شاهد به تدریج از طرف بم به طرف زیر جابجا می‌شود. آهنگی که وسعت صوتی آن از همه پایین‌تر است به عنوان اولین آهنگ قرار می‌گیرد و «درآمد» نامیده می‌شود. مثلاً در درآمد و گوشه‌های اولیه دستگاه شور، نغمه سل به عنوان نغمه شاهد و نخستین درجه تسلسل خواهد بود. سپس آهنگ‌ها و گوشه‌های دیگر به دنبال درآمد یا درآمدها قرار گرفته و به ترتیب ردیف می‌شوند و «دستگاه» یا «مقام» شور را به وجود می‌آورند. در هر یک از این گوشه‌ها، به تدریج و معمولاً به ترتیب، نغمه شاهد از درجه اول بر روی درجات بالاتر (دوم، سوم و بالاتر) انتقال داده می‌شود و در انتهای دستگاه، دوباره به درجه اول بازگشت (فرود) پیدا می‌کند (صفوت و کارن، ۱۳۹۱: ۱۳۵).

^۱ به انگاره اصلی در موسیقی کلاسیک غربی، که کوچک‌ترین واحد دارای مفهوم در موسیقی است، «موتیف» (Motif یا motive) نیز گفته می‌شود.

^۲ Phrase

^۳ Melody Model

به دیگر سخن، در ردیف‌نوازی روش کلی این است که گردش ملودی ابتدا در منطقه پائین (حضیض) یا درآمد با محوریت «مایه» یا «تنالیت»^۱ اصلی (معمولاً درجه اول گام) شروع می‌شود و بعد با گوشه‌های دیگر درجات گام به ترتیب پله پله بالا رفته و بسط و پرورش می‌یابد و به درجه هشتم (اوج) می‌رسد و سپس به «مایه اصلی» فرود می‌آید. گوشه‌ها هم معمولاً بعد از گردش لازم به «مایه اصلی» فرود می‌آیند تا وحدت و یگانگی دستگاه با مایه اصلی حفظ شود (فخرالدینی، ۱۳۹۲: ۳۰).

دو رویکرد اساسی در شناسایی خودکار دستگاه و گوشه در موسیقی سنتی ایرانی

تفاوت گوشه‌ها با یکدیگر، به تفاوت آن‌ها در هر یک از سه مؤلفه مد اشاره دارد. بدین معنا که گوشه‌ها حداقل در یکی از سه مؤلفه اشل صوتی، نقش درجات و یا الگوهای (فرمول‌های) ملودیک با یکدیگر تفاوت دارند. می‌توان برای این مؤلفه‌ها، سلسله‌مراتبی قائل شد که در آن اشل صوتی بیشترین اهمیت و الگوهای ملودیک از کمترین اهمیت برخوردارند. از این رو است که گوشه‌هایی که دارای اشل صوتی مشترکی هستند، گام یا دستگاه مجزایی را تشکیل می‌دهند. گوشه‌هایی که دارای اشل صوتی یکسانی هستند، نقش درجات آن‌ها با یکدیگر متفاوت خواهد بود. مثلاً تفاوت گوشه شهناز، رضوی و حسینی با یکدیگر (که هر سه در دستگاه شور و دارای اشل صوتی یکسانی هستند)، در این است که نغمه شاهد آن‌ها متفاوت بوده و در مایه شور سل، به ترتیب نغمه‌های دو، فا و سل نقش نغمه شاهد را ایفا می‌کنند.

به منظور بررسی الگوی لحنی و ملودیک در هر گوشه، طلایی (۱۳۹۴: ۴۱-۳۶) به تجزیه و تحلیل ردیف میرزا عبدالله (۱۲۲۲-۱۲۹۷ ش.) که در حال حاضر رایج‌ترین ردیف آموزشی برای سازها بوده و نزدیک به ۲۵۰ گوشه دارد، پرداخته است. بر این اساس، یک گوشه کامل را به ترتیب به جمله‌ها و بخش‌های آغازین، معرف، گسترشی، تکمیلی، پایانی

و ختم قابل تجزیه و تفکیک می‌داند. در این تبیین از آناتومی گوشه، جمله معرف مهم‌ترین بخش گوشه است. این ویژگی انحصاری می‌تواند الگویی ریتمیک باشد، مثل کرشمه؛ یا ملودی‌های خاصی باشد، مثل لیلی و مجنون یا غم‌انگیز یا می‌تواند تحریر خاصی باشد، مثل جوادخانی.

بر این اساس و با بررسی پژوهش‌های مشابه، می‌توان دو رویکرد اساسی در شناسایی خودکار دستگاه و گوشه در موسیقی سنتی ایرانی متصور شد: پژوهش‌هایی که صرفاً بر اساس اشل صوتی و فواصل در پنج دستگاه اصلی، سعی در تفکیک و شناسایی خودکار این دستگاه‌ها از یکدیگر داشته‌اند (رویکرد ماکرو^۱). می‌توان گفت که عمده پژوهش‌هایی که تاکنون منتشر شده‌اند، این رویکرد را انتخاب کرده‌اند. رویکرد دوم، شناسایی خودکار گوشه‌های مختلف در ردیف دستگاهی و قطعات ساخته‌شده در موسیقی سنتی ایرانی است (رویکرد میکرو^۲). تا حدود زیادی می‌توان برای گوشه‌ها اصالت قائل بود و از آنجا که ملودی مدل اصلی در هر گوشه، در جمله معرف آن گوشه تجلی می‌یابد، بر اساس جمله معرف هر گوشه که در یکی از سه حالت سیر ملودی، الگویی ریتمیک یا تحریر، مختص و متمایزند، می‌توان آن‌ها را از یکدیگر بازشناخت.

در این پژوهش سعی شده پس از بررسی روش‌شناسی و الگوریتم‌های ارائه شده در دیگر پژوهش‌های مرتبط با شناسایی خودکار دستگاه‌ها، الگوریتم و روش پیشنهادی برای شناسایی خودکار گوشه‌های ردیف دستگاهی موسیقی ایرانی ارائه شود.

پیشینه پژوهش

روش‌شناسی پژوهش‌های مرتبط

از حیث معرفت‌شناسی، بیشتر پژوهش‌های انجام‌شده در حوزه شناسایی خودکار گوشه‌ها و دستگاه‌ها در موسیقی سنتی ایرانی را می‌توان عینی‌گرا دانست. این پژوهش‌ها در تفسیر نتایج، تابع رویکرد استقرایی بوده و از نظر امکان استفاده از نتایج، عموماً کاربردی به

¹ Macro Approach

² Micro Approach

حساب می‌آیند؛ زیرا یافته‌های آن را می‌توان در توسعه فناوری شناسایی خودکار موسیقی سنتی ایرانی مؤثر دانست. راهبرد حاکم بر آن‌ها، روش شناسی پژوهش در عملیات است. پژوهش در عملیات، به زبان ساده، رفتار پدیده‌ها و سامانه‌ها را مطالعه می‌کند تا بتواند تغییرات آن‌ها را برنامه‌ریزی و کنترل کند. رویکرد غالب پژوهش‌ها نیز از نوع کمی بوده و ناظر بر استفاده از قوانین علمی برای تصمیم‌گیری است. باید توجه داشت که انتخاب ویژگی‌های صوتی که بتواند منجر به دسته‌بندی خودکار دستگاه‌ها و گوشه‌ها شود، از اهمیت ویژه‌ای برخوردار است.

الگوریتم و فرایند اجرا در پژوهش‌های مرتبط

پیش فرض پژوهش‌ها در تشخیص هوشمند یا خودکار دستگاه، اساساً این بوده است که هر نغمه موسیقی، فرکانس خاص خود را داشته و از طرف دیگر، امکان ارائه الگوی منحصر به فرد برای هر یک از دستگاه‌های موسیقی سنتی ایرانی (بر اساس فواصل فرکانسی منحصر به فرد در هر دستگاه) وجود دارد.

الگوریتم‌هایی که معمولاً در شناسایی خودکار موسیقی غربی به کار گرفته می‌شوند، به دلیل ماهیت متفاوت آن با موسیقی شرقی به طور عام و موسیقی ایرانی به طور خاص، عموماً قابل تعمیم نیستند و از این رو، شاید نتوان نتایج آن را به موسیقی سنتی ایرانی تعمیم داد. برای مثال، برخی الگوریتم‌های نظارت شده یادگیری ماشین^۱، مانند پرسپترون چندلایه^۲، نزدیک‌ترین مجاور^۳ و ماشین‌های بردار پشتیبان که معمولاً در دسته‌بندی سبک (ژانر) موسیقی غربی از آن‌ها استفاده می‌شود، برای شناسایی خودکار دستگاه‌ها و گوشه‌ها در موسیقی ایرانی نتایج مطلوبی نخواهد داشت.

فرایند و الگوریتم اجرای پژوهش‌های حوزه شناسایی و دسته‌بندی خودکار دستگاه‌های موسیقی سنتی ایرانی، معمولاً بین پنج تا هفت مرحله انجام شده است که بر اساس ترتیب زمانی، به برخی از مهم‌ترین آن‌ها اشاره می‌شود:

¹ Supervised Machine Learning Algorithms

² Multi-layer perceptron (MLP)

³ k-nearest neighbor (KNN) classifier

دارابی (۱۳۸۲): ۱. دریافت سیگنال قطعه موسیقی، پنجره‌بندی و اعمال تبدیل فوریه سریع بر روی هر پنجره به منظور به دست آوردن طیف و محتوای فرکانسی سیگنال ورودی. ۲. نگاشت طیف فرکانسی به بازه فرکانسی یک اکتاو از طریق جمع جبری تمامی داده‌ها ۳. تحلیل دامنه طیف و به دست آوردن فرکانس قله‌های اصلی^۱ (فرکانس‌های غالب به کار رفته در قطعه به عنوان نغمه‌های اصلی) و حذف داده‌های حشو^۲ و فرکانس قله‌هایی که از یک حد و آستانه انرژی کمتر است ۴. مرتب کردن و قراردادن نغمه‌ها و قله‌های اصلی در آرایه‌ای به نام توالی ملودیک ۵. ساختن آرایه فواصل فرکانسی بر اساس فواصل قله‌های طیف اکتاو به دست آمده ۶. تطبیق آرایه نرمال شده با آرایه‌های بانک اطلاعاتی تعریف شده^۳ برای دستگاه موسیقی، جهت نسبت دادن آرایه به دست آمده از قطعه به نزدیک‌ترین دستگاه. به عبارت دیگر، ابتدا آرایه‌ای عددی از فواصل موسیقی یک قطعه موسیقی استخراج و سپس با هر یک از آرایه‌های دستگاه‌ها مقایسه شده است. نزدیک‌ترین و شبیه‌ترین آرایه به یکی از این دستگاه‌ها، به عنوان خروجی موردنظر (نام یکی از دستگاه‌ها) در نظر گرفته شده است.

گواهیان جهرمی (۱۳۸۹): ۱. محاسبه اسپکتروگرام (تبدیل CQT)^۴ برای سیگنال صوتی، با توجه به رزولوشن مورد نیاز برای موسیقی سنتی ایرانی (ده بین برای هر نغمه ربع پرده‌ای) ۲. تبدیل ماتریس اسپکتروگرام (از مرحله قبل) به بردار از طریق یک مرحله کاهش بعد (زمان): در این مرحله تنها بین‌های فرکانسی^۵ و درصد حضور هر یک از آن‌ها در اسپکتروگرام نشان داده شده است، بدون آن که هیچ‌گونه وابستگی به زمان و تغییرات آن داشته باشد ۳. تنظیم پیچ یا نواک^۶: تعیین نوع کوک و اختلاف آن با کوک دیپازن. در این مرحله، پس از استخراج پیک‌های بردار، با تعیین پیچ مرجع (بر اساس کوک دیپازن)، بین‌های فرکانسی در موقعیت‌های تعریف شده، برای فواصل تعدیل شده قرار گرفته است.

¹ Peak detection

² Redundant Data

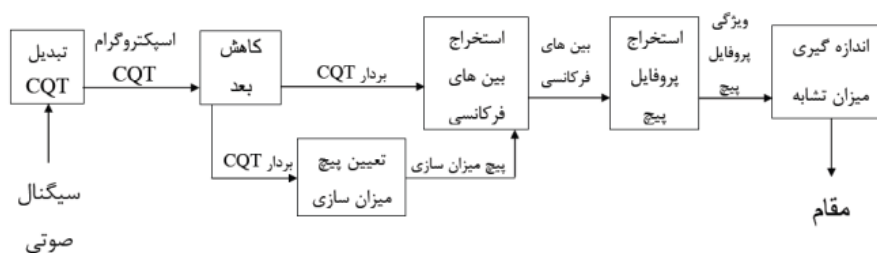
³ Array comparator

⁴ Constant Q-Transform

⁵ Bins

⁶ Tuning Pitch

به عبارت دیگر، پس از تعیین میزان اختلاف با پیچ استاندارد، همه بین‌های فرکانسی، به میزان مشخص تغییر موقعیت داده شده است ۴. استخراج آن دسته از بین‌های فرکانسی که احتمال بیشتری در تشکیل گام/دستگاه قطعه دارند از بردار CQT. در این مرحله همچنین، از تکنیک فیلترینگ ساده بر روی بردار انجام گرفته است تا از تداخل نغمه‌های مجاور و میزان نویزی بودن سیگنال کاسته شود ۵. تشکیل پروفایل پیچ بر اساس بین‌های فرکانسی استخراج شده از مرحله قبل به منظور بررسی خواص مشابه در کلاس‌های (دستگاه‌های) مختلف موسیقی ایرانی ۶. محاسبه میزان شباهت بین پروفایل پیچ استخراج شده با هر یک از چهار الگوی فواصل گام در موسیقی ایرانی، به منظور ارزیابی کارایی تشخیص دستگاه با استفاده از فاصله مینکفسکی^۱.



شکل ۱. الگوریتم کلی تشخیص مقام برای موسیقی ایرانی (گواهیان جهرمی، ۱۳۸۹: ۷۴؛ با الگو گرفتن از مقاله (Zhu, Y. & Kankanhalli, M.S, 2006)

محمودان (۱۳۹۰): ۱. انتخاب ۲۰ قله برتر که بیانگر فرکانس‌های غالب قطعه است: استفاده از ویژگی نواک در مرحله پیش‌پردازش ۲. اعمال فرکانس‌های نظیر به قله‌های به دست آمده از مرحله قبل در قالب ماتریس ورودی با ابعاد $۱۳۵ * ۲۰$ (به عنوان داده‌های آموزش شبکه) ۳. آموزش داده‌ها با استفاده از مدل توابع پایه شعاعی ۴. تفکیک داده‌های آموزشی و آزمون (به نسبت ۷۰ به ۳۰ درصد) و ارزیابی شبکه بر روی داده‌های آزمون ۵. دسته‌بندی و تشخیص دستگاه ماهور از غیر ماهور.

بیات (۱۳۹۲): ۱. اعمال تبدیل فوریه سریع بر روی طیف کل سیگنال ورودی به منظور به دست آوردن دیدی کلی نسبت به قطعه مورد بررسی ۲. پنجره‌بندی قطعات به طول ۱۰۰

¹ Minkowski Distance

میلی ثانیه و اعمال تبدیل فوریه سریع (محاسبه تخمین طیف پریودوگرام) بر روی هر پنجره، جهت محاسبه فرکانس نغمه‌ها در هر پنجره ۳. ذخیره‌سازی فرکانس نغمه‌های به دست آمده از هر پنجره زمانی در سطرهای یک ماتریس ۴. حذف نغمه‌های تکراری پشت سر هم (به عنوان ریز که در فراوانی تعداد کل نغمه‌ها لازم است یک نغمه در نظر گرفته شود) و ذخیره آن در بردار جدید ۵. محاسبه میانگین بردار و مرتب‌سازی نغمه‌ها بر حسب تعداد (فراوانی) از بزرگ به کوچک ۶. مقایسه چهار یا پنج نغمه اول قطعه (که بیشترین فراوانی را داشته‌اند) با تراکورد یا پنتاکوردهای نظیر هر گوشه (بر اساس مبانی نظری موسیقی ردیف ایرانی) یا از طریق شناسایی قله به روش شابلون‌گذاری ۷. شناسایی گوشه و دستگاه با استفاده از روش حداقل فاصله و نیز روش شبکه عصبی (تشکیل بردار ویژگی بر اساس الف. بردار پنتاکورد و در نظر گرفتن تنها پنج نغمه پرتکرار در هر قطعه ب. بردار فراوانی برای کلیه نغمه‌های به کار رفته در قطعه موسیقی).

عبدالله‌زادگان (۱۳۹۳): ۱. استخراج (و به دست آوردن) پوش دامنه تمام نغمه‌های بالقوه دو اکتاو برای ساز تار (۳۶ نغمه: از نغمه $F3$ با فرکانس ۱۸۰ هرتز تا نغمه $F5$ با فرکانس ۶۹۶ هرتز. بر این اساس، هر یک از این ۳۶ نغمه، به عنوان یک ویژگی صوتی در نظر گرفته شده است) بر اساس کوک دیاپازن، به منظور یافتن فرکانس و مشخص کردن محل ضربه خوردن مضراب بر روی سیم (یافتن فازهای حمله در طول یک قطعه موسیقی، چرا که قله‌های حاصل از لحظه ضربه خوردن مضراب، می‌تواند بیانگر زمان و فرکانس نغمه‌های نواخته شده در یک قطعه از موسیقی باشد) ۲. پنجره‌بندی سیگنال حول نقاط به دست آمده از پوش دامنه، به طوری که در یک پنجره، مرحله حمله و آزادسازی مربوط به نواخته شدن یک نغمه را در بر می‌گیرد ۳. یافتن فرکانس پایه (غالب) در هر یک از پنجره‌ها ۴. اعمال تابع خودهمبستگی بر روی هر یک از پنجره‌ها به منظور محاسبه هرچه دقیق‌تر فرکانس‌های پایه نغمه‌ها در طول قطعه ۵. استخراج گام و دستگاه: بر اساس فواصل توالی نغمه‌های به دست آمده از مرحله قبل در یک قطعه موسیقی نواخته شده (و در نظر گرفتن نغمه‌ای که از بین ۳۶ نغمه به تعداد بیشتری در قطعه ظاهر شده به عنوان نغمه شاهد)،

گام و دستگاه موسیقی تشکیل شده و سپس میزان شباهت آن با بردار فرکانسی هر یک از پنج گام اصلی دستگاه‌ها مقایسه شده است.

پیوندی (۱۳۹۴): ۱. تقسیم ۳۰ ثانیه اول سیگنال موسیقی، به پنجره‌هایی با طول ۳۰ میلی ثانیه (به تبعیت از پایگاه داده GTZAN) ۲. استخراج چهار ویژگی (ضرایب کپسترال مل، نرخ عبور از صفر، رول طیفی و شار طیفی) در هر پنجره به طور مجزا ۳. تشکیل بردار ویژگی ۴. آموزش هر یک از هفت دستگاه موسیقی ایرانی یا ده سبک موسیقی غربی بر اساس مدل مخفی مارکوف^۱ ۵. استفاده از الگوریتم Baum-Welch جهت یافتن مدل بهینه ۶. ارزیابی سیستم با استفاده از روش اعتبارسنجی ده تکه‌ای^۲.

عبدلی (۲۰۱۱): ۱. مرحله پیش‌پردازش، شامل الف. حذف نویز و فرکانس‌های غیرمرتبط: پس از تبدیل برخی قطعات از نوار کاست خوانندگان قدیمی موسیقی سنتی ایرانی به فرمت دیجیتالی، با استفاده از الگوریتم SWIPE نویزهای سفید^۳ که به آن‌ها غیرنواک^۴ گفته می‌شود، از نواک‌ها و نغمه‌های اصلی قطعه تفکیک شده‌اند ب. شناسایی نواک ج. اعمال الگوریتم خوشه‌بندی کسری برای حذف فرکانس‌های زائد در آواز^۵. نگاهت نغمه‌ها به یک اکتاو^۶ و سپس تبدیل فواصل فرکانسی به مقیاس سنت: به منظور سهولت فرایند دسته‌بندی و شناسایی دستگاه‌ها ۳. مرحله پس از خوشه‌بندی^۷: با استفاده از فاصله ماهالانوبیس^۸، نقاط اکتاو مرجع، به نغمه مربوطه نسبت داده شده‌اند ۴. استفاده از

¹ Hidden Markov model (HMM)

² 10-fold Cross Validation

³ White Noises

⁴ Unpitch

^۵. یکی از چالش‌های این پژوهش، انتخاب قطعات آوازی (به جای قطعات سازی) بوده است؛ چرا که ادا کردن و تولید نغمات با نواک و فرکانس معین در آواز، سخت‌تر از ساز است. عبدلی در پژوهش خود نشان داده که خواننده از لحظه‌ای که شروع به ادا کردن نغمه می‌کند تا لحظه‌ای که می‌خواهد به فرکانس مورد نظر برسد، چند میلی‌ثانیه زمان می‌برد. به عبارت دیگر، در این چند میلی‌ثانیه، فرکانس‌هایی غیر از نغمه اصلی ایجاد می‌شوند که ضروری است حذف یا نادیده گرفته شوند. از طرف دیگر و به همین منوال، در انتهای هر نغمه نیز فرکانس‌های غیرمرتبط تولید می‌شوند. از این رو، برای حذف این فرکانس‌های زائد (حشو)، از روش خوشه‌بندی کسری^۵ استفاده شده است (عبدلی، ۲۰۱۱، ص ۲۷۶).

⁶ Folding Notes

⁷ Post-Clustering

⁸ Mahalanobis Distance

فازبند^۱ یا طبقه‌بند جهت مدیریت عدم قطعیت بین هر یک از فواصل نغمات مستخرج از داده‌ها (قطعات اجراشده) و داده‌های نظری: بر این اساس، ابتدا مرز هر یک از نغمات مشخص شده (در این پژوهش ۶۷ سنت در نظر گرفته شده است) و سپس با استفاده از «مقیاس شباهت فازی»، میزان شباهت بین الگوی برگرفته از مبانی نظری (که در این پژوهش با عنوان نمونه اولیه ذکر شده) و الگوهای ناشناخته (اطلاعات مستخرج از فایل‌های صوتی قطعات) محاسبه شده است. ۵. شناسایی دستگاه بر اساس بیشترین شباهت بین مجموعه‌های فازی فاصله نوع دوم قطعه موسیقی^۲ و الگوی نظری فواصل آن دستگاه.

عباسی لایق؛ حقی‌پور و نجفی سارم (۲۰۱۳): ۱. بخش‌بندی و جداسازی ۲۰ ثانیه اول فایل صوتی هر یک از گوشه‌ها^۳. ۲. نرمال‌سازی^۴ این فایل‌های ۲۰ ثانیه‌ای، به منظور یکدست کردن و همگن ساختن^۵ داده‌های ورودی برای طبقه‌بند^۶. ۳. کاهش تعداد نمونه^۷ با ضرب چهار ۴. محاسبه مقادیر هر یک از چهار ویژگی صوتی با طول پنجره ۳ ثانیه و ایجاد بردار ویژگی ۵. استفاده از پنجره همینگ^۷ با ۹۰ درصد همپوشانی پنجره‌ها ۶. اعمال طبقه‌بند ماشین بردار پشتیبان بر روی بردار ویژگی.

بیگ‌زاده و بلالی کوچ اصفهانی (۲۰۱۶): ۱. استخراج ویژگی نواک و طیف فرکانسی قطعه نواخته شده با استفاده از تبدیل فوریه سریع ۲. نرمال‌سازی داده‌ها در بازه (+۱، -۱). ۳. تقسیم بازه فرکانسی به فواصل بیست‌تایی ۴. تخصیص بیشترین مقدار هر یک از فواصل به عنوان نماینده آن فاصله ۵. آموزش شبکه عصبی بر روی نتایج (خروجی) به دست آمده از تحلیل تبدیل فوریه سریع با روش پرسپترون چندلایه. روش و الگوریتم به کار رفته در این پژوهش، مشابه پژوهش محمودان (۱۳۹۰) است و به نظر می‌رسد از آن الگو گرفته است.

¹ Fuzzifier

² IT2FSs

³ Segmentation

⁴ Normalization

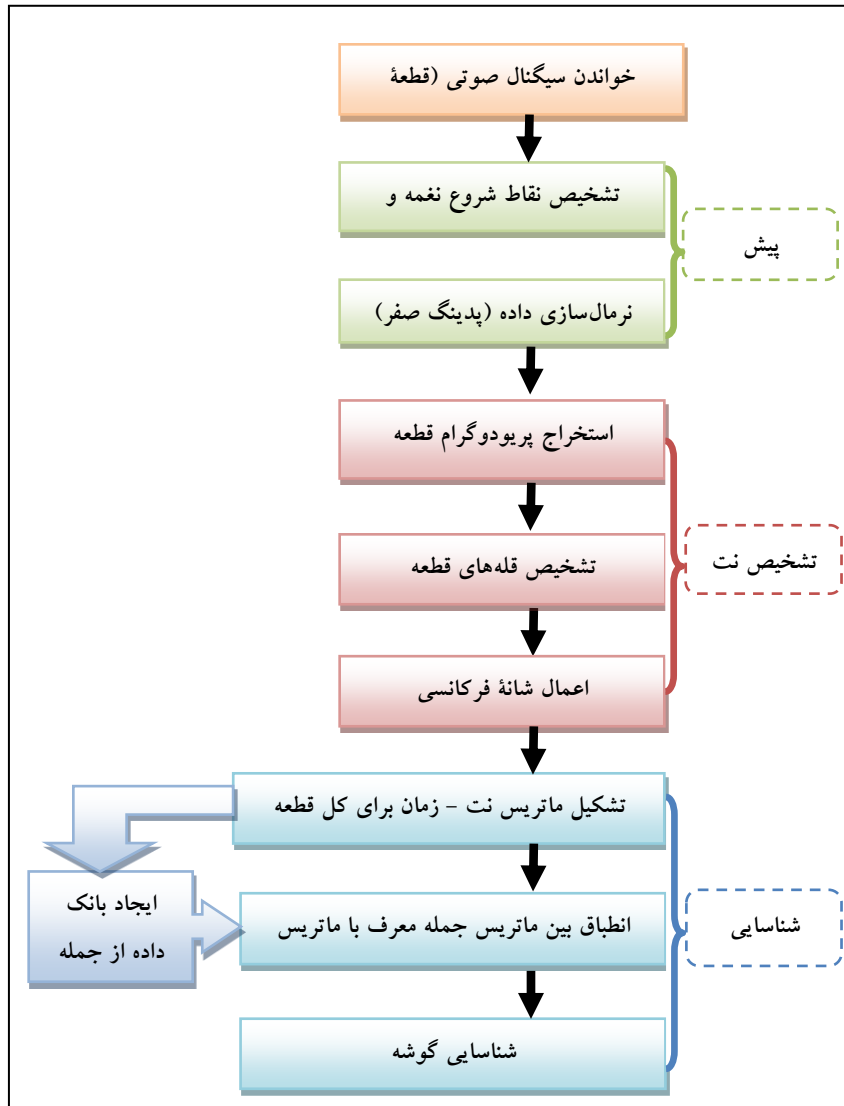
⁵ Homogenize

⁶ Down sampling

⁷ Hamming

روش شناسایی پژوهش

همان‌طور که اشاره شد، تقریباً تمامی پژوهش‌های انجام شده برای شناسایی خودکار موسیقی سنتی ایرانی، بر اساس رویکرد ماکرو بوده و تنها اشل صوتی و فواصل را برای تشخیص و تفکیک پنج دستگاه اصلی از یکدیگر ملاک قرار داده‌اند. در این پژوهش، سعی شده با استفاده از رویکرد میکرو و تشخیص جمله معرف در هر گوشه، امکان شناسایی خودکار گوشه‌های مختلف در ردیف دستگاهی فراهم شود. بر این اساس، در ادامه به ارائه روش و الگوریتم پیشنهادی (شکل ۱) می‌پردازیم.



شکل ۲. الگوریتم پیشنهادی برای شناسایی گوشه‌ها در موسیقی سنتی ایرانی

- در این روش، ابتدا سیگنال موسیقی مورد نظر پنجره‌بندی شده و سپس در هر پنجره زمانی، نغمه نواخته شده شناسایی و تشخیص داده می‌شود. پس از این مرحله، ماتریس نت-زمان برای سیگنال کل قطعه موسیقی را تشکیل می‌دهیم. گام بعدی، تشکیل ماتریس نت-زمان برای جمله معرف موسیقی مورد نظر است. سپس، با شیفت دادن ماتریس نت-

زمان جمله معرف بر روی ماتریس نت-زمان کل قطعه موسیقی و تطبیق این دو ماتریس، گوشه مورد بررسی تشخیص داده خواهد شد؛ بدین صورت که اگر ماتریس نت-زمان برای جمله معرف یک گوشه، با بخشی از ماتریس نت-زمان کل قطعه تطبیق یابد، نشان‌دهنده این است که این قطعه مربوط به همان گوشه است و اگر تطبیق داده نشود، این قطعه مربوط به آن گوشه نیست. به همین ترتیب، برای همه گوشه‌های مورد نظر، یک ماتریس نت-زمان برای جمله معرف تشکیل می‌دهیم و آن‌ها را در یک بانک داده^۱ از جملات معرف، ذخیره می‌کنیم.

روش پیشنهادی برای تشخیص نغمه و شناسایی گوشه

الگوریتم پیشنهادی را می‌توان روی گوشه‌های اصلی دستگاه شور (برای مثال درآمد اول شور و پنج گوشه اوج، شهناز، مقدمه قرچه، قرچه و رضوی) و روی سازهای سنتی ایرانی، مانند تار، سه‌تار، سنتور، بربط، کمانچه و نی اعمال کرده و میزان دقت و خطا در تفکیک هر یک از این شش گوشه از یکدیگر را بررسی کرد؛ انتخاب دستگاه شور از آن جهت است که بر اساس گفته بسیاری از اساتید و توافق جمعی موسیقی‌دانان ایرانی، شور را می‌توان دستگاه مادر در موسیقی سنتی ایرانی نامید (دستگاه‌های موسیقی ایرانی، ۲۰۱۶؛ فرهنگ، ۱۳۹۱: ۵۳؛ نعل، ۱۳۹۳: ۷۸؛ فخرالدینی، ۱۳۹۴: ۱۲۶) و از مجموع تصانیف و آهنگ‌هایی که تاکنون در موسیقی سنتی ایرانی ساخته شده است، به جرأت می‌توان گفت که بیش از یک‌سوم از آن‌ها در دستگاه شور بوده است^۲. چنانچه جداسازی در این گوشه‌ها به درستی جواب دهد، به راحتی می‌توان نتایج آن را به سایر دستگاه‌ها و گوشه‌ها نیز تعمیم داد.

از طرف دیگر، اساس شناسایی گوشه در مرحله اول به نغمه‌های آن گوشه بستگی دارد و اگر شناسایی نغمه‌های اصلی و نواک به درستی و با دقت قابل قبول انجام نشود، به

^۱ Database

^۲ همچنین بنا به اظهار نعل (۱۳۹۳، ص ۷۸)، بیشترین اجراها طی دهه‌های ۱۳۵۰ و ۱۳۶۰ شمسی مربوط به مجموعه دستگاه شور و متعلقات آن بوده است.

طور یقین، نتایج تشخیص گوشه و دستگاه نیز مطلوب نخواهد بود. از این رو، در این پژوهش به طور ویژه‌ای به این مرحله توجه شده است. در ادامه، مراحل و فرایند انجام تشخیص نغمه و سپس روش پیشنهادی برای شناسایی خودکار گوشه گوشه‌ها از یکدیگر، به تفصیل توضیح داده می‌شود.

مرحله اول در تشخیص نغمه، خواندن سیگنال موسیقی موردنظر است؛ به این منظور از تابع `audioread` یا `miraudio` در نرم‌افزار متلب استفاده می‌شود. در مرحله بعد، فریم‌بندی سیگنال موسیقی انجام می‌شود. برای محاسبه طول پنجره در موسیقی، لازم است سریع‌ترین نغمه‌ای که معمولاً در سازهای غیرکوبه‌ای نواخته می‌شود را ملاک قرار داد تا حیثاً در بررسی و شناسایی نغمه‌های اصلی در یک قطعه موسیقی، هیچ موردی از دست نرفته و از نظر دور نماند. از آنجا که کوتاه‌ترین نغمه در موسیقی، از نظر کشش و ارزش زمانی، مربوط به چهارلاچنگ بوده و مدت زمان کشش آن حدود ۶۴ میلی‌ثانیه (یا حدود یک شانزدهم ثانیه) است^۱، این عدد مبنای انتخاب طول پنجره زمانی ($T_{window}=64ms$) قرار می‌گیرد. چنانچه فرکانس نمونه‌برداری ثابتی در کل قطعه موسیقی اعمال و انتخاب شود، تعداد نمونه‌های سیگنال در هر پنجره نیز ثابت و عدد مشخصی خواهد بود. با توجه به این که مقدار فرکانس نمونه‌برداری (F_s) برای بیشتر قطعات ضبط شده ۴۴۱۰۰ هرتز است؛ بنابراین تعداد ۲۸۲۲ نمونه در هر پنجره زمانی قرار دارد ($N_{window}=2822$). تجربه نشان می‌دهد که برای شناسایی نغمه و گوشه، بهتر است بر روی فریم‌ها، تابعی اعمال نشود که همان تابع پنجره مستطیلی خواهد بود.

از طرف دیگر، برای این که هیچ اتفاقی در نغمه از دست نرود و هرگونه تغییر احتمالی (چه به لحاظ حرکت از نغمه‌ای به نغمه دیگر و چه به لحاظ تغییر در سرعت اجرا) مدنظر قرار گیرد، میزان همپوشانی^۲ پنجره‌ها ۹۰ درصد انتخاب شده است که در این حالت،

^۱ مدت زمان یک شانزدهم ثانیه تقریبی بوده و در حالتی است که متروم (سرعت یا تمپو) نغمه سیاه برابر ۶۰ ضربه در هر دقیقه در نظر گرفته شود. آنچه مسلم است این که فاصله زمانی و سرعت توالی نغمه‌ها در عمل و هنگام نوازندگی، هر چقدر هم سریع نواخته شوند، کمتر از این مقدار نخواهد بود.

^۲ Overlap

میزان گام یا پرش^۱ برای هر پنجره برابر با $6/4$ میلی ثانیه (یا $0/064$ ثانیه) خواهد بود. بر این اساس، تعداد نمونه‌ها در هر گام برابر با 282 نمونه خواهد بود. همچنین، تعداد کل پنجره‌ها در یک قطعه موسیقی (N_w)، از تفاضل تعداد کل نمونه‌ها در آن قطعه موسیقی با 282 ، تقسیم بر 282 به دست خواهد آمد.

اما در یک قطعه موسیقی ساخته شده، از کشش‌های زمانی مختلف (مانند سیاه، چنگ، دولاچنگ و غیره) استفاده می‌شود و از این رو، انتخاب طول پنجره با فواصل زمانی یکسان معقول و منطقی به نظر نمی‌رسد؛ بنابراین، روش پنجره‌بندی با طول یکسان بدون توجه به کشش‌ها و فواصل زمانی واقعی در موسیقی بوده و روشی کورکورانه و ناکارآمد است؛ به عبارت دیگر، پژوهش‌های مشابهی که به تبعیت از حوزه‌های دیگر و موضوعات غیرموسیقایی، پنجره‌بندی سیگنال موسیقی را با طول و فواصل یکسان انجام داده‌اند، به ماهیت متغیر بودن کشش نغمه‌ها توجه نداشته‌اند. این روش در بیشتر سیگنال‌های غیرموسیقایی کارآمد بوده است، ولی در سیگنال‌های موسیقی باید به دنبال محل نواختن هر نت و لحظه شروع هر نغمه باشیم. از این رو، در این پژوهش از روش تشخیص شروع نغمه^۲ استفاده شده است.

در این پژوهش نیز به منظور تشخیص نقاط قله^۳ و تشکیل ماتریس نت-زمان، ابتدا از روش اول یعنی پنجره‌بندی سیگنال موسیقی با طول یکسان (64 میلی ثانیه در هر فریم) استفاده شد و نتایج خوبی حاصل نشد؛ بنابراین، از روش تشخیص شروع نغمه و قطعه‌بندی سیگنال^۴، با توجه به زمان و لحظه ضربه خوردن مضراب استفاده شد. روش قطعه‌بندی را می‌توان نوعی فریم‌بندی دانست؛ با این تفاوت که در قطعه‌بندی طول فریم‌ها یکسان نیست و با توجه به زمان و لحظه ضربه خوردن مضراب تعیین می‌شود. به عبارت دیگر، شروع فریم، لحظه ضربه خوردن مضراب و پایان آن لحظه ضربه خوردن مضراب بعدی است. برای یافتن لحظه‌های ضربه خوردن مضراب، از تابع Mironset که برگرفته از جعبه‌ابزار

1 Step
2 Onset detection
3 Peaks
4 Segmentation

MIRToolbox در نرم افزار متلب است، استفاده شد. سپس برای قطعه بندی سیگنال بر اساس زمان ضربه خوردن مضراب، از تابع Mirsegment استفاده شد. از مرحله قطعه بندی سیگنال موسیقی به بعد، در کلیه مراحل تحلیل ها روی هر پنجره (قطعه^۱) به صورت جداگانه اعمال و اجرا شده است و نه کل قطعه موسیقی.

پیش از مرحله تحلیل فرکانسی بر روی سیگنال، لازم است داده ها را نرمال سازی کرد. برای این منظور، با استفاده از روش زیرسازی صفر (یا پدینگ صفر^۲)، با اضافه کردن نه نقطه صفر به انتهای سیگنال، تعداد نمونه ها را ده برابر کرده تا هر نمونه، نشان دهنده یک هرتز تغییرات باشد و نه ده هرتز. این نقاط در شکل تبدیل فوریه تأثیری نداشته و فقط تعداد نمونه ها را افزایش می دهند. این مرحله دقت پرلودوگرام (که نوعی از تبدیل فوریه است) را بیشتر می کند، به طوری که هر گام (پرش) در تبدیل فوریه، به جای این که ۱۰ هرتز را نشان دهد، ۱ هرتز را نشان می دهد. برای مثال، فرض کنید که یک قله در ۷۸ هرتز داشته باشیم و تبدیل فوریه هر ۱۰ هرتز تغییرات را نشان دهد. پس ما تغییرات فرکانسی در ۷۰ هرتز و ۸۰ هرتز را خواهیم داشت؛ ولی چون گام ما ۱۰ هرتز است، از روی قله ۷۸ هرتز عبور کرده و نتوانسته ایم آن را تشخیص دهیم؛ در صورتی که دقت ۱ هرتز باشد، تغییرات در هر ۱ هرتز داده شده و می توان قله ۷۸ هرتز را نیز شناسایی کرد.

گام بعدی، تحلیل فرکانسی فریم های سیگنال موسیقی است. بدین منظور، برای به دست آوردن رزولوشن فرکانسی و شکل تخمین طیف بهتر، بهتر است به جای تبدیل فوریه سریع، از سیگنال پرلودوگرام استفاده شود تا از چگالی طیف توان سیگنال، تخمینی به دست آید. طیف توان سیگنال به صورت زیر تعریف می شود:

$$S(w) = 1/N[x(n) \sum_{n=0}^{\infty} x(n) e^{-i\omega t}] \quad \text{رابطه (۱)}$$

که در آن w نشان دهنده فرکانس مربوطه بوده و $x(n)$ نیز سیگنال مورد نظر است.

همچنین، N تعداد کل نمونه های سیگنال را نشان می دهد.

1. Segment
2. Zero padding

در مرحله سوم، لازم است فرکانس‌هایی که دامنه آن‌ها حداکثر است (نقاط قله)، در هر پنجره شناسایی شوند. فرکانس این نقاط، معمولاً همان نواک نغمه‌ها (یا نام نغمه‌ها) هستند. برای این منظور، الگوریتم تشخیص قله به شرح زیر دنبال شده است:

۱- محاسبه تفاضل مرتبه اول تبدیل فوریه سیگنال گسسته ورودی:

$$d(n) = \text{Diff}(\text{fft}(x(n))) = \text{fft}(x(n)) - \text{fft}(x(n-1)) \quad (\text{رابطه ۲})$$

که در آن $x(n)$ سیگنال گسسته ورودی، diff نشان‌دهنده دیفرانسیل و fft هم نماد تبدیل فوریه است.

۲- محاسبه حاصل ضرب نقاط متوالی تابع به دست آمده در مرحله قبلی:

$$m(n) = d(n) \times d(n-1) \quad (\text{رابطه ۳})$$

۳- محاسبه نقاطی (n هایی) که در آن تابع منفی است که با تابع علامت قابل انجام است:

$$s(n) = \text{sign}(m(n)) \quad (\text{رابطه ۴})$$

۴- نقاطی که تابع علامت منفی شده است، در تبدیل فوریه سیگنال همان فرکانس‌هایی هستند که در آن‌ها قله سیگنال اتفاق افتاده است.

۵- با اعمال آستانه مناسب روی دامنه تبدیل فوریه، فرکانس‌هایی که دامنه آن‌ها از آستانه بیشتر باشد، به عنوان نغمه‌های اصلی یا هارمونیک‌های آن‌ها شناسایی می‌شوند. در مرحله آخر، باید دقت داشت که تنها نقاط حداکثری (ماکزیمم) انتخاب شوند. از این رو، فرکانس‌های به دست آمده از مرحله پنج، با نقاط مجاورشان مقایسه شده و در صورت بزرگ‌تر بودن نسبت به آن‌ها، به عنوان نقطه قله در نظر گرفته می‌شوند (بیات، ۱۳۹۲: ۵۶-۵۷).

پس از شناسایی تمامی نقاط قله، در نهایت سه قله برتر به دست آمده در یک بردار ذخیره می‌شوند. این بردار، با هر ستون ماتریس استاندارد سه سطری، شامل فرکانس هر نغمه و هارمونیک‌های آن، مقایسه می‌شود. اگر سه عنصر از بردار ذخیره شده، در هر یک

از ستون‌های ماتریس استاندارد حضور داشته باشد، نغمه مربوط به آن ستون علامت زده می‌شود.

همان‌طور که گفته شد، تنها تعداد معین و مشخصی از نواک‌ها، در محدوده بسامد بین ۱۶ تا ۷ هزار هرتز در موسیقی استفاده می‌شود، به طوری که تعداد نواک‌هایی که برای کلیه قطعات موسیقی غربی در یک ارکستر بزرگ استفاده می‌شود، حدود ۹۰ نغمه است. طی بررسی انجام شده^۱، تعداد نغمه‌هایی که برای نواختن این شش گوشه اصلی دستگاه شور با سازهای سنتی ایرانی استفاده می‌شود، حدود ۱۶ نغمه است (از نغمه F_3 تا G_5). با توجه به این که هر شش گوشه در دستگاه شور قرار دارند، نسبت فواصل همه این گوشه‌ها یکسان است.

با فرض این که در همه این قطعات، بیش از ۱۶ نغمه مشخص استفاده نشده و با کوک استاندارد دیپازن نواخته شوند، می‌توان تشخیص فرکانس را به این مقادیر محدود کرد که هر یک مبین نغمه‌ای خاص است؛ بنابراین، با تشکیل فیلتر نغمه‌ها، امکان تشخیص نغمه موردنظر فراهم می‌آید؛ به این معنا که این بار با شناخت بردار قله‌ها، یعنی فرکانس اصلی و هارمونیک‌های آن که با توجه به فرکانس استاندارد نغمه‌ها به دست آمده است، وجود این قله‌ها در روی طیف سیگنال بررسی می‌شوند و اگر دامنه آن‌ها از یک مقدار آستانه مشخص بیشتر باشد، دلیل بر وجود نغمه خاص است. به عبارت دیگر، در این روش فیلتری در حوزه فرکانس ساخته شده است که با قرار دادن آن روی طیف سیگنال، می‌توان از بودن یا نبودن یک نغمه مطلع شد. این فیلتر بر اساس فرکانس نواک نغمه و دو هارمونیک اصلی آن ساخته می‌شود. در این حالت نیز بر روی پنجره‌های زمانی مشخص فرکانس هر نغمه استاندارد و هارمونیک‌های دوم و سوم آن، فیلتر فرکانسی را ایجاد کرد.

^۱ در رساله دکتری، با عنوان: «شناسایی خودکار دستگاه شور در موسیقی سنتی ایرانی بر پایه تک‌نوازی سازهای تار، سه‌تار، سنتور و برپه»، اساتید راهنما: دکتر کیوان برنا و دکتر حامد ساجدی. اساتید مشاور: دکتر داریوش علی‌محمدی و دکتر پویا سرایی. گروه علم اطلاعات و دانش‌شناسی، دانشگاه خوارزمی.

از این رو، روند کار به این ترتیب است که برای هر نغمه، یک فیلتر خاص در نظر گرفته و در ستون یک ماتریس، ذخیره می‌شود. سپس، این ماتریس (که هر ستون آن حاوی پنجره فرکانسی یا فیلتر هر نغمه است) در طیف سیگنال ضرب می‌شود. اگر میانگین حاصل ضرب هر ستون از یک حدی بیشتر باشد، نشان‌دهنده وجود نغمه مربوط به همان فیلتر است.

یافته‌های پژوهش

در این پژوهش برای تشخیص نغمه، انواع روش‌ها برای تشکیل فیلتر فرکانسی انجام شد، یعنی فیلتر فرکانسی با یک فرکانس اصلی و دو هامونیک بعدی، فیلتر فرکانسی با یک فرکانس اصلی و ۳ هامونیک بعدی، فیلتر فرکانسی با یک فرکانس اصلی و ۴ هامونیک بعدی، تشخیص ۴ تا از این ۵ فرکانس و غیره که بهترین نتیجه برای تشخیص نغمه، مربوط به فیلتر فرکانسی با یک فرکانس اصلی و دو هامونیک بعدی تشخیص داده شد.

با توجه به این که در شش گوشه مورد بررسی ۱۶ نغمه بکار رفته است، برای هر نغمه یک ماتریس نغمه (در مجموع ۱۶ ماتریس نغمه) تشکیل می‌دهیم. این ماتریس به گونه‌ای عمل می‌کند که خود نغمه و دو هامونیک بعدی آن را در نظر گرفته، بقیه طیف را در نظر نمی‌گیرد و تنها در همین فرکانس‌ها به دنبال قله می‌گردد. برای افزایش کارایی فیلتر فرکانسی، پهنای آن متناسب با مرتبه هامونیک تنظیم می‌شود. به این منظور یک بازه $0/01 \pm$ برای هر فرکانس و هامونیک‌ها در نظر گرفته می‌شود. به عبارت دیگر، اندازه این فیلتر در فرکانس‌های کم‌تر و در فرکانس‌های زیر بزرگ‌تر خواهد بود؛ زیرا دامنه نغمه‌ها و بازه فرکانسی آن‌ها، در فرکانس‌های پایین، کوچک‌تر و در فرکانس‌های بالا، بزرگ‌تر است. برای مثال اگر فرکانس نغمه سل را 200 ± 2 (Hz) در نظر گرفته می‌شود. برای هامونیک دوم آن که دو برابر بازه $2 \pm$ (Hz) در نظر گرفته می‌شود. برای هامونیک سوم نیز بازه 6 ± 600 تشکیل خواهد شد.

مرحله چهارم، تشکیل ماتریس نت-زمان است. این ماتریس نشان‌دهنده ترتیب نواختن نغمه‌ها و زمان نواخته شده برای هر نغمه است؛ همانند آنچه در موقع نواختن یک قطعه موسیقی اتفاق می‌افتد. در واقع، یک قطعه موسیقی شامل فرایندی دوجهی از زمان و فرکانس است که در آن توالی فرکانس‌های معینی، بر اساس یک زمان‌بندی از پیش تعیین شده‌ای، نواخته و اجرا می‌شوند.

برای تشکیل ماتریس نت-زمان لازم است به برخی نکات توجه شود. اول این که، سکوت بعد از هر نغمه، جزء زمان همان نغمه در نظر گرفته شود. دوم این که اگر نغمه‌ای تنها در یک پنجره زمانی تشخیص داده شد و در پنجره‌های قبل و بعد از آن تشخیص داده نشد، به عنوان نویز در نظر گرفته شود؛ چرا که با توجه به مقدار گام پنجره تعریف شده (۶/۴ میلی‌ثانیه)، امکان ندارد که یک نغمه به تنهایی در این زمان کوتاه نواخته شود. نکته سوم این که وقتی در یک پنجره تلاقی دو نغمه مختلف اتفاق می‌افتد، دیگر نباید زمان را برای نغمه قبلی محاسبه کرد و همان‌جا که نغمه بعدی شروع شد، حتی اگر نغمه قبلی ادامه داشته باشد، باید این پنجره به عنوان پایان نغمه قبلی و شروع نغمه جدید در نظر گرفته شود؛ چرا که به دلیل داشتن طنین ساز، هر نغمه ممکن است تا زدن نغمه بعدی میرا نشده باشد. از این رو، لازم است این نکات را در تشکیل ماتریس نت-زمان لحاظ کرده و در نظر گرفت. برای تشخیص نغمه‌های یک قطعه موسیقی به منظور شناسایی دستگاه یا گام در موسیقی، ذکر این نکته ضروری است که درجه‌ها و فواصل بین آن‌ها، در یک بازه تکرارپذیر (مانند اکتاو یا دانگ) اهمیت دارد. به عبارت دیگر، تنها نام نغمه نواخته شده اهمیت دارد و نه شماره اکتاو نغمه. از این رو، لازم است تا نغمه‌های یکسانی که در اکتاوهای مختلف نواخته می‌شوند را به یک نغمه و در یک اکتاو مشخص کاهش داد. برای مثال اگر نغمه فا در سه اکتاو مختلف سوم، چهارم و پنجم نواخته شوند، در شناسایی نغمه به طور جداگانه بررسی نشده و همه آن‌ها را نغمه فا در نظر می‌گیریم؛ بنابراین، قبل از تشکیل ماتریس نت-زمان، همه نغمه‌های همسان که تنها در اکتاو با هم تفاوت دارند را

یکی در نظر می‌گیریم. بدین ترتیب ماتریس نت-زمان خلاصه‌تر شده و میزان محاسبات نیز کاهش می‌یابد.

در مرحله پنجم، ماتریس نت-زمان برای جمله معرف هر گوشه تشکیل می‌شود. این ماتریس را به دو طریق می‌توان تشکیل داد: روش اول، تشکیل این ماتریس بر اساس نغمه‌نگاری یکی از ردیف‌های رایج (برای مثال، ردیف میرزا عبدالله) است. روش دوم، تشکیل چنین ماتریسی با استفاده از قطعه نواخته شده و جداسازی آن از روی خود قطعه است. در صورتی که شش گوشه اصلی دستگاه شور انتخاب شوند، بر همین اساس نیز لازم است شش ماتریس نت-زمان برای جمله معرف هر یک از گوشه‌ها، به طور جداگانه ایجاد و آن‌ها را در یک بانک داده، ذخیره و نگهداری کرد.

در مرحله آخر، با شیفت دادن و لغزاندن ماتریس نت-زمان جمله معرف بر روی ماتریس نت-زمان کل قطعه، میزان تطبیق بین این دو ماتریس به دست آورده شده است؛ در صورتی که ماتریس جمله معرف هر یک از گوشه‌ها، در قسمتی با سیگنال اصلی تطبیق داشته باشد، می‌توان آن قطعه را به گوشه مربوطه نسبت داد. به منظور تطبیق بین دو ماتریس جمله معرف و قطعه موسیقی، از فرمول عکس خطای میانگین مربعات^۱ استفاده شده است.

$$\text{رابطه (۵)} \quad \frac{1}{(X - X_m)^2}$$

که در آن X نشان‌دهنده ماتریس نت-زمان قطعه موسیقی، X_m بیانگر ماتریس نت-زمان جمله معرف است.

نتیجه‌گیری و پیشنهادها

نکته بسیار مهمی که در دیگر پژوهش‌های مرتبط مورد غفلت واقع شده و در نهایت منجر به افزایش میزان خطای تشخیص خودکار و تفکیک دستگاه‌ها از یکدیگر شده است، تلاش برای شناسایی دستگاه‌ها تنها بر اساس میزان اشل صوتی آن‌ها بوده است. به عبارت دیگر، از میان هر یک از سه مؤلفه مد برای تفکیک گوشه‌ها از یکدیگر، یعنی اشل صوتی،

^۱ MSE (Mean Squared Error)

نقش درجات یا الگوهای ملودیک، تنها مؤلفه اول، یعنی اشکل صوتی مورد بررسی قرار گرفته است.

این درحالی است که اشکل صوتی و گام بالفعل بیشتر دستگاه‌های موسیقی سنتی ایرانی، در یکی از دانگ‌های اول یا دوم با یکدیگر اشتراک دارند. به عبارت دیگر، فواصل گام‌ها در دستگاه‌های موسیقی سنتی، بسیار به یکدیگر شبیه و نزدیک هستند و این خود مانعی جدی برای تفکیک دستگاه‌ها از یکدیگر است. از این رو، با اندکی تغییر در فواصل یک گام می‌توان به دستگاه دیگری انتقال یافت. برای مثال، چنانچه درجه سوم دستگاه شور را نیم‌پرده افزایش دهیم، بدون هیچ تغییر دیگری در دیگر درجات این گام، به دستگاه جدید همایون و آواز اصفهان خواهیم رسید؛ یا با کاهش ربع‌پرده در درجه سوم گام شور، دستگاه سه‌گاه ایجاد خواهد شد. نتایج پژوهش عبدلی (۲۰۱۱: ۲۷۹) نیز مؤید این مسئله است؛ به نحوی که بیشترین شباهت، بین دو دستگاه ماهور و چهارگاه (۷۳ درصد شباهت) و کمترین شباهت بین چهارگاه و سه‌گاه (با ۴۳ درصد شباهت) است. شباهت کلی فواصل دستگاه‌ها به یکدیگر نیز ۸۵ درصد عنوان شده است که رقم قابل توجهی است. همچنین، عبدالله‌زادگان (۱۳۹۳: ۴۴) در پژوهش خود، میزان شباهت بین پنج دستگاه اصلی موسیقی ایرانی را بررسی کرده است. نتایج این پژوهش نشان داد که همه دستگاه‌ها بالای ۴۳ درصد به یکدیگر شباهت دارند^۱.

از این رو، با روش‌هایی که تاکنون برای شناسایی خودکار دستگاه‌ها ارائه شده، تمایز نه تنها گوشه‌هایی که در یک مایه یا دستگاه قرار دارند امکان‌پذیر نیست؛ بلکه تفکیک کامل دستگاه‌ها از یکدیگر صرفاً بر اساس اشکل صوتی آن‌ها نیز ممکن نیست. ضمن این که، همان‌طور که گفته شد، تقسیم‌بندی دستگاه‌ها از اصالت لازم برخوردار نبوده و در این خصوص بین نظریه‌پردازان و موسیقی‌دان‌ها، از نظر تعداد دستگاه و مرز بین آن‌ها، اتفاق نظر وجود ندارد.

^۱ مثلاً میزان شباهت دستگاه ماهور به چهارگاه ۷۳ درصد، همایون به چهارگاه ۶۳ درصد، ماهور و سه‌گاه ۶۰ درصد، شور و همایون ۵۹ درصد و... عنوان شده است.

به باور پژوهشگر، جمله معرف هر گوشه می‌تواند معیار درست و مناسبی برای شناسایی خودکار و اندازه‌گیری وجه افتراق گوشه‌ها از یکدیگر باشد. به واسطه شناسایی جمله معرف در هر گوشه، امکان تفکیک گوشه‌هایی که در یک دستگاه یا گام قرار دارند فراهم خواهد آمد؛ مسئله‌ای که تاکنون در هیچ پژوهش مشابه دیگری به آن نپرداخته‌اند. از طرف دیگر، به واسطه دسته‌بندی و تشخیص گوشه‌ها از یکدیگر، دستگاه مربوط به آن را نیز می‌توان شناسایی کرد.

در روش پیشنهادی این پژوهش، در درجه نخست تلاش بر بهبود پارامترهای شناسایی نغمه بوده است. پس از شناسایی نغمه‌های یک قطعه موسیقی، با استفاده از رویکرد میکرو، یک گوشه را می‌توان به جمله‌های (انگاره‌های) تشکیل دهنده آن تجزیه کرد و سپس با تعریف و بررسی سیر ملودی مهم‌ترین جمله آن، یعنی جمله معرف و ذخیره آن در یک بانک داده، به انطباق جمله معرف با کل قطعه پرداخت. قطعه‌ای که بیشترین انطباق را با جمله معرف آن گوشه داشته باشد، به‌عنوان گوشه مورد نظر شناسایی خواهد شد.

از این رو، انجام پژوهش‌هایی با رویکرد میکرو که در آن تأکید بر شناسایی خودکار گوشه‌های اصلی از طریق بررسی سیر ملودی جمله‌های معرف هر گوشه است و همچنین، یافتن روش‌هایی به‌منظور بهبود عملکرد سامانه‌ها در تشخیص نغمه، نقش درجات و تنالیت، به‌عنوان جهت‌گیری آینده پژوهش‌ها پیشنهاد می‌شود.

منابع

- Abbasi Layegh, M., Haghypour, S. & Najafi Sarem, Y. (2013). Classification of the Radif of Mirza Abdollah a canonic repertoire of Persian music using SVM method. *Gazi University Journal of Science, Part A: Engineering and Innovation*, Vol. 1, Issue 4: 57-66.
- Abdoli S. (2011). Iranian traditional music Dastgah classification, *12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 275-280.
- Abdollahzadegan, S. (2014). *Automatic classification of Iranian traditional music based on scale (Dastgah)*. Supervisor: Shahram Jafari. Master's Dissertation, Department of Electrical Engineering, Shiraz University, Shiraz. (In Persian)
- Asadi, H. (2004). Theoretical foundations of Persian classical music: Dastgah as a multi-modal cycle. *Mahoor Music Quarterly*, 6 (22): 43-56. (In Persian)
- Bayat, H. (2013). *Classification and Identification of scale (Dastgah) in Iranian music*. Supervisor: Hamed Sajedi. Master's Dissertation, Department of Electronic and Electrical Engineering, Shahed University, Tehran. (In Persian)
- Behnamnia, B. (2010). *Structural study of the Gushes in Nava scale in Mirza Abdollah's repertoire*. Supervisor: Hamidreza Ardalan. Master's Dissertation, Iranian Music Performance, Faculty of Music, University of Art, Tehran. (In Persian)
- Beigzadeh, B. & Belali Koochesfahani, M. (2016). Classification of Iranian traditional musical modes (Dastgāh) with artificial neural network. *Journal of Theoretical and Applied Vibration and Acoustics*, Vol. 2, Issue 2: 107-118.
- Darabi, N. (2003). *Producing and analyzing music digital signals: designing Nava scale programming language to produce music pieces and automatic recognition of scales (Dastgahs) for Persian music*, Bachelor's Dissertation, Department of Electrical Engineering, Khaje Nasir Toosi University of Technology, Tehran. (In Persian)
- *Dastgah, Iranian Music* (2016). Wikipedia, the free encyclopedia. 28th January. Available via: <https://en.wikipedia.org/wiki/Dastgah>
- Fakhraddini, F. (2013). *Analysis and description of Radifs (repertoires) in Iranian Music*. Tehran: Moin and Music Museum of Iran Publishing Co. (In Persian)
- Fakhraddini, F. (2015). *Iranian Music Harmony*. Tehran: Moin Publishing Co. (In Persian)
- Farhat, H. (1990). *The Dastgah Concept in Persian Music*. Translator: Mehdi Poormohammad. Tehran: Part publication. (In Persian)
- Gavahian Jahromi, E. (2010). *Persian Music Classification using Pitch Profile Feature*. Supervisor: Hossein Marvi. Advisor: Ali Soleimani. Master's Dissertation, Faculty of Electrical and Robotic Engineering, Shahrood University of Technology, Shahrood. (In Persian)
- Hannaneh, M. (1988). *Lost Scales: A research on the foundations and concepts of Iranian music*. Soroush Publishing Co. (In Persian)
- Mahmoodan, S. (2012). *Automatic classification of audio signals using artificial neural network*. Supervisor: Ayoob Banooshi. Master's Dissertation in Audio Engineering, Department of Broadcasting, IRIB University, Tehran. (In Persian)
- Nettle, B. (2014). *The Radif of Persian Music: Studies of Structure and Cultural Context*. Translator: Ali Shadkam. Tehran: Sooremehr Publishing Co. (In Persian)
- Peivandi, J. (2015). *Design an efficient System to Detect Traditional Iranian Music style*. Supervisor: Hadi Soltanzadeh. Master's Dissertation in Artificial Intelligence, Faculty of Electrical and Computer Engineerin, Semnan University, Semnan. (In Persian)
- Safvat, D. and Caron, N. (2012). *Iranian National Music*. Translator: Soussan Salimzadeh. Tehran: Aras Publishing Co. (In Persian)
- Talai, D. (2015). *Radif analysis: based on the Notation of Mirza Abdollah's Radif with Annotated visual description*. Tehran: Ney Publishing Co. (In Persian)
- *Theoretical Fundamentals of Iranian music* (2009). Hossein Alizadeh [et. al]. Tehran: Mahoor Institute of Culture and Arts. (In Persian)

شناسه دیجیتال (DOI): 10.22091/jemsc.2018.1268

استناد به این مقاله:

وفائیان، امیر، بُرنا، کیوان، ساجدی، حامد، علیمحمدی، داریوش، سرایی، پویا. (۱۳۹۷). «روش پیشنهادی برای شناسایی خودکار گوشه‌ها در ردیف موسیقی سنتی ایرانی با رویکرد میکرو». مدیریت مهندسی و رایانش نرم، ۶ (۲)، ۱۳۸-۱۱۳.