بسم الله الرحمن الرحيم

دانشکده آموزش های مجازی

پایان نامه کارشناسی ارشد مهندسی کامپیوتر

شناسایی نت های پایه ای موسیقی و اجرای آن از روی تصاویر دیجیتال

نگارنده : کامران شکری

استاد راهنما

دکتر حمید حسن پور

شهریور ۱۳۹۶

شماره: ۰۴/۲۷۱۱ الف
تاریخ: ۱۳۹۷/۹/۱۰

باسمه‌تعالی

دانشگاه صنعتی شاهرود
مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد آقای کامران شکری با شماره دانشجویی ۹۲۰۸۵۱۴ رشته مهندسی کامپیوتر گرایش هوش مصنوعی تحت عنوان شناسایی نت های پایه موسیقی و اجرای آن از روی تصاویر دیجیتالی که در تاریخ ۱۳۹۶/۰۶/۲۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با امتیاز ۱۸/۰۰ درجه بسیار خوب) ☑ **مردود** ☐

نوع تحقیق: نظری ☐ **عملی** ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استادراهنمای اول	حمید حسن پور	استاد	
۲- استادراهنمای دوم			
۳- استاد مشاور			
۴- نماینده تحصیلات تکمیلی	منصور فاتح	استادیار	
۵- استاد ممتحن اول	علیرضا احمدی فرد	دانشیار	
۶ـ استاد ممتحن دوم	مرتضی زاهدی	استادیار	

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

این تحقیق پیشکشی است به **خانواده** عزیز و فداکارم، که عاشقانه و

صبورانه همیار و همراه زندگی ام هستند

و

تمامی عزیزانی که لبخند زندگی را بر وجودم نشاندند.

تشکر و قدردانی:

بر خود لازم می‌دانم از زحمات استادان عزیزم جناب آقای دکتر حمید حسن پور به عنوان استاد راهنما که در تمامی مراحل این تحقیق مرا صمیمانه یاری کرده و در مسیر به سرانجام رساندن این پروژه نقش قابل توجهی را ایفا کردند تشکر نمایم.

از همه دوستان و اساتیدی که در طی انجام این پروژه مرا از لطف خود بی بهره نگذاشتند کمال تشکر را می نمایم و برای تمام عزیزان آرزوی موفقیت دارم.

تعهد نامه

اینجانب کامران شکری دانشجوی دوره کارشناسی ارشد رشته مهندسی کامپیوتر دانشکده دانشکده آموزش های مجازی دانشگاه صنعتی شاهرود نویسنده پایان نامه **شناسایی نت های پایه ای موسیقی و اجرای آن از روی تصاویر دیجیتال** تحت راهنمائی استاد محترم جناب آقای دکتر حمید حسن پور متعهد می‌شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود است و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بود اند در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی

چکیده

در میان علائم موجود جهت انتقال مفاهیم, علائمی نیز وجود دارد که در گروه زبان موسیقی قرار می گیرند. زبان موسیقی زبانی است با قواعد و دستورات خاص خود که طی سال های متمادی و در هر دوره با گذشت صد ها سال از شکلی به شکل دیگر در حال تغییر و تکامل می باشد که برای شناخت آن نیاز به تمرین چندین ساله می باشد.

هدف پایان نامه شناسایی و اجرای نت های پایه ای موسیقی می باشد. در راستای این هدف تصویر ورودی که شامل قطعه ای از موسیقی که به زبان موسیقی بیان شده به صورت تصویر از ورودی دریافت می شود.

پیش پردازش های لازم روی تصویر از جمله تصحیح زاویه تصویر ورودی و حذف نویز انجام شده و به منظور شناسایی اجزای تشکیل دهنده قطعه بندی صورت می گیرد و قطعات به اجزای کوچکتر غیر قابل تجزیه تقسیم می گردند. در خلال قطعه بندی, در صورت نیاز, اعوجاجات معنایی ناشی از کیفیت تصویر نیز اصلاح می گردند.

پس از شناسایی هر المان به همراه مختصات آن , به کمک شناسایی کننده های شبکه عصبی مصنوعی , اقدام به به شناسایی علائم می نمائیم . باید در نظر داشت جهت خواندن نت موسیقی از چپ به راست بوده لذا کلیه علائم و نشانه ها به ترتیب قرار گیری در خط حامل شناسائی شده و در نهایت به کمک ابزار MIDI اقدام به ایجاد ماتریس MIDI نموده و کلیه خصوصیات و نشانه ها , زمانبندی ها , فاصله ها را به زبان MIDI ترجمه کرده و ماتریس داده آن را تشکیل میدهیم. در نهایت با کمک ابزاری که نرم افزار متلب جهت پخش فایل های MIDI در اختیارمان قرار داده قطعه موسیقی را اجرا می کنیم.

واژه‌های کلیدی:

زبان موسیقی, نت موسیقی, خطوط حامل, تبدیل رادون, شبکه های عصبی مصنوعی, MIDI

فهرست مطالب

فصل اول : مقدمه ۱

۱-مقدمه ۲

۱-۱-تعریف مسئله : ۴

۱-۲- کاربرد های شناسایی زبان موسیقی : ۶

۱-۳-تفاوت های ساختاری زبان موسیقی با زبان محاوره: ۷

۱-۴-سازمان دهی پایان نامه: ۷

فصل دوم : پیشینه تحقیق

۲-۱- مقدمه ۱۰

۲-۲- سیستم های شناسایی نوری موسیقی ۱۰

۲-۲- ۱- معرفی معماری یک سیستم OMR ۱۱

۲-۲-۲- فاصله مرجع در خطوط حامل ۱۴

۲-۲- ۳- قطعه بندی و شناسایی المان های موسیقی ۱۶

۲-۲- ۴- تولید نهایی بازنمایی از مرحله قطعه بندی و شناسایی ۱۷

۲-۳- شناسایی الگو ۱۹

۲-۳-۱- یادگیری با نظارت ۲۰

۲-۳-۲- یادگیری بدون نظارت ۲۰

۲-۴- انواع نمونه ها ۲۱

۲-۵- انتخاب ویژگی ۲۱

۲-۵-۱- خصوصیات یک مشخصه مناسب ۲۲

۲-۶- عملکرد کلی سیستم های شناسایی الگو ۲۳

۲-۶-۱- داده های آموزش ۲۵

۲-۶-۲- داده های تست ۲۵
۲-۶-۳- داده های اعتبار سنجی ۲۵
۲-۶-۴- انواع شناسایی کننده ها ۲۶
۲-۶-۴-۱- شناسایی کننده های بر مبنای مدل ۲۶
۲-۶-۴-۲- شناسایی کننده های بدون مدل ۲۷
۲-۶-۵- ترکیب شناسایی کننده ها ۲۸
۲-۷- معرفی سیستم های بازشناسی نوری علائم (OCR) ۲۹
۲-۷-۱- انواع سیستم های OCR ۳۰
۲-۷-۲- کاربردهای OCR ۳۱
۲-۷-۳- معرفی بخشهای مختلف در پیاده سازی سیستم های OCR ۳۲
۲-۷-۳-۱- پیش پردازش ۳۲
۲-۷-۳-۱-۱-کاهش نویز ۳۳
۲-۷-۳-۱-۲- اصلاح کجی تصویر ۳۴
۲-۷-۳-۲- باینری کردن تصویر ۳۴
۲-۷-۳-۳- قطعه بندی تصویر ۳۵
۲-۷-۳-۳- ۱- استخراج ویژگی ۳۶
۲-۷-۳-۳-۲- نحوه استخراج ویژگی ۳۷
۲-۷-۴- طبقه بندی و بازشناسی ۳۷
۲-۷-۴-۱- روشهای تشخیص الگو ۳۹
۲-۷-۵- پس پردازش ۴۱
۲-۸- جمع بندی و نتیجه گیری ۴۱

فصل سوم : روش های پیشنهادی

۳-۱- مقدمه ۴۴

۳-۲- تصحیح زاویه تصویر ورودی (نرمال کردن تصویر) ۴۴

۳-۲- ۱- تبدیل رادون ۴۶

۳-۲-۱-۱- تاریخچه تئوری تبدیل رادون ۴۷

۳-۲-۱-۲- اصول تبدیل رادون و روابط حاکم بر آن ۴۸

۳-۲-۱-۳- استفاده از تبدیل رادون جهت تصحیح کجی تصویر ۴۹

۳-۳- به کار گیری تبدیل رادون در شناسایی خط ۵۱

۳-۴- قطعه بندی تصویر ۵۲

۳-۴-۱- آستانه گذاری به روش Otsu: ۵۳

۳-۴-۲- قطعه بندی المانهای تصویر ۵۴

۳-۴-۳- قطعه بندی کل تصویر ۵۶

۳-۴-۴- قطعه بندی المانها از قطعات کلی با استفاده از معادلات تفاضلی ۵۷

۳-۴-۴-۱- بکارگیری مشتق مرتبه اول در جداسازی المانهای موسیقی از قطعات کلی ۵۷

۳-۴-۵- شناسایی و رفع مشکل چسبندگی برخی المانها: ۵۹

۳-۴-۵-۱- شناسایی چسبندگی ۵۹

۳-۴-۵-۲- روش پیشنهادی جهت رفع مشکل ۶۰

۳-۵- شبکه های عصبی مصنوعی ۶۳

۳-۵-۱- شبکه عصبی هاپفیلد ۶۴

۳-۵-۱-۱- مفهوم انرژی در شبکه عصب هاپفیلد ۶۶

۳-۶- جعبه ابزار MIDI ۶۷

۳-۶-۱- خواندن , ایجاد و تغییردادن فایلهای MIDI ۶۸

۳-۶-۱-۱- خواندن , ایجاد و تغییردادن فایلهای MIDI ۶۸

۳-۶-۱-۲- ایجاد یک رشته از نت ها ۷۰

۳-۶-۱-۳- تغییر دادن ماتریس نت ها : ۷۱

۳-۷- نتیجه گیری ۷۲

فصل چهارم: تشریح روش های بکار رفته در پیاده سازی

۴ـ ۱ـ مقدمه ۷۴

۴ـ ۲ـ اصلاح زاویه تصویر: ۷۴

۴ـ ۲ـ ۱ـ محاسبه زاویه تصویر ورودی ۷۵

۴ـ ۲ـ ۲ـ اصلاح زاویه تصویر ۷۷

۴ـ ۳ـ پیدا کردن خطوط حامل ۷۸

۴ـ ۳ـ ۱ـ محاسبه خطوط اصلی و خطوط در سایه: ۸۳

۴ـ ۳ـ ۱ـ ۱ـ محاسبه مکان خطوط اصلی و خطوط حامل ۸۳

۴ـ ۳ـ ۱ـ ۲ـ محاسبه مکان خطوط در سایه ۸۴

۴ـ ۳ـ ۲ـ محاسبه مکان خطوط غیراصلی ۸۵

۴ـ ۳ـ ۳ـ حذف خطوط میزان ۸۶

۴ـ ۳ـ ۴ـ حذف خطوط حامل ۸۸

۴ـ ۴ـ استخراج مرزهای هر دسته از خطوط حامل ۹۱

۴ـ ۵ـ قطعه بندی کلی تصویر: ۹۲

۴ـ ۶ـ قطعه بندی جزئی جهت استخراج المانهای موسیقی ۹۴

۴ـ ۶ـ ۱ـ جداسازی هر یک المانهای موجود در قطعه کلی ۹۴

۴ـ ۶ـ ۲ـ بررسی قابلیت تفکیک عمودی هر قطعه جزئی ۹۷

۴ـ ۶ـ ۳ـ بررسی تفکیک پذیری افقی هر زیر قطعه ۱۰۰

۴ـ ۶ـ ۳ـ ۱ـ یافتن قطعه معیوب ۱۰۱

۴ـ ۶ـ ۳ـ ۲ـ اصلاح قطعه دارای چسبندگی: ۱۰۲

۴ـ ۶ـ ۴ـ ذخیره سازی المانهای تفکیک شده ۱۰۴

۴ـ ۵ـ شناسایی کننده المانها ۱۰۴

۴ـ ۶ـ اجرای موسیقی ۱۰۷

۴ـ ۷ـ خلاصه و نتیجه گیری ۱۰۷

فصل پنجم: خلاصه و نتیجه گیری

۵ـ ۱ـ خلاصه پایان نامه ۱۱۰

۵ـ ۲ـ پیشنهاد ادامه کار ۱۱۳

منابع

ضمیمه : معرفی تئوری بنیادین موسیقی

۱- تعریف موسیقی ۱۱۸

۱-۱-صدا ۱۱۸

۱-۱-۱- بسامد ۱۱۹

۱-۲- صداهای خوشایند ۱۱۹

۱-۳- صداهای ناخوشایند ۱۱۹

۲- نت موسیقی ۱۲۰

۳- الفبای موسیقی ۱۲۰

۳-۱- خطوط حامل ۱۲۱

۳-۲- کلید ۱۲۲

۳-۳- خطوط اضافه ۱۲۴

۴- زمان در موسیقی ۱۲۵

۴-۱- نقطه ونقش آن ۱۲۸

۴-۲- خط اتحاد و خط اتصال ۱۲۹

۴-۲-۲- خط اتصال ۱۳۰

۴-۳- سکوت ۱۳۰

۴-۴- تاکید در موسیقی ۱۳۱
۴-۵- میزان ۱۳۲
۴-۵-۱- وزن در موسیقی ۱۳۲
۴-۵-۲- گونه‌های میزان ۱۳۳
۴-۵-۳- کسر میزان ۱۳۳
۴-۵-۴- میزان‌های ترکیبی ۱۳۴
۵- فاصله در موسیقی ۱۳۵
۵-۱- پرده و نیم پرده ۱۳۵
۵-۲- فاصله در موسیقی ۱۳۵
۵-۲-۱- فاصله ساده- فاصله ترکیبی ۱۳۶
۵-۲-۲- نشانه‌های تغییر دهنده ۱۳۷
۵-۲-۲-۱- دیز (Diese ,#) ۱۳۷
۵-۲-۲-۲- بمل (b, BEMOL) ۱۳۷
۵-۲-۲-۳- دوبل دیز (# #, X, Double Dies) ۱۳۸
۵-۲-۲-۴- دوبل بمل (bb, Double, be mol) ۱۳۸
۵-۲-۲-۵- بکار (Becerra) ۱۳۹
۵-۲-۲-۶- سُری (Sori) و کرن (P, korn) ۱۳۹
۵-۲-۲-۷- تکرار میزان ۱۴۰
۵-۲-۲-۸- تکرار دو میزانی ۱۴۰
۵-۲-۲-۹- گزش ۱۴۱
۵-۲-۲-۱۰- علامت تریل (tr) ۱۴۱
۵-۲-۲-۱۱- برگشت ساده ۱۴۲
۵-۲-۲-۱۲- برگشت متنوع ۱۴۲
۶-خلاصه ۱۴۳

فهرست اشکال

شکل۲-۱: معماری کلی یک سیستم OMR ۱۲

شکل ۲-۲: فاصله بین خطوط حامل و عرض هر خط ۱۵

شکل ۲-۳ : شناسایی کننده آماری الگو ۲۴

شکل ۲-۴- کاهش توزیع خطا با ترکیب دسته بندها ۲۸

شکل ۲-۵: نمونه تصویر از نت موسیقی که به علت کجی تصویر, خطوط حامل در راستای افق نمی باشند. ۳۴

شکل ۳-۱: تبدیل رادون جهت شناسایی خطوط ۵۲

شکل ۳-۲ نمونه خط اتحاد چسبیده به نت ۵۵

شکل ۳-۴ نمونه قطعه موسیقی و مشتق مرتبه اول آن ۵۸

شکل ۳-۵ نمونه المان قطعه بندی شده و خط اتحاد چسبیده به آن ۶۱

نمودار ۳-۱ دو نمونه پروفایل از دو نت هم شکل جهت یافتن نقطه انفصال ۶۲

شکل ۳-۶ نقطه انفصال مطابق مقدار ماگزیمم یافت شده در نمودار ۳-۲ ..... ۶۳

شکل ۳-۷ شمای کلی یک شبکه عصبی هاپفیلد ۶۵

شکل ۳-۸ نمونه تابع Hard lim ۶۵

شکل ۳-۹ دو نقطه تعادل نواحی جذب آنها ۶۶

شکل ۳-۹ نمونه محتوای یک فایل ۶۸

شکل ۳-۱۰ نام هر یک از ستون های فایل ۶۹

جدول ۳-۱ نام گذاری نت های موسیقی در زبان های مختلف و نیز اطلاعات فرکانسی و شماره نت در فایل نت ۷۱

شکل ۴-۱ـ نمونه خط غیرمستقیم که در زمان باینری شدن پله پله شده است. ۷۵

شکل ۴ـ ۳ـ نمودار ماتریس تک ستونی R. ۷۹

شکل ۴ـ ۵ـ نمودار میله ای به همراه نقاط علامت گذاری شده ماتریس R پس از استاندارد سازی ۸۰

شکل ۴ـ ۶ بزرگنمایی سگمنت اول و دوم ماتریس ۸۰

شکل ۴ـ ۷ـ نمودار بزرگنمایی شده $R2$ که مرز بین هر سگمنت را مشخص نموده است. ۸۲

شکل ۴ـ ۸ـ نقاط یافت شده حول هر نقطه پیک با شرایط اشاره شده ۸۵

۴ـ ۹ـ نمونه نت موسیقی که در محلی غیر از خطوط حامل واقع شده است. ۸۵

شکل ۴ـ ۱۰ـ نمونه نتی که ممکن است به اشتباه خط میزان فرض شود ۸۷

شکل ۴ـ ۱۱ـ نمونه خط اتحاد که خط میزان را قطع نموده است. ۸۷

شکل ۴ـ ۱۲ـ نمونه خط حامل حذف شده از تصویر ۸۸

شکل ۴ـ ۱۳ـ نمونه خط حذف شده و عملیات بازسازی المانهای منقطع شده ۸۹

شکل ۴ـ ۱۴ـ حذف قسمتهایی از خط اتحاد به دلیل حذف خط حامل ۸۹

نمودار ۴ـ ۲ـ نمودار مربوط به متغیر D_3 و متغیر D_4 پس از عملیات مشتق گیری ۹۰

شکل ۴ـ ۱۵ـ نمونه خط اتحاد منقطع شده و تصویر پس از ترمیم قسمت منقطع شده ۹۱

شکل ۴ـ ۱۶ـ وجود علائم و نشانه هایی که از منظر عمودی دارای نقطه مشترک با دیگر علائم می باشند. ۹۲

شکل ۴ـ ۱۷ـ نمونه ۲ قطعه دارای یک و بیش از یک المان ۹۳

شکل ۴ـ ۱۸ـ نمونه یک قطعه کلی و اجزای تشکیل دهنده هر یک ۹۵

شکل ۴ـ ۱۹ـ الگوریتم کلی قطعه بندی جزئی المانهای موجود در قطعه های کلی ۹۶

شکل ۴ـ ۲۰ـ قطعه استخراج شده و نمودار محاسبه مشتق مرتبه اول ۹۸

شکل ۴ـ ۲۱ـ نمودار حاصل از مشتق مرتبه اول روی علامت دیز.....۹۸

شکل ۴ـ ۲۲ـ نمونه قطعه کلی و المانهای تفکیک شده.....۹۹

شکل ۴ـ ۲۳ـ نمونه سگمنت که به خوبی المانهای آن تجزیه نگردیده است.....۱۰۰

شکل ۴ـ ۲۴ـ نمونه دو سیگنال حامل مشتق گیری روی مجموع قطعات..... ۱۰۲

شکل ۴ـ ۲۵ سیگنال مربوط به شکل ۴-۲۴ پس از شیفت و انطباق.....۱۰۳

شکل۴ـ۲۶ـ نمونه یک المان ذخیره شده در ماتریس 30×30 (المان مربوط به نت سفید) ۱۰۵

شکل۴ـ۲۷ـ نمونه خروجی شبکه هاپفیلد(المان مربوط به نت سفید) ۱۰۶

فصل اول : مقدمه

۱-مقدمه

یادگیری ماشین یکی از موضوعاتی است که امروزه توجه بسیاری به آن گردیده و با توجه به گستره و کاربرد های آن , پژوهش های فراوانی توسط محققین و علاقه مندان به این حوزه صورت گرفته است. امروزه یادگیری ماشین در زمینه های فراوانی جایگاه خود را یافته و با گذشت اندک سالیان از پیدایش آن کاربردهای,بسیاری در اغلب صنایع و حتی حوزه های غیر صنعتی پیدا نموده است.

یادگیری ماشین در واقع شاخه ای از هوش مصنوعی است که با وجود نو پا بودن , پیشرفت های چشم گیری هم از لحاظ تئوری و نیز از لحاظ عملی شاهد بوده است. گستره مقوله یادگیری ماشین دارای طیف های وسیعی نیز می باشد. به این معنی که یادگیری ماشین در هر بخشی که نیاز به یک انسان خبره داشته باشد قابلیت ورود داشته و می توان با استفاده از داده هایی که از یک انسان خبره دریافت می نماید , سیستم را آموزش داده و رفتارهای مشابه فرد خبره از خود نشان دهد. در ادامه به یکی از کاربرد هایی که در حوزه یادگیری ماشین کمتر به آن توجه شده می پردازیم.

بازشناسایی حروف و اعداد, در واقع توانمند سازی ماشین جهت شناسایی علائم نوشتاری است که امروزه به یک کاربرد مهم در بسیاری زمینه ها تبدیل شده است. علائم نوشتاری (حروف و اعداد) که بشر جهت انتشار افکار خود از آن بهره می برد, علائمی است که درصد بالایی از جمعیت جهان با توجه به موقعیت جغرافیایی که در آن قرار دارند, با یکی از رسم الخط های مرسوم مانند فارسی, عربی, انگلیسی, آسیای شرقی, روسی و دیگر زبان ها, آشنایی داشته و آن را درک نموده و استفاده می نمایند.

در میان علائم موجود جهت انتقال مفاهیم, علائمی نیز وجود دارد که در گروه زبان موسیقی[1] قرار می گیرند. زبان موسیقی زبانی است با قواعد و دستورات خاص خود که طی سال های متمادی و در هر دوره با گذشت صد ها سال از شکلی به شکل دیگر در حال تغییر و تکامل می باشد و این هنر را دارای پیچیدگی های خاص خود نموده که برای شناخت آن نیاز به تمرین چندین ساله می باشد.

شخصی را تصور کنید بدون دارا بودن سواد خواندن و نوشتن برای اولین بار تصمیم به یادگیری زبانی گرفته است. با وجود اینکه ممکن است به واسطه استعداد بالای شخص, در کمترین زمان علائم زبان را در حافظه بسپرد , اما زمانی که تصمیم به خواندن یک صفحه از کتاب نماید , با لکنت و کندی این عمل را انجام خواهد داد. زیرا خواندن فرآیندی است که برای تسلط بر آن به تمرین و تکرار های فراوان جهت یادگیری نیاز دارد.

با مثالی که آورده شد ؛ کاملا واضح است در صورتی که شخص تصمیم به یاد گیری نوشتن نیز نماید , علاوه بر هماهنگی چشم و مغز, نیاز به هماهنگی بسیار بالا در کمترین زمان بین اعضاء و مغز از طریق شبکه های عصبی که در بدن انسان قرار دارد نیز دارد.

حال با مقایسه فرآیند یادگیری برای خواندن و نوشتن با یادگیری موسیقی به این موضوع پی می بریم که در فرآیند یادگیری موسیقی بخشی از کار مشابه یادگیری خواندن و نوشتن یک زبان است. با توضیحاتی که در ادامه در خصوص قواعد و اصول موسیقی به این موضوع پی خواهیم برد که به دلیل پیچیدگی های زبان موسیقی یادگیری آن فرآیندی بس مشکل تر از زبان محاوره ای خواهد بود.

مهارت در خواندن قطعات موسیقی و نواختن آن مستلزم دارا بودن پیش نیازهایی می باشد :

۱- آشنایی با الفبای موسیقی. نت ها به عنوان الفبای موسیقی شناخته می شوند.

---

[1]- در بخش ضمائم به تفضیل در خصوص زبان موسیقی و آشنایی با علائم و قواعد آن صحبت گردیده است.

۲- توانایی خواندن زبان موسیقی. نت ها در کنار هم تشکیل قطعات موسیقی را داده که مشابه کلمات در زبان محاوره می باشند.

۳- آشنایی با فواصل در موسیقی. که یکی از مهمترین ارکان موسیقی است.

۴- آشنایی با یکی از آلات موسیقی.

موارد اشاره شده در بالا که به عنوان پیشنیازهای یادگیری معرفی شدند در واقع همه آن چیزی است که برای درک موسیقی و اجرای آن به آن نیاز داریم. دلیل اینکه این موارد پیش نیاز معرفی نمودیم این است که هر چقدر در هر کدام از موارد اشاره شده به لحاظ تئوری دارای دانش باشیم, اما به هنگام اجرای موسیقی, در صورتی که هماهنگی مورد نیاز بین چشم, مغز و اعضای ما نباشد , در نهایت نتیجه کار ارزشی نخواهد داشت.

مهمترین بخش , مهارتی است که در همزمانی خواندن و نواختن یک قطعه موسیقی به واسطه تمرین های زیاد بدست می آید. که این مرحله مشکل ترین و سخت ترین مرحله در آموزش موسیقی است.

۱-۱-تعریف مسئله :

با توجه به مطالبی که در بخش قبل به آن اشاره شد , در این پایان نامه, مسئله شناسایی نت های پایه ای موسیقی و اجرای آن توسط ماشین مورد تحقیق و بررسی قرار گرفته است. هدف نهایی این است که کامپیوتر تصویری از یک قطعه موسیقی دریافت شده را از ورودی دریافت و اقدام به شناسایی نت ها و اجرای دقیق آن بر اساس قوانین و قواعد حاکم بر زبان موسیقی نماید. به طوری که از لحاظ صحت اجرای نت ها و نیز فواصل زمانی, هیچ اشکالی در اجرای آن در مقایسه با یک فرد خبره در موسیقی وجود نداشته باشد.

در خصوص شناسایی علائم موسیقی مسائلی چالش برانگیز زیادی وجود دارد . یکی از چالش بر انگیزترین مسائل, شناسایی اولیه خطوط حامل در قطعات موسیقی می باشد که این مسئله به تنهایی مورد بحث نیست. همانگونه در ضمائم و در معرفی قواعد و زبان موسیقی به آن اشاره شده است, ملاک عمل و شناسایی انواع نت, موقعیت آن روی خطوط حامل می باشد . از آنجایی که زاویه تصویر ورودی, به هر دلیلی ممکن است با کمی انحراف همراه باشد, این خود باعث شکستگی خطوط حامل در طول تصویر گردیده که تاثیر بسیار نامطلوبی در شناسایی نت ها دارد. برای این منظور و جهت جلو گیری خطا به واسطه زاویه نادرست تصویر از روش تبدیل رادون استفاده نموده ایم که بر اساس آن اقدام به تصحیح زاویه تصویر با خطای ۰.۰۵ درجه نموده ایم.

از دیگر مواردی که در فرآیند شناسایی نت ها بایستی در نظر داشت, عدم وجود قاعده خاصی در فاصله بین المان های موسیقی است. این فواصل گاه بسیار زیاد بوده و گاه بسیار کم از یکدیگر قرار دارند. برای غلبه بر این چالش از روش معادلات تفاضلی مراتب اول و دوم و تواناهایی که این معادلات در استخراج ویژگی دارند بهره بردیم.

یکی دیگر از چالش های پیش رو, قابل تغییر بودن نام نت ها بر اساس آن چیزی است که سیستم بایستی از علامت کلید در قطعه موسیقی درک نماید. در نتیجه در کلید های مختلف نقش خطوط حامل و در نتیجه نام های نت ها با یکدیگر متفاوت بوده که بایستی به آن توجه ویژه نمود.

از دیگر چالش های پیش رو, چسبندگی برخی المان های موسیقی است و باعث اختلال در قطعه بندی تصویر شده که این خود چالشی است که در صورت عدم کنترل و اصلاح آن به منتج به نتایج نا مطلوب می گردد. برای حل این چالش از روش های مقایسه ای بر اساس معیار های تشابه استفاده نموده ایم که نتایج آن به تفضیل در بخش مربوطه توضیح داده خواهد شد.

همینطور بزرگترین چالش پس از شناسایی نت ها و المان های موسیقی , تفسیر آن به زبان کامپیوتر جهت اجرا بوده که این مقوله خود بسیار جای تامل خواهد داشت . زیرا یکی از مهمترین شاخص های تفاوت قطعات موسیقی با یکدیگر , فواصل زمانی است که هر نت به خود اختصاص داده یا فاصله بین دو نت می باشد.

۱-۲- کاربرد های شناسایی زبان موسیقی :

همانگونه که در ابتدای فصل بیان شد , یادگیری موسیقی تنها دانستن قواعد آن نبوده و برای تبدیل شدن به یک فرد خبره در موسیقی که توانایی خواندن نت ها و اجرای آن به طور هم زمان را داشته باشد, نیاز به تمرین و ممارست فراوان می باشد. جدا از نیاز به هماهنگی چشم, مغز و اعضا جهت خواندن و اجرای همزمان قطعات موسیقی , رعایت نکات کلیدی مانند فواصل و زمانبندی های در زمان اجراست که تفاوت بین یک موسیقی دان حرفه ای را از یک موسیقی دان آماتور مشخص می نماید. از آنجایی که میزان آگاهی عموم از زبان موسیقی بسیار بسیار کمتر از میزان آگاهی کسانی است که به زبان محاوره ای آشنایی دارند , لذا در صورت نیاز به راهنمایی در این خصوص بایستی هزینه و زمان زیادی را صرف نمود . این پایان نامه قصد دارد تا با پیاده سازی سیستمی که قابلیت شناسایی و اجرای موسیقی را دارد به دانش آموزان علاقه مند به موسیقی در هر لحظه امکان دسترسی را داده تا برای تصحیح خطای خود از آن بهره مند گردند.

۱-۳-تفاوت های ساختاری زبان موسیقی با زبان محاوره:

با وجود شباهت های بسیار بین زبان موسیقی و زبان محاوره , از دید ساختاری تفاوت های عمیقی بین زبان موسیقی و زبان محاوره وجود دارد. هر زبان محاوره ای از الفبایی تشکیل شده که این الفبا در کنار هم تشکیل کلمه را داده و کلمه ها در کنار هم تشکیل جمله را داده که برای ما معنی دار است.

در موسیقی نیز همین قاعده با اسامی متفاوت وجود دارد اما از لحاظ رسم الخط تفاوت زبان موسیقی با زبان محاوره بسیار زیاد است . اغلب زبان های محاوره ای از تعداد زیادی الفبا تشکیل شده که از ترکیب آنها روی یک خط کلمه به وجود می آید. در صورتی که تعداد الفبای موسیقی در حالت استاندارد ۷ بوده و این الفبا در شرایط استاندارد بروی ۵ خط و در شرایطی خاص بروی بیش از ۵ خط که به خطوط حامل معروفند قرار می گیرند.

به لحاظ ظاهری اکثر علائم موسیقی به هم شبه بوده و این محل استقرار آنهاست که نام آنها را مشخص می نماید.

در صورتی که در الفبای محاوره هیچ دو الف بایی کاملا شبه هم نبوده و در واقع به وضوح با هم تفاوت های شکلی و ساختاری دارند.

۱-۴- سازمان دهی پایان نامه:

این پایان نامه در ادامه در ۵ فصل تهیه شده است. در فصل دوم کارهای پیشین و مرتبط با مبحث آورده شده است. از آنجا که شناسایی نت ها و قطعات موسیقی در بسیاری از زمینه ها مشابهت هایی با سیستم

نویسه خوان نوری[1] داشته , در این بخش به توضیح آن پرداخته شده است. درفصل سوم الگوریتم های بکار رفته و نیز الگوریتم های پیشنهادی شرح داده شده است که شامل ۵ بخش است.

در بخش اول تصحیح زاویه تصویر , بخش دوم شناسایی خطوط حامل و حذف آن از تصویر به همراه مکانیزم نگه داری آدرس آنها به منظور شناسایی نت ها و نیز پیدا کردن خطوط حامل فرضی که از خطوط حامل استاندارد تجاوز نموده اند, بخش سوم شناسایی قطعات کلی و قطعه بندی المان های غیر قابل تجزیه را مورد بررسی قرار می دهیم. در بخش چهارم به شبکه های عصب مصنوعی[2] میپردازیم و مکانیزم و کارایی آنها را در دسته بندی مورد بررسی قرار خواهیم داد و در بخش پایانی این فصل در خصوص ابزار MIDI[3] و تشکیل ماتریس داده مربوط به اصوات و اجرای آنها به تفضیل توضیحاتی ارائه خواهد شد.

در فصل چهارم به توضیح مراحل پیاده سازی پرژژه می پردازیم و در فصل آخر نتیجه گیری از تحقیق پیش رو ارائه خواهد شد و در نهایت نیز پیشنهاداتی برای ادامه این تحقیق برای علاقه مندان به موضوع پایان نامه ارائه خواهد شد.

---

[1]- OCR: Optical Character Recognition.
[2] - Artificial Neural Networks: ANN.
[3] - Music Instrument Device Interface.

فصل دوم : پیشینه تحقیق

۲-۱- مقدمه

تحقیقات در زمینه شناسایی نت های موسیقی اگر چه عمر طولانی ندارد اما در همین زمان کم, تحقیقاتی توسط علاقه مندان به این موضوع صورت گرفته است. زبان موسیقی هم مانند زبان های محاوره از قواعد رسم الخط خاص خود برخوردار است. این رسم الخط اگر چه از دید انسان کاملا متفاوت از رسم الخط محاوره ای است اما از دید ماشین و الگوریتم هایی که برای آن در زمینه شناسایی متون زبان محاوره ای ایجاد شده دارای نقاط مشترک بسیاری می باشد. در این فصل ابتدا به بیان تاریخچه ای از تحقیقات صورت گرفته در خصوص شناسایی زبان موسیقی می پردازیم سپس با توجه به اینکه در بسیاری مراحل, روش های بکار گرفته شده با روشهایی که در شناسایی متون زبان محاوره ای از آنها استفاده شده شباهتهای بسیاری دارد و نیز دارای منابع بیشتری نیز می باشند, به بیان پیشینه سیتم های شناسایی نوری علائم خواهیم پرداخت.

۲-۲- سیستم های شناسایی نوری موسیقی[۱]

تحقیقات در زمینه OMR توسط پروسلین[۲] [۱۶] و پراراو[۳][۱۶] در سال ۱۹۶۶ آغاز و از آن دوره به بعد به طور جدی مورد توجه پژوهشگران و علاقه مندان به این موضوع قرار گرفت و بررسی های بسیاری در زمینه OMR در محافل علمی صورت گرفت.

بلاستین[۴] [۱۶] و بایرد[۵] [۱۶] اولین مقالات منتشره شده شان را در رابطه با شناسایی نوری زبان موسیقی بین سالهای ۱۹۹۲ـ ۱۹۶۶ منتشر نمودند. برین بریدج[۱] [۱۶] و بل[۲] [۱۶] در سال ۲۰۰۱ اولین

[1] - Optical Music Recognition
[2] - Pruslin
[3] - Prerau
[4] - Blostine
[5] - Baird

چهارچوب کلی را در این باره ایجاد کردند. در سال ۲۰۰۵ دبلیو هومندا[۳][۲۰] و همکارانش شناسایی نوری زبان موسیقی را به عنوان موضوع تحقیق در زمینه شناسایی الگو عنوان کردند.

نرم افزارهای تجاری زیادی تا کنون در زمینه OMR بوجود آمده اند. اما هیچکدام نتایج رضایت بخشی نداشته اند. به طور مشخص در خصوص شناسایی نت های دستنویس کارآئی لازم در هیچ کدام از این نرم افزارها دیده نمی شود و در بسیاری مواقع نرم افزارهای موجود همه علائم موسیقی را در برنمی گیرند. لازم به ذکر است اگر چه نرم افزارهای تجاری موجود تمامی علائم موسیقی را پوشش نمی دهند اما اغلب آنها در خصوص اسناد چاپی به خوبی عمل می نمایند اما تحقیقات صورت گرفته در خصوص اسناد دستنویس موسیقی بسیار ناقص بوده و جای کار زیادی دارند.

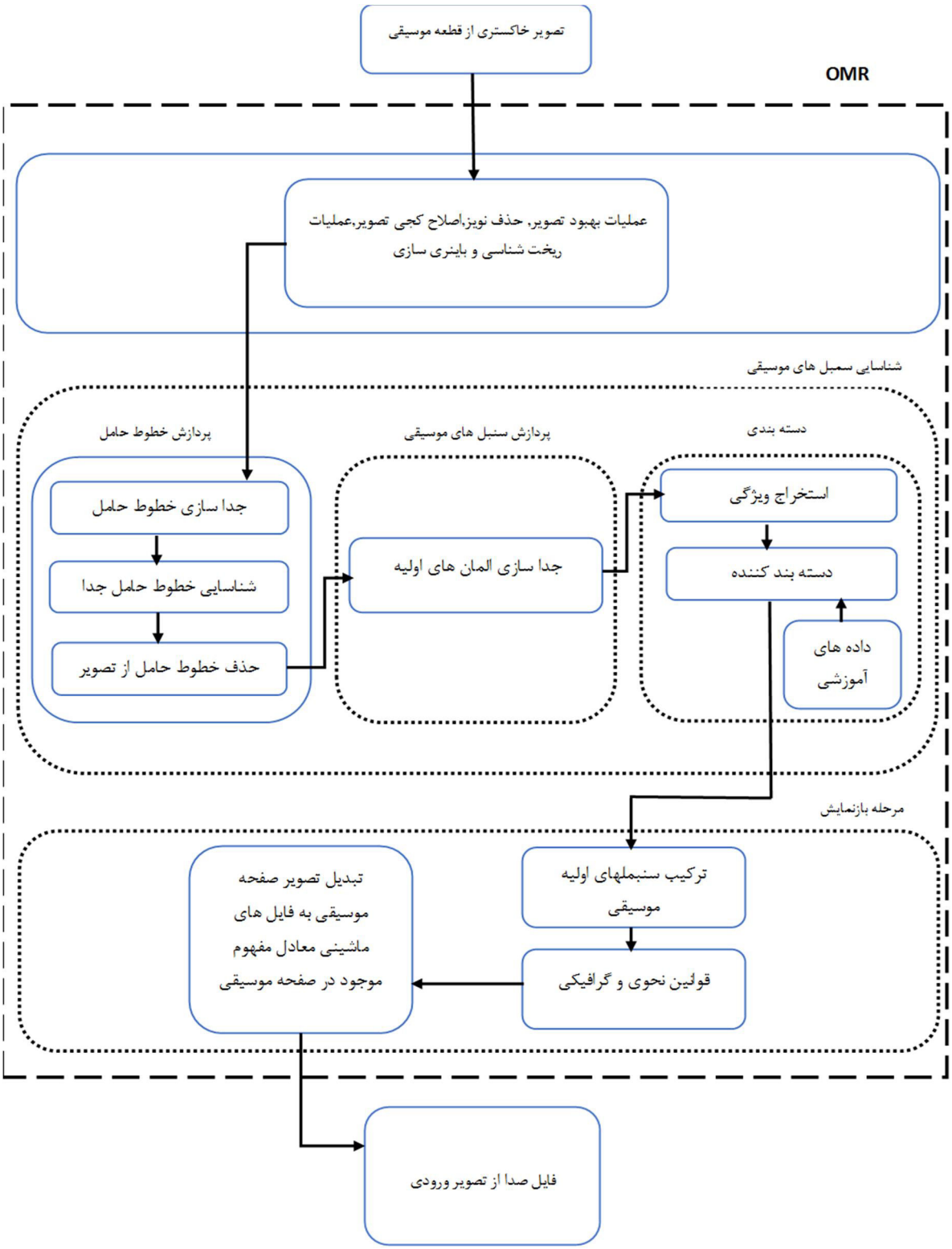
۲-۲- ۱- معرفی معماری یک سیستم OMR

به طور کلی تحقیقات روی سیستم های OMR در راستای سیستم های OCR می باشند به این معنا که هدف هر دو سیستم شناسایی پدیده های درون تصویر است که طبق قواعدی بکار گرفته شده است. حال در OMR هدف شناسایی نت موسیقی است ولی در OCR هدف شناسایی زبان محاوره است. لذا معماری آنها نیز شبیه به هم بوده البته با کمی تغییرات.

دیاگرام زیر، مراحل و معماری یک سیستم OMR را نشان می دهد:

---

[1] - Brain Bridge
[2] - Bell
[3] - W.Homenda



شکل۲-۱: معماری کلی یک سیستم ***OMR***

هماهنگونه در شکل ۲-۱ مشاهده می شود ورودی یک سیستم OMR یک تصویر خاکستری است که از یک قطعه موسیقی دریافت شده است.

مرحله بعدی شامل مراحل مورد نیاز در پیش پردازش است. این عملیات شامل تصحیح زاویه تصویر ورودی، حذف نویز، تاری تصویر، باینری سازی و برخی عملیات ریخت شناسی[1] است. معماری فوق جهت شناسایی قطعات دستنویس طراحی شده و ممکن است در خصوص نسخه های چاپی به برخی از پیش پردازشهای اشاره شده نیازی نباشد.

مرحله بعدی شامل کلیه پردازش هایی است که جهت شناسایی, جداسازی و حذف خطوط حامل صورت می پذیرد. مرحله بعدی شامل جدا سازی المان های زبان موسیقی از درون تصویر می باشد و در آخرین بخش از مراحل شناسایی سمبلهای موسیقی، استخراج ویژگی المانهای جداسازی شده مد نظر بوده و در نهایت دسته بندی المان های جدا سازی شده به منظور شناسایی الگوی آنها مورد توجه قرار می گیرد.

در بخش نهایی, ترکیب سمبلهای شناسایی شده و تبدیل آنها به زبان ماشین صورت می پذیرد. در نهایت خروجی که سیستم OMR شامل فایل گرافیکی از صفحه قطعه موسیقی یا یک فایل صوتی حاصل از شناسایی و ترجمه قطعه موسیقی می گردد که می تواند به صورت یک فایل MIDI یا Music XML ذخیره گردد.

روشها و الگوریتم های مختلفی جهت پیاده سازی هر قدم در مراحل مختلف توسط علاقمندان به موضوع OMR پیشنهاد شده است که می تواند باعث حصول نتایج بهتر در نتیجه نهایی شناسایی گردد.

اغلب روشهای بکار گرفته شده در تحقیقات قبلی در مرحله پیش پردازش، روشهایی عمومی و مطابق با روشهایی است که در سیستم های OCR بکار گرفته می شود و وجه تمایز ویژه ای در آن دیده نمی شود. اما الگوریتم و روشهای خاصی در مراحل بعدی بکار گرفته شده است که به آنها اشاره می گردد.

[1] - Morphological

۲-۲-۲- فاصله مرجع در خطوط حامل

اغلب سیستم های OMR مطالعه شده, براساس فاصله بین خطوط حامل و نیز عرض هر خط طراحی شده اند و این دو مقدار اساس کار شناسایی محسوب می شوند. در برخی از تحقیقات شناسایی خطوط حامل، حاصل شمارش ستون هایی است که بیشترین مقدار را در امتداد محور x دارا می باشند. از این رو بدست آوردن دنباله ای از ۰ و ۱ ها باعث ایجاد دنباله ای می گردد که در نهایت می توان فاصله خطوط حامل از یکدیگر را محاسبه کرد. این روش در اغلب ورودی ها نتیجه مطلوبی نخواهند داشت. همانگونه که در ادامه مباحث این فصل به آن اشاره خواهد شد بروز نویز در تصویر ورودی نه فقط در زمان دریافت تصویر بلکه در زمان باینری سازی تصویر نیز حاصل می شود. هماهنطور که در معماری کل یک سیستم OMR به آن اشاره شد، اغلب مطالعات صورت گرفته بروی سیستم های OMR، بدنبال یافتن و حذف خطوط حامل می باشند از این رو در زمان حذف خطوط حامل ممکن است به دلیل جلوگیری از حذف پیکسلهایی که المانهای موسیقی را تشکیل می دهند و روی خط حامل نیز قرار دارند باعث ایجاد نویز بیشتر و تخریب المان مربوطه گردند و چه بسا در برخی مواقع این عملیات خود باعث ایجاد نقاطی در تصویر گردد که در ادامه سیستم شناسایی کننده این اختلالات را به عنوان المان موسیقی شناسایی کند.

از طرفی در طی دریافت تصویر از ابزارهای مختلف امکان اینکه خطوط حامل در تصویر به صورت خط راست و کاملا در راستای افق نباشد نیز وجود دارد و کمتر مشاهده شده که در تصاویر قطعات موسیقی خطوط حامل کاملا موازی با هم قرار گرفته باشند.

ساده ترین دیدگاه شامل پیدا کردن ماکزیمم محلی روی ماتریس تبدیل رادون به صورت پویش افقی در پیکسلهای سیاه تصویر می باشد. با فرض راست و کاملاً افقی بودن خطوط حامل، این ماکزیمم های محلی مانریس تبدیل رادون, خطوط حامل را به ما نشان می دهند.

تعداد زیادی مقدار در راستای افق و در نتیجه چرخش تصویر حاصل می شود که می توان در زاویه ای که ماکزیمم محلی بالاترین مقدار را داراست را به عنوان مرجع انتخاب کرد. این روش در شرایطی که خطوط حامل کاملاً راست و در راستای افق باشند کاربرد دارند و در غیر اینصورت کارآئی نخواهد داشت.

میااو[1] و ناکانو[2] از روش انتقال هاف جهت پیدا کردن خطوط حامل استفاده کرده اند. اساس کار و استراتژی آنها براساس پویش عمودی خطوط بوده است. این فرآیند براساس گراف خط مجاور[3] ایجاد شده است. اغلب کارهای صورت گرفته براساس همین روش بنا نهاده شده اند.

شکل ۲-۲: فاصله بین خطوط حامل و عرض هر خط

فوجیناگا[4][۲۰] روشهای کدگذاری طول رانش[5] [۲۰] و آنالیز اجزای متصل[6] [۲۰] و روشهای بروجکشن را با هم یکی کرد. در این روش پس از پیاده سازی روش RCL جهت یافتن میزان نازکی خطوط حامل و فاصله بین آنها هر نقطه سیاه که بیش از دو برابر ارتفاع خطوط حامل باشد را از تصویر اصلی حذف می نمود. سپس به کمک روش اجزا متصل هر جزئی که عرض آن کمتر از میزان فاصله بین خطوط حامل بود را برش و سپس حذف می کند. پس از تصحیح کجی عمومی تصویر، اجزای بلند مانند خطوط اتحاد و نت های بلند حذف می شد که این برای سیستم OMR مناسب نبود. تمامی تحقیقات انجام شده به دلیل

[1] - Miyao.
[2] - Nakano.
[3] - Line Adjacency Graph (LAG).
[4] - Fujinaga.
[5] - Run-Length Coding (RLC).
[6] - Connected component.

تمرکز روی پردازشهای محلی[1] و نه سراسری[2] تصویر جهت پیدا کردن و حذف خطوط حامل دارای نواقصی می باشد که در نتیجه نهایی تأثیرات نامطلوبی می گذارند.

۲-۲- ۳- قطعه بندی و شناسایی المان های موسیقی

استخراج المانهای موسیقی براساس حذف خطوط حامل و بعد از آن اتفاق می افتد. فرآیند استخراج المانها براساس جداسازی و مکان یابی آنها در تصویر قطعه موسیقی صورت می پذیرد. در این مرحله مشکل اصلی کیفیت اجزای تصویر یا به عبارت دیگر اجزای تشکیل دهنده المانهای قطعه موسیقی است (غیر از خطوط حامل که می بایست پیدا و حذف گردند) که در زمان چاپ یا فرآیند دیجیتال کردن و یا باینری کردن تصویر اتفاق می افتد. پیچیدگی این مرحله نه فقط اعوجاج[3] ذاتی خطوط حامل، که انقطاع و روی هم افتادگی[4] و تفاوت در اندازه و شکل و منطقه قرارگیری اجزای تشکیل دهنده تصویر است.

معمول ترین رویکرد در قطعه بندی المانهای قطعات موسیقی، تجزیه سلسله مراتبی قطعات موسیقی است. یک قطعه موسیقی, ابتدا به کمک جداسازی خطوط حامل آنالیز شده و سپس اشکال گرافیکی تشکیل دهنده زبان موسیقی استخراج می شوند. اشکالی مانند سر نت ها[5] ، خطوط اتحاد و اتصال[6]، نقطه ها[7]، دنباله های نت[8] ، پرچم ها[9] و از این قبیل.

---

[1] - Local processing
[2] - Global processing
[3] - Distortion
[4] - Over Lap
[5] - Note head
[6] - Rests
[7] - Dots
[8] - Stems
[9] - Flags

گر چه در برخی مواقع خطوط اتحاد به سر نت ها چسبیده اند اما در زمان قطعه بندی لازم است این اتصالات از یکدیگر جدا شوند. ماهونی[1] [۲۰] با ساختن یک مجموعه کاندید از یک یا بیش از یک از انواع سمبل ها و بکارگیری یک توصیف کننده[2]، سعی در انتخاب کاندیدهایی که از تصویر استخراج شده اند و با مجموعه ساخته شده یکی می باشند، کرد.

روشهای دیگری نیز جهت قطعه بندی بکار برده شده، روشهایی مانند LAG, KNN، شبکه های عصبی، روشهای فازی، استفاده از تکنیکهای گشتاور مرکزی[3]، مدل مخفی مارکوف و نیز استفاده از اطلاعات هندسی در ورودیهای مختلف تصویر.

۲-۲- ۴- تولید نهایی بازنمایی از مرحله قطعه بندی و شناسایی

آخرین مرحله در سیستم های OMR ساخت خروجی است که یا به صورت فایلهای گرافیک قابل ویرایش باشد و یا به صورت ساختار داده موسیقی جهت اجرا. در اصل هدف این مرحله ایجاد ساختاری است که براساس زبان و قواعد زبان موسیقی که خط محور است (خطوط حامل) دقیقاً ترجمه و تفسیر نسخه اصلی تصویر دریافت شده از یک قطعه موسیقی باشد.

اگر با نگاه سیستم های OCR به این مقوله نگاه کنیم، این کار بسیار ساده به نظر می رسد. به دلیل اینکه طرح کار بطور برجسته ای به هم شبیه و تک بعدی است. اما در سیستم های OMR طرح کار بسیار پیچیده تر است. موسیقی ذاتاً دارای دو بعد است، صدا در محور عمودی و زمان در محور افقی. به عبارت دیگر موسیقی ایجاد فرکانسهایی در راستای زمان است که این زمان عنصر بسیار مهم در زیبایی قطعات موسیقی است. پس در نتیجه اطلاعات موقعیت هر المان در قطعات موسیقی بسیار حائز اهمیت

[1] - Mahoey
[2] - Descriptor
[3] - Central moment

است. اشکال گرافیکی که در ظاهر با هم شبیه هستند، ولی زمانی که در موقعیت ها و مکانهای مختلف استفاده می شود، دارای معانی مختلفی می باشند.

تحقیقات گسترده ای تا بحال در این زمینه صورت گرفته است. در بسیاری از آن محققین در تلاش جهت یافتن گرامر زبان موسیقی و شناساندن آن به سیستم بودند. گرامرهایی جهت ساخت خروجی های گرافیکی یا فایلهای صوتی که نتایج دقیقی را بتواند تولید کند. از جمله این سیستم ها می توان به CANTOR, CMN, DCG , اشاره کرد.

تفکیک دیگری که برای ترجمه قطعات موسیقی بکار گرفته شده و تحقیقات گسترده ای نیز برای آن صورت گرفته سیستم های مبتنی بر ترکیب قواعد موسیقی و اکتشافی و بخشهای هیستوگرام عمومی روی ردیف ها، ستونها برای هر جفت المان موسیقی می باشد. روسانت[1] [۲۰] و بلاچ[2] [۲۰] دو مرحله برای این تکنیک در نظر گرفته اند:

- پیدا کردن و جداسازی اشیاء و محاسبه توابع فرضیه
- اطلاعات وابسته به قرائن و قواعد موسیقی

مرحله اول شامل پردازش های سطح پایین بوده که به محاسبه میزان مطابقت بین یک شئی و اشیاء ای که در مجاورت آن قرار دارد می پردازد. برخی از این تطابقات و رابطه ها به شرح زیر می باشند:

ـ علائم عرضی[3] و نت: علائم عرضی قبل از سر نت ها قرار می گیرند.

ـ سر نت ها و نقاط: نقطه ها بعد یا بالای سر نت ها قرار دارند و با فاصله های مختلفی نیز نسبت به سر نت نمایش داده می شوند.

---

[1] - Rossant
[2] - Bloch
[3] - Accidental

ـ بین هیچ دو شئی دیگری, هیچگونه روی هم افتادگی وجود ندارد.

در مرحله قواعد نحوی هدف معرفی قوانین در ارتباط با تن های صدا، علائم عرضی و وزن[1] قطعه موسیقی می باشد. این گروه از علائم در تصاویر قطعات موسیقی به صورت یک دنباله مرتب و در مکانهای تصادفی بعد از کلید[2] آورده می شوند. در انتها, وزن شعر (تعداد ضرب در هر میزان) کنترل می شود.

۲-۳- شناسایی الگو

یادگیری ماشین، شاخه ای از هوش مصنوعی است که با طبقه بندی و توصیف مشاهدات سر و کار دارد. یکی از مباحث اساسی در یادگیری ماشین، شناسایی الگوست.

شناسایی الگو به طور کلی شامل دریافت داده های خام و تصمیم گیری براساس دسته بندی داده هاست. عملیات شناسایی الگو در واقع عملی است که انسانها در زندگی روزمره خود در زمینه های مختلف با آن سر و کار داشته و با توجه به مشاهدات خود، داده ها را براساس الگوهایی که در سیستم مدیریت دانش مغز خود وجود دارد، پردازش و نگهداری می نمایند. داده ها می توانند، بزرگی، کوچکی، سرما و گرما، مرغوبیت، عدم مرغوبیت و به طور کلی مشاهداتی که از دو منظر کیفی و کمی قابل دسته بندی باشند را شامل گردند.

در صورتی که برای داده های خام هیچ الگویی در ذهن انسان وجود نداشته باشد؛ امکان دسته بندی آن نیز در ذهن وجود نخواهد داشت لذا با آموزش و دریافت نمونه های مختلف از داده می توان الگوهایی را جهت دسته بندی در ذهن ایجاد نمود.

با توجه به توضیحات ارائه شده، اغلب مطالعات و پژوهش ها در حیطه یادگیری ماشین را می توان به دو دسته کلی تقسیم نمود.

---

[1] - Beat

[2] -Key

۱ـ یادگیری با نظارت

۲ـ یادگیری بدون نظارت

۲-۳-۱ یادگیری با نظارت

شناسایی کننده هایی که براساس یادگیری با نظارت عمل می نمایند، یک مدل پارامتریک را برای عملکرد خود به صورت پیش فرض تعریف می کنند و با استفاده از داده های آموزشی سعی می کنند آن پارامترها را تخمین بزنند. با تخمین پارامترهای مجهول، مدل سیستم شناساسی کننده شکل گرفته می تواند برای شناسایی نمونه های جدید مورد استفاده قرار گیرد.

۲-۳-۲- یادگیری بدون نظارت

در نقطه مقابل، شناسایی کننده هایی که براساس یادگیری بدون نظارت عمل می نمایند, از داده های آموزشی به صورت خام استفاده نموده و با دریافت یک نمونه آزمایشی از اطلاعات خام و بدون مدل داده های آموزشی, جهت شناسایی آن نمونه جدید بهره می گیرند.

در ادامه این فصل به معرفی انواع داده ها و نمونه هایی که در یک سیستم شناسایی کننده مورد استفاده قرار می گیرند می پردازیم. سپس با توجه قرابت موضوع پایان نامه با سیستم های شناسایی کننده نوری، به معرفی روشهای بکار رفته در اینگونه سیستم ها می پردازیم و خصوصیات هر یک از روشها را تشریح می نماییم.

۲-۴- انواع نمونه ها

داده های خام به کار رفته در سیستم های شناسایی کننده به طور کلی به ۲ دسته کلی تقسیم می شوند. این داده ها در واقع از انواع داده هایی است که ماشین قابلیت پردازش آن ها را داشته و منطبق با معماری ماشین می باشند. بسیاری از نمونه هایی که از جهان پیرامون درک می گردد، دنباله ای پیوسته از داده ها می باشند. پیوستگی این داده ها هم در زمان و هم در دامنه است. به همین دلیل برای انتقال آنها به ماشین نیاز به سیستمی جهت تبدیل داده های پیوسته به داده های گسسته جهت درک ماشین داریم.

با توجه به توضیحاتی که در بالا ارائه شد, ۲ دسته کلی داده ها به شرح زیر می باشند:

- سیگنالها

  مانند سیگنالهای صوتی و گفتار , سیگنالهای EEG , سیگنال های ماهواره ای و راداری.

- تصاویر دو بعدی

  مانند تصاور , متون تایپی , تصاویر پزشکی , تصویر پلاک خودرو و ...

در هر دو دسته بندی از داده های ورودی به سیستم شناسایی کننده؛ هدف شناسایی الگو است که شامل خصوصیاتی است از آن شی که در تمام نمونه های مربوط به یک دسته یا طبقه شبیه به هم هستند. با توجه به هدف یک سیستم شناسایی کننده، انتخاب خصوصیات, از پارامترهای مهم در رسیدن به هدف می باشد.

۲-۵- انتخاب ویژگی[1]

ویژگی یا خصوصیات یک نمونه در قالب یک بردار به نام بردار ویژگی[1] شناخته می شوند. مشخصه، یک ویژگی قابل اندازه گیری و منحصر بفرد در یک پدیده است. مشخصه ها گاه به طور مستقیم از یک پدیده

[1] - Feature selection

دریافت می شوند. به عنوان مثال در خصوص شناسایی میوه های کال و رسیده از مشخصه هایی مانند رنگ یا وزن که به صورت مستقیم از نمونه ها قابل دریافت است, استفاده می شود. ولی در برخی کاربردها مشخصه ها با انجام پردازشهایی قابل درک می باشند. به طور مثال ریتم ضربان قلب که برای دست آوردن ویژگی های آن نیاز به انجام پردازشهایی از جمله اعمال معادلات تفاضلی در نمونه ها می باشد.

با توجه به توضیحات ارائه شده درمی یابیم انتخاب صحیح ویژگیها از یک پدیده گامی اساسی در رسیدن به هدف سیستم شناسایی کننده است.

۲-۵-۱ خصوصیات یک مشخصه مناسب

۱ـ بازنمایی[۲] : یک مشخصه باید به خوبی پدیده را توصیف کند. و اطلاعات مختصر و مفید از آن پدیده را در اختیار ما قرار دهد.

۲ـ منحصر به فرد بودن[۳] : یک مشخصه باید مقادیر متفاوت برای کلاس های متفاوت و مقادیر مشابه برای پدیده های مشابه داشته باشد.

۳ـ قابل توصیف[۴] : یک مشخصه بایستی کاملاً قابل تفسیر و ترجمه برای یک فرد خبره باشد.

۴ـ مفید و مناسب[۵] : یک مشخصه مناسب بایستی دارای معنی و متناسب با خصوصیات یک پدیده باشد.

---

[1] - Feature vector
[2] - Representation
[3] - Characteristic
[4] - Interpretable
[5] - Suitable

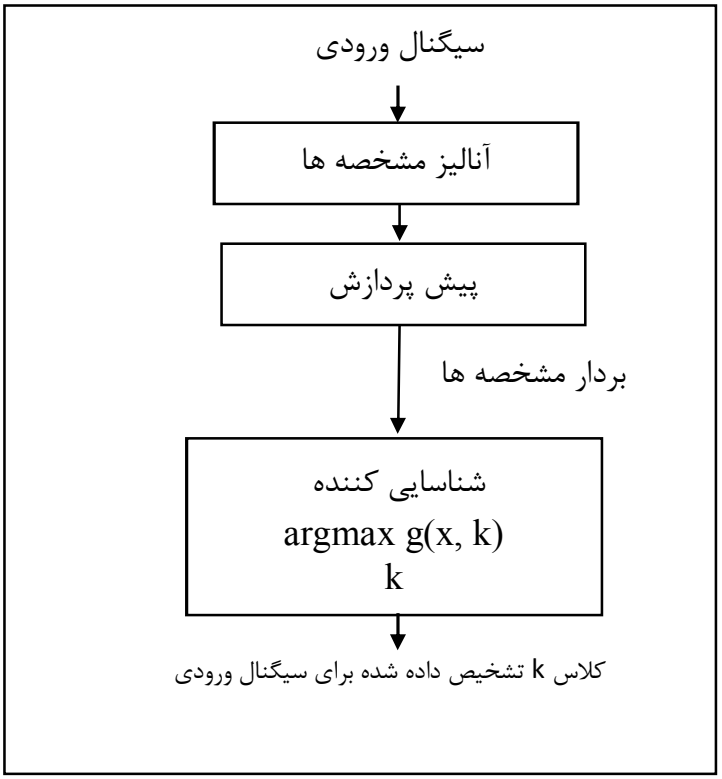
۲-۶- عملکرد کلی سیستم های شناسایی الگو

با توجه به توضیحات ارائه شده در بخشهای قبلی این فصل, حال به تشریح نسبتاً کامل ولی مختصر گونه عملکرد شناسایی کننده ها می پردازیم.

سیستم های شناسایی کننده یک سیگنال را که حاوی داده های خام از یک پدیده خاص است را از طرق مختلف دریافت و با آنالیز سیگنال ورودی و پیش پردازش های لازم، اقدام به تشکیل یک بردار x به عنوان بردار ویژگی سیگنال ورودی می نماید که در واقع نام این مرحله استخراج ویژگی[1] است.

پس از استخراج ویژگی ها یک تابع تشخیص دهنده به نام g (x, k) تعیین می کند که احتمال اینکه سیگنال ورودی به کلاس k تعلق داشته باشد، چقدر است. دسته k که بتواند بیشترین مقدار را برای تابع تشخیص دهنده g با بردار ویژگی x به وجود آورد، کلاس مربوط به سیگنال ورودی خواهد بود.

[1] - Feature Extraction



شکل ۲-۳ : شناسایی کننده آماری الگو

اغلب روشهای دسته بندی، از مراحل آموزش و آزمون کمک می گیرند. داده های دسته های مختلف به دو بخش داده های آموزشی[۱] و داده های آزمون[۲] تقسیم می شوند. جهت بالا بردن درجه اطمینان به سیستم شناسایی کننده هر دو مجموعه داده های آموزش و آزمون بایستی کاملاً از یکدیگر مستقل بوده و نمی بایستی هیچ گونه اشتراکی با یکدیگر داشته باشند. سپس براساس یک الگوریتم مشخص و بر اساس داده های آموزش, الگوریتم آموزش و در نهایت میزان دقت و خطای کارکرد الگوریتم با داده های آزمون سنجیده می شود.

[1] - Train set
[2] - Test set

۲-۶-۱ داده های آموزش

مجموعه داده های آموزش تا حد امکان بایستی کامل بوده و تمامی حالات ممکن یک پدیده را در خود جای داده باشد. بدیهی است جهت طراحی یک سیستم شناسایی کننده برای یک پدیده نمی توان برای هر تغییر در داده ها یک سیستم مستقل طراحی نمود, لذا در طراحی یک سیستم کامل شناسایی کننده لازم است سیستم را به گونه ای تعریف کرد که برای دسته بندی داده های مختلف مورد استفاده قرار گیرد.

۲-۶-۲- داده های تست

یک سیستم شناسایی کننده، براساس میزان کارآئی آن در دسته بندی داده های ورودی مورد ارزیابی قرار می گیرد. در صورتی که شناسایی کننده, داده ها را به خوبی دسته بندی کنند و در دسته های صحیح جای دهد، دارای کارآئی مطلوب بوده و در غیر اینصورت کارآئی آن مورد تردید خواهد بود.

برای ارزیابی یک سیستم شناسایی کننده از یک سری داده هایی که داده های تست نامیده می شوند استفاده می شود. همانگونه که توضیح داده شده این داده ها بایستی کاملاً مستقل از داده های آموزش بوده تا بتوان میزان کارآئی سیستم را در مواجه با داده های دیده نشده مورد ارزیابی قرار داد.

۲-۶-۳-داده های اعتبار سنجی

در برخی اوقات سیستم شناسایی کننده پس از آموزش توسط داده های آموزش، نتایج ضعیفی را در مرحله آزمون و بروی داده های تست از خود نشان می دهند. یکی از مهمترین دلایل این امر می توان

عدم در نظر گرفتن یک یا چند پارامتر تأثیرگذار دیگر در سیستم باشد که سیستم شناسایی کننده با استفاده از داده های آموزشی، آموزش ندیده است.

همچنین ارزیابی روشهای آموزش سیستم نیز باعث بهبود کارآئی گردیده و در مواردی که سیستم شناسایی بسیار منطبق[1] بر داده های آموزشی می گردد, این اتفاق منجر به نتایج نامطلوب در زمان تست سیستم با داده های آموزشی می گردد.

لذا در چنین حالاتی از مجموعه داده ها به نام داده های اعتبار سنجی[2] استفاده می شود. داده های اعتبار سنج مانند داده های تست برای ارزیابی استفاده شده و با هدف بهینه کردن نتیجه عملیات شناسایی بر روی مجموعه اعتبار سنجی, پارامترها را تنظیم می کند.

۲-۶-۴ انواع شناسایی کننده ها

از انواع شناسایی کننده ها با توجه به نحوه کار شناسایی کننده ها و منابعی که در اختیار دارند می توان به دو مدل کلی شناسایی کننده بر مبنای مدل[3] و بدون مدل[4] اشاره کرد.

۲-۶-۴-۱- شناسایی کننده های بر مبنای مدل

شناسایی کننده هایی که بر مبنای مدل طراحی شده اند, از داده های آموزشی برای تخمین پارامتر های مدل در نظر گرفته شده استفاده می نمایند و سپس به کمک داده های تست اقدام به ارزیابی مدل بدست آمده می کنند. از شناسایی کننده های معروف بر مبنای مدل می توان به شناسایی کننده های بیز[5] اشاره

---

[1] - Over fit
[2] - Validation data
[3] - Model base
[4] - Model free
[5] - Bays

نمود که اساس کار آن شناسایی از طریق روش های آماری است. این دسته از شناسایی کننده های بیز از نظر میزان دقت و سرعت دارای قابلیت های بالا و حتی قابل رقابت با برخی شناسایی کننده ها مانند درخت های تصمیم می باشند.

از دیگر شناسایی کننده های بر مبنای مدل می توان به شناسایی کننده های بر مبنای فاصله اشاره کرد. اساس کار شناسایی کننده های بر مبنای فاصله مشابه شناسایی کننده های بیز بوده با این تفاوت که فاصله نمونه آزمایشی تا بردار میانگین یک کلاس را مبنای کار خود قرار می دهد. معیار های مختلفی جهت سنجش فاصله وجود دارد که از جمله آنها می توان به فاصله های اقلیدسی , فاصله بلوکی اشاره کرد که بنابر کاربرد و صورت مسئله می توان از یکی از این روشها استفاده نمود.

۲-۶-۴-۲ شناسایی کننده های بدون مدل

در شناسایی کننده های بدون مدل تعداد دسته با اعضای مشخص وجود دارد که بعضی ویژگی های این اعضاء نیز مشخص است. با توجه به اینکه برخی ویژگی های اعضاء مشخص است, تمام نمونه های آموزشی معنی پیدا می کنند.

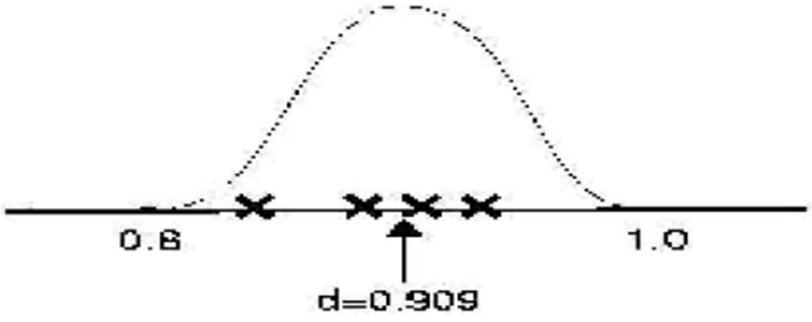
به عنوان مثال فرض کنید اطلاعاتی از رنگ و وزن ۱۰ نوع میوه مختلف در اختیار داریم. این اطلاعات شامل ۱۰۰۰۰ نمونه می باشد. با توجه به خصوصیات رنگ و وزن, برداری دوبعدی از داده ها در اختیار داریم. بر طبق این نوع شناسایی کننده ها, مدل بر اساس همین داده های ورودی که به عنوان داده های آموزشی نیز استفاده می شوند ایجاد شده و دسته بندی داده های جدید بر اساس همین نمونه های آموزشی انجام می گردد و هیچ مدلی برای بررسی داده های ورودی وجود نخواهد داشت.

از پر کاربرد ترین شناسایی کنند های بدون مدل می توان به شناسایی کننده های نزدیک ترین همسایه[1] و شناسایی کننده های K نزدیکترین همسایه[2] اشاره نمود.

۲-۶-۵- ترکیب شناسایی کننده ها

انگیزه اصلی ترکیب دسته بندی کننده ها این است که ما هنگام طراحی یک سیستم یادگیر، انتخاب های فراوانی داریم.

مانند نحوه نمایش، پارامترهای یادگیر, داده های آموزشی. این تنوع باعث می شود نوعی واریانس در عملکرد سیستم وجود داشته باشد. در نتیجه اگر سیستم های مختلفی داشته باشیم و از نتایج آن استفاده شود این امکان وجود دارد که توزیع خطا حول آن به حداقل برسد.



شکل ۲-۴- کاهش توزیع خطا با ترکیب دسته بندها

[1] - Nearest neighbor
[2] - K Nearest neighbor

۲-۷- معرفی سیستم های بازشناسی نوری علائم (OCR)

نویسه خوانهای نوری[1] به تکنیک هایی اطلاق می شود که در تصاویری که از طریق اسکن و دوربینهای دیجیتال بدست می آید نواحی متنی تصویر ورودی را تشخیص داده, سپس این نواحی از تصویر را به متن تبدیل می نماید. از مهمترین و اصلی ترین مزایای تبدیل تصویر به متن، امکان جستجو، ویرایش، افزایش دسترسی به اطلاعات و کاهش فضای ذخیره سازی می باشد.

اسناد حاوی متن اگر چه برای یک انسان حاوی اطلاعات و مطالب مفید است ولی همین اسناد برای ماشین جز یک تصویر که متشکل از پیکسلها با طیف رنگ های گوناگون است, نمی باشد. از نظر ماشین تصویر یک سند حاوی متن با یک تصویر منظره یکسیت. زیرا هر دو تصویر نمایش دو بعدی از پیکسلها می باشند.

با توجه مقدمه ای که بیان شد چرائی اهمیت بینایی ماشین و شناسایی الگو نمایان می شود. ایجاد قابلیت ماشین جهت شناسایی و تبدیل تصاویر حاوی متن, به متن قابل ویرایش یکی از مسائل مورد توجه در زمینه بینایی ماشین و شناسایی الگو می باشد.

مانند بسیاری از سیستم های هوشمند، سیستم های OCR نیز دارای پیچیدگی های بسیاری می باشند. این پیچیدگی در زبانهای مختلف نیز دارای درجات مختلف است. به عنوان مثال در زبانهای لاتین مانند انگلیسی، به دلیل اینکه حروف از یکدیگر جدا بوده و اتصالی بین آنها برقرار نیست به مراتب آسانتر از زبانهای مانند فارسی و عربی است. زیرا زبانهای فارسی و عربی به دلیل چسبندگی حروف فرآیند شناسایی و قطعه بندی[2] را با مشکل مواجه می سازند و درک آنها توسط ماشین از پیچیدگی های بیشتری برخوردار است.

---

[1] - Optical Character Recognition

[2] - Segmentation

۲-۷-۱- انواع سیستم های OCR

سیستم های OCR را می توان از منظر نوع ورودی به دو دسته تقسیم نمود.

۱ـ سیستم های OCR بر خط[1]

۲ـ سیستم های OCR برون خط[2]

در سیستم های بر خط همراه با نوشتن به شناسایی صورت می گیرد و در سیستم های برون خط از یک سند از قبل نوشته شده استفاده می شود و از آن یک متن قابل ویرایش ساخته می شود.

به طور کلی اغلب سیستم های نویسه خوان برون خط بوده و فرآیند آنالیز متون را به صورت ایستا[3] دنبال می کنند. اما سیستم های نویسه خوان بر خط هم جایگاه ویژه ای داشته و در واقع فرآیند پیچیده تر را نسبت به سیستم های برون خط دنبال می کنند.

درک حرکت قلم، قطعاتی که در آن قلم فشرده و مکانهایی که قلم در آن رها می شوند، درک جهت نوشتن قلم، و الگوهای فشردن و رها کردن قلم، از جمله پردازشهایی است که سیستم های نویسه خوان برخط بایستی انجام دهند. نام های دیگری نیز برای این گونه سیستم ها در نظر گرفته شده است مانند، سیستم های هوشمند شناسایی کاراکتر[4]، سیستم های بر خط شناسایی کاراکتر[5]، سیستم های پویای شناسایی کاراکتر[6].

---

[1] - Online
[2] - Offline
[3] - Static
[4] - Intelligent Character Recognition
[5] - Online Character Recognition
[6] - Dynamic Character Recognition

۲-۷-۲- کاربردهای OCR

سیستم های نویسه خوان نوری برای بسیاری کاربردها طراحی شده اند. صنایع و بخش های مختلفی از سیستم های نویسه خوان استفاده می نمایند. صنایعی مانند: بیمه، بانک ها، سیستم های کنترل ترافیک و بسیاری از بخش هایی که برای مقاصد خاص از این سیستم ها بهره می گیرند.

در زیر به چند نمونه از کاربردهای آن اشاره می کنیم:

۱ـ کنترل ترافیک (خواندن اتوماتیک پلاک خودروها)

۲ـ سیستم های تبدیل متن به گفتار (برای کمک به نابینایان و یا کم بینایان)

۳ـ فرم خوانی (مانند اطلاعات درج شده در چک های بانکی، فرم های استخدامی)

۴ـ تبدیل کتب قدیمی به متون قابل جستجو و قابل ویرایش

۵ـ تبدیل دستنوشته در یک سیستم بلادرنگ[1] جهت کنترل کامپیوترها

در راستای کاربردهایی که برخی از آنها در این بخش شرح داده شد، بعضی کاربردها نیز وجود دارد که مبتنی بر عملیاتی است که از روی اسناد چاپ شده صورت گرفته و از لحاظ عام بودن نسبتاً به متون چاپی به زبانهای رایج دنیا دارای عمومیت بینابینی می باشند. یکی از این کاربردها در این پایان نامه محور اصلی موضوع بوده و آن شناسایی نت های چاپ شده موسیقی است. همانگونه که در فصل اول این پایان نامه به آن اشاره شد,

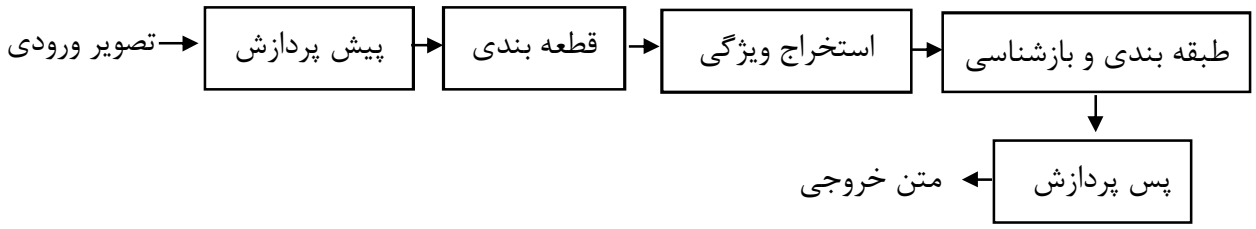
تفسیر و ترجمه متون موسیقی خود به تنهایی نیازمند به داشتن مهارت های خاص در زمینه دانش و هنر موسیقی است. از این رو کاربردهای سیستم های مشابه در زمینه OCRها را می توان جهت پیاده سازی نویسه خوانهای موسیقی[2] نیز در نظر گرفت.

[1] - Real-time

[2] - OMR : Optical Music Recognition

۲-۷-۳ معرفی بخشهای مختلف در پیاده سازی سیستم های OCR

بلوک دیاگرام کلی یک سیستم OCR از ۵ بخش تشکیل شده است. در برخی منابع سیستم های OCR گاه به بیش از ۵ بخش تقسیم شده است اما به طور کلی می توان اینگونه سیستم ها را به ۵ بخش مطابق دیاگرام زیر تقسیم نمود :



شکل ۲-۴- شمای کلی یه سیستم نویسه خوان نوری

۲-۷-۳-۱-پیش پردازش[1]

فرآیند تحلیل داده ها با بارگذاری داده ها و جدا سازی داده های خوب از داده های بد یا همان پیش پردازش داده ها آغاز می شود. این مرحله گام دوم از تحلیل داده ها بوده و تأثیر به سزایی در میزان کارآئی مراحل بعدی تحلیل دارد.

در مباحث مربوط به پردازش تصویر با داده های دو بعدی سر کار داریم که عوامل مختلفی می تواند باعث ایجاد داده های نامطلوب در آن گردد.

عملیات مختلفی در مرحله پیش پردازش ممکن است انجام شود تا خروجی آن یک تصویر مناسب جهت انجام مراحل بعدی گردد. از جمله این مراحل می توان از: باینری[2] کردن تصویر، تصحیح زاویه تصویر ورودی، حذف نویز، هموار سازی، نازک سازی، فیلترگذاری مکانی و فرکانسی نام برد.

[1] - Preprocessing
[2] - Binary

۲-۷-۳-۲-کاهش نویز[1]

نویز، اعوجاج ناخواسته ای است که در یک سیگنال و یا یک تصویر به وجود می آید. به عبارت دیگر، نویز عاملی است که ممکن است در عملیات یک سیستم, اختلال ایجاد کند. عوامل متعددی باعث ایجاد نویز در تصویر می گردد.

در تصاویر دیجیتال ممکن است نویز از راههای متعددی به تصویر تحمیل شود. در تصاویر دیجیتال مکانیزم اخذ داده, می تواند منبع نویز باشد. همچنین انتقال الکترونیکی داده های تصویر نیز می تواند منبع دیگری برای ایجاد نویز باشد. ابزاری های پویشگر[2] نیز از جمله منابعی است که می تواند تواند باعث ایجاد نویز در تصویر گردد.

یکی از مهمترین عوامل و مؤثرترین آنها در زمان انجام عملیات باینری کردن تصویر اتفاق می افتد. کلیه روشهای باینری سازی با توجه به کیفیت تصویر ورودی می تواند باعث ایجاد نویز گردد. رفع نویز از تصویر می تواند به صورت سراسری[3] و یا محلی[4] پیاده سازی سازی شود. نویزهای سراسری را می توان در همان مرحله پیش پرداز حذف نمود ولی برخی از نویزهای محلی را بهتر است در حین قطعه بندی و در برخی موارد در زمان بازشناسی کاهش داد و یا از بین برد.

کاهش نویز همچنین باعث کم شدن حجم فایل تصویر که به نوبه خود، کاهش زمان مورد نیاز برای پردازش و ذخیره سازی های آینده را در پی خواهد داشت.

---

[1] - Noise Reduction
[2] - Scanner
[3] - Global
[4] - Local

۲-۷-۳-۲-۱اصلاح کجی تصویر

در بسیاری از متون چاپی و دستنویس , در زمان تصویر برداری با مشکل کجی یا چرخش تصویر ورودی مواجه می شویم. چرخش تصویر ورودی خود می تواند باعث ایجاد اختلال در عملیات قطعه بندی گردد.

شکل ۲-۵: نمونه تصویر از نت موسیقی که به علت کجی تصویر, خطوط حامل در راستای افق نمی باشند.

با توجه به اینکه در اغلب روش های شناسایی متون خط مبنا (خط کرسی) مبنای قطعه بندی و سایر عملیات می باشد, وجود چرخش می تواند باعث ایجاد نتایج نامطلوب گردد. روش های مختلفی جهت اصلاح چرخش و کجی تصویر وجود دارد. از جمله آنها می توان از اطلاعات هیستوگرام[1] تصویر و یا تبدیلاتی مانند تبدیلات هاف[2] و یا تبدیلات رادون[3] نام برد.

۲-۷-۳-۲-۲-باینری کردن تصویر

برخی سیستم های نویسه خوان نوری با سطوح خاکستری تصویر کار می کنند اما اغلب نویسه خوان ها, ابتدا تصویر را باینری کرده و سپس اقدام به استخراج ویژگی می نمایند.

فرآیند باینری کردن در برخی مواقع باعث حذف نویز از تصویر شده و در برخی موارد نیز عملیات باینری کردن خود باعث ایجاد نویز می گردد. البته باید اشاره نمود که کیفیت باینری نمودن تصویر به کیفیت

[1] - Histogram
[2] - Hough transform
[3] - Radon transform

تصویر ورودی بستگی فراوان دارد. روش های بسیاری برای باینری کردن تصاویر وجود دارد که از مهترین و بارزترین آنها می توان به روش اتسو[1] اشاره کرد. خروجی الگوریتم اتسو در واقع حد آستانه[2] بوده و در نهایت پیکسل هایی را که شدت روشنایی آنها از آستانه بیشتر است روشنتر و پیکسلهایی که از حد آستانه کمتر می باشند را خاموش می نماید.

۲-۷-۳-۲- قطعه بندی تصویر[3]

قطعه بندی، تصویر را به عناصر یا نواحی سازنده تصویر تقسیم می کند. به عبارت دیگر قطعه بندی جداسازی بخشهایی از تصویر است که ویژگی و یا خاصیت مشترکی با یکدیگر دارند. به این ترتیب به هر بخش از تصویر برچسبی داده می شود و آن بخشهایی که یک برچسب را افراز می کنند، دارای مشترکات ویژه ای می باشند.

معیاری که قطعه بندی براساس آن انجام می شود بستگی به مسئله مورد نظر دارد. در بسیاری منابع قطعه بندی را یکی از مشکل ترین مراحل پیش پردازش معرفی می کنند. دقت قطعه بندی احتمال نمایش موقعیت یا شکست روشهای تحلیل کامپیوتری را تعیین می کنند.

برای قطعه بندی تصویر می توان از معیارهای مختلفی مانند سطح خاکستری، رنگ، بافت , عمق و حرکت استفاده کرد. به طور کلی در تقسیم تصویر I به x قطعه $R_1, R_2, ..., R_n$ خواص زیر می بایست برقرار باشد.

- با اجتماع تمام قطعه ها، تصویر I به طور کامل پوشش داده شود، یعنی پیوستگی زیادی در هر یک از قطعه ها وجود داشته باشد. $U_{i=1}^{n} = R_i = I$

[1] - Otsu
[2] - Threshold
[3] - Image Segmentation

- فضای تحت پوشش دو قطعه j,i $(i \neq J)$، با هم هیچ اشتراکی نداشته باشند؛ یعنی $R_i \cap R_j = \phi$، (نماد ϕ به معنای تهی است).
- برای هر قطعه R_i $(i \in \{1,2,...,n\})$، باید ویژگی $Q(Ri) = true$ برقرار باشد. این ویژگی بیان می کند که تمامی پیکسلهای قرار گرفته در قطعه R_i، معیار خاصی را برآورده می کند. به عنوان مثال، تمامی پیکسلهای یک قطعه، رنگ یکسانی داشته باشند.
- برای هر دو قطعه همسایه R_j, R_i $(i \neq j)$ ویژگی $Q(R_i \cup R_j) = False$ برقرار باشد. یعنی دو قطعه متفاوت، نباید معیار یکسانی را برآورده کند. برای مثال، دو قطعه متفاوت باید رنگ متفاوتی داشته باشند.

۲-۷-۳-۳- استخراج ویژگی

مسئله انتخاب ویژگی یکی از مسائل است که در مبحث یادگیری ماشین و همچنین شناسایی آماری الگو مطرح است. این مسأله در بسیاری از کاربردها (مانند طبقه بندی) اهمیت به سزائی دارد، زیرا در این کاربردها تعداد زیادی ویژگی وجود دارد، که بسیاری از آنها بلااستفاده هستند و یا اینکه بار اطلاعاتی چندانی ندارند. حذف نکردن این ویژگی ها مشکلی از لحاظ اطلاعاتی ایجاد نمی کند ولی بار محاسباتی را برای کاربرد مورد نظر بالا می برد و علاوه بر آن باعث می شود اطلاعات غیرمفید نهایی را به همراه داده های مفید ذخیره کنیم.

۲-۷-۳-۳-۱ نحوه استخراج ویژگی

در استخراج ویژگی ها از یک پدیده، می توان بردار ویژگیها را از کل سیگنالی که به عنوان نمونه دریافت می کنیم استخراج نمود. در مقابل می توان ابتدا اقدام به قطعه بندی نمونه کرد و با جداسازی بخش مورد نظر از درون پدیده، ویژگیهای مورد نظر خود را از درون آن قطعه جدا شده, استخراج نمود.

مشخصه های استخراج شده با این دو روش را به نام مشخصه های براساس ظاهر و مشخصه های براساس قطعه می نامند. از بارزترین مشخصه هایی که براساس قطعه بندی تصویر بسیار پر کاربردند می توان به مشخصه های هندسی[1] و مشخصه های گشتاوری[2] اشاره نمود. که پس از آنکه در یک تصویر یک شیئ قطعه بندی شد و محدوده آن از تصویر اصلی جدا گردید، می توان مشخصه های هندسی و گشتاوری توصیفی از شکل ظاهری آن قطعه ارائه کرد.

مشخصه های هندسی اطلاعاتی خاص را در ارتباط با موقعیت اجسام درون یک تصویر، مانند شکل آنها یا پیکربندیشان نشان می دهند. از انواع مشخصه های هندسی می توان به مرکز ثقل، مستطیل پیرامون شی اشاره نمود. مشخصه های گشتاوری، مشخصه هایی هستند که با استفاده از ممانهای هندسی به دست آمده و نسبت به تبدیلات هندسی مانند تغییر اندازه، چرخش و ثابت می مانند.

۲-۷-۴- طبقه بندی و بازشناسی

طبقه بندی و بازشناسی برای انسان به ظاهر امری بسیار ساده است. انسان به راحتی می تواند چهره ها را تشخیص دهد، به وسیله حس بویایی فرق بین رایحه خوب و بد را از هم تمیز دهد حتی با حس چشایی مزه ها و طعم های مختلف را در مغز خود طبقه بندی و سپس شناسایی نماید. همه انسان ها در طول روز

[1] - Geometric segmentation
[2] - Moment segmentation

بارها و بارها اقدام به طبقه بندی و شناسایی پدیده های پیرامون خود می نمایند و حتی شاید کمتر در خصوص نحوه و مکانیزم این قدرت, فکر و تعمل نموده باشند.

در زمینه دسته بندی و بازشناسی نیاز به یکسری مشخصه هایی داریم که از دل یک پدیده استخراج می شود (در خصوص مشخصه ها و نحوه استخراج آنها در بخش قبل توضیحات ارائه شد). حال که مشخصه های مفید را از یک پدیده جدا کردیم و در بردار مشخصه ها یا بردار ویژگی ذخیره نمودیم، نیاز به روشهایی داریم تا به کمک آنها عملیات دسته بندی و طبقه بندی را انجام دهیم در نهایت با نوعی برچسب گذاری روی هر دسته برای آن نامی در نظر می گیریم. در سیستم های OCR که هدف آن تشخیص کلمات و حروف می باشد نیز از همین رویه استفاده می گردد.

روشهای متنوع بسیاری تا بحال توسط محققان و پژوهشگران ابداع گردیده که به برخی از آنها که در OCR نیز کاربردهای فراوانی دامنه اشاره می کنیم.

بازشناسی نوشتار به دو شاخه بسیار کلی تقسیم می شوند:

- بازشناسی نویسه های گسسته
- بازشناسی نویسه های پیوسته

بازشناسی نویسه های گسسته به دلیل ساده تر بودن و امکان مقید کردن کاربر در نحوه نگارش، موفقیت بیشتری داشته است و امروزه سیستم های مبتنی بر بازشناسی حروف گسسته و ارقام رایج شده اند. به عنوان مثال اسناد چاپی که به زبان لاتین می باشند یا فرم هایی که به صورت گسسته و در خانه های مجزا از یکدیگر تکمیل می گردند.

نوشتار پیوسته خود به دو دسته دستنویس و چاپی تقسیم می شوند. بازشناسی در نوشتار پیوسته به دلیل پیچیدگی زیاد و کاربرد محدود، هنوز تا رسیدن به یک وضعیت مطلوب فاصله دارد. اما بازشناسی متون پیوسته چاپی، به دلیل کاربرد زیاد و پیچیدگی کمتر، مورد توجه بیشتری قرار گرفته و موقعیتهای

بسیاری داشته اند به طوری که در حال حاضر برای بسیاری از زبانهای دنیا، سیستم های تجاری نسبتاً کارآمدی وجود دارد.

طبقه بندی و بازشناسایی آخرین مرحله پردازش در یک نویسه خوان نوری است. در این مرحله هر حرف شناسایی می گردد و به کلاس حرفی مناسب ارجاع داده می شود. در ادامه دو رویکرد مختلف برای دسته بندی و بازشناسی حروف مورد نظر را مورد بررسی قرار می دهیم. ابتدا، بازشناسی از طریق تئوری تصمیم گیری مورد بررسی قرار می گیرد.

این روشها زمانی مورد استفاده قرار می گیرند که مشخصه های یک کاراکتر بتوانند به صورت عددی توسط بردار ویژگی معرفی گردند.

ممکن است ما مشخصه های الگو داشته باشیم که از ساختار فیزیکی حروف مشتق شده اند که نمی توانند به آسانی به صورت عددی اندازه گیری شوند. در این حالات، ارتباط بین حروف در هنگام تصمیم گیری روی عضویت کلاس باید مورد توجه قرار گیرد. برای مثال اگر ما بدانیم که یک حرف از یک پاره خط عمودی و یک پاره خط افقی تشکیل شده است، ممکن است این حرف L یا T باشد. در این صورت ارتباط بین دو پاره خط نیز برای تشخیص حرف لازم است لذا یک رویکرد ساختداری نیاز خواهد بود.

۲-۷-۴-۱ روشهای تشخیص الگو

روش های معمول در تشخیص الگو به طور اعم و نوشته ها به طور خاص را میتوان به دودسته روش های ساختاری[1] و روش های آماری[2] تقسیم بندی کرد . در روشهای ساختاری برای تشخیص الگو از خواص هندسی و توپولوژی آن نظیر قوس ها, حلقه ها, زوایا و امتداد های موجود در تصویر استفاده می گردد. در روش های آماری به هر الگو یک بردار در یک فضای برداری نسبت داده می شود که به آن فضای ویژگی

[1] - Structural
[2] - Statistical

می گویند و مولفه های بردار ویژگی الگو با استفاده از اندازه گیری های متعددی که روی الگو انجام می شود به دست می آیند. فضای ویژگی الگو با استفاده از اطلاعاتی که قبلا در اختیار سیستم تشخیص الگو قرارداده شده , به طبقات و نواحی متعددی تقسیم شده است. هدف از تشخیص یک الگو بخصوص تعیین طبقه یا ناحیه ای است که بردار الگو در آن قرار دارد. به این عمل طبقه بندی[1] و به سیستم تشخیص الگو, طبقه بندی کننده[2] می گویند. اطلاعات لازم جهت تقسیم بندی فضای ویژگی الگو شامل توابع چگالی احتمال شرطی و غیر شرطی طبقات و الگوهاست. اما این توابع همواره موجود نیستند و تنها اطلاعاتی که از طبقات و الگو ها در دسترس است, تعدادی زوج متناظر به صورت طبقه یا الگو است. در روش های آماری سعی می شود با استفاده از این اطلاعات توابع چگالی احتمال لازم تقریب زده شود. اما تقریب توابع چگالی شرطی مشکل است , خصوصا اگر ابعاد فضای ویژگی بزرگ باشد.

روش دیگری که به لحاظ سادگی ودقت, توجه زیادی در سال های اخیر به خود جلب کرده, استفاده از شبکه های عصبی مصنوعی[3] است. هر شبکه های عصبی مصنوعی از تعداد زیادی پردازنده ساده تشکیل شده است که به دلیل تشابه با شبکه عصبی بیولوژیکی, نرون نامیده می شود. بر حسب نوع ساختاری که برای این شبکه ها انتخاب می شود. هر پردازنده می تواند به تعدادی یا به تمام پردازنده های شبکه متصل شود. ارتباط بین هر دو پردازنده با یک وزن خاص صورت می گیرد, به طوری که W_{ij} وزن اتصال بین خروجی نرون j ام و ورودی نرون i ام است. بسته به اینکه نرون jام در تحریک نرون i ام اثر محرک یا بازدارنده داشته باشد, W_{ij} مثبت یا منفی خواهد بود و اگر نرون i ام از نرون jام هیچگونه سیگنالی دریافت نکند, W_{ij} برابر صفر خواهد شد.

[1] - Classification
[2] - Classifier
[3] - Artificial Neural Network

در خصوص شبکه های عصبی, انواع آن و نیز نحوه کار آنها در شناسائی الگو و دسته بندی در فصل بعدی به تفضیل خواهیم پرداخت.

۲-۷-۵- پس پردازش

با بازشناسی تمام تصاویر اجزای متنی یک بلوک، متن متناظر با هر یک از آنها مشخص شده است. برای بدست آوردن متن نهایی باید پس پردازشهایی روی این متن انجام شود. به عنوان مثال از به هم چسباندن حروف بازشناسی شده در مرحله قبل، زیر کلمات و کلمات تشکیل می شوند. تشکیل این زیر کلمات و کلمات با استفاده از موقعیت آنها در تصویر، یک مرحله پس پردازش است. مرحله بعدی می تواند سنجش اعتبار این کلمات با استفاده از واژه نامه باشد.

۲-۸- جمع بندی و نتیجه گیری

در این فصل به کلیات سیستم های نویسه خوان نوری موسیقی و علائم و سیستم های شناسایی الگو پرداخته شد. همچنین پیچیدگی هایی که در شناسایی قطعات موسیقی و متون چاپی برای ما وجود دارد نیز بیان گردید.

منابع بسیار محدودی در خصوص شناسایی زبان موسیقی توسط ماشین تا بحال ارائه گردیده که هر یک دارای

کم و کاستی هایی می باشند از طرفی مشابهت هایی را بین شناسایی کننده های زبان محاوره ای با زبان موسیقی میتوان متصور شد که در واقع برخی از آنها اساس کار ما در این پروژه خواهد بود. اما قواعد و قوانین حاکم در زبان موسیقی در برخی مواقع باعث ایجاد پیچیدگی هایی به مراتب بالاتر از اسناد چاپی است. در بسیاری از علائم موجود در زبان موسیقی نکاتی نهفته است که در صورت عدم کنترل و شناسایی

آنها ادامه کار را با مشکل مواجه نموده و یا اثرات سوء در نتایج کار خواهد داشت اما به لحاظ کلی همانگونه که بیان شد مشترکاتی نیز در این خصوص وجود دارد که میتواند ما را در رسیدن به هدف یاری نماید.

فصل سوم : روش های پیشنهادی

۳-۱- مقدمه

همانگونه که در فصل قبل و در بحث معرفی معماری کلی یک سیستم OMR به اشاره شد, در این فصل به جزئیات و روش هایی که از آنها به منظور انجام مراحل مختلف پیاده سازی و معماری سیستم از جمله اصلاح کجی تصویر , باینری سازی , قطعه بندی کلی, پیدا کردن خطوط حامل, حذف خطو حامل, قطعه بندی بمنظور جدا اسازی المان های موسیقی و نیز سیستم شناسایی کننده و در نهایت روشی که برای اجرای نهایی قطعه موسیقی بکار گرفته شده می پردازیم و نحوه کار هر مرحله به طور خاص بررسی می نمائیم.

به دلیل اهمیت بالای خطوط حامل در زبان موسیقی لازم است قبل از هر عملیات نسبت به صحت خطوط مذکور در تصویر ورودی اطمینان حاصل نمائیم. یقینا یکی از مهمترین ارکان موجود در زبان موسیقی وجود خطوط حامل می باشد لذا در صورت وجود هر گونه اعوجاج در ایندسته از خطوط قطعا در مراحل شناسایی المان های دیگر موسیقی با مشکل مواجه خواهیم شد.

لذا قبل از هر عملیاتی لازم است تصویر را از منظر صحت زاویه ورودی مورد بررسی دقیق قرار داده و سپس اقدام به عملیات تقطیع تصویر نمائیم. از آنجا که موقعیت دقیق قرار گیری خطوط حامل کمک شایانی به یافتن نت های موسیقی می نماید.

۳-۲- تصحیح زاویه تصویر ورودی (نرمال کردن تصویر)

همانگونه که در فصل ۲ و در بخش پیش پردازش آورده شد، اعمال برخی عملیات پیش از پردازش اصلی روی تصاویر، ما را در رسیدن به هدف کمک کرده و بسیاری از خطاهای حین پردازش اصلی را از بین می

برد. یکی از این پیش پردازشها تصحیح زاویه تصویر ورودی بوده که در صورت اعمال آن روی تصویر باعث می شود بسیاری از عملیات از جمله تقطیع تصویر با سهولت بیشتری صورت پذیرد.

یکی از مباحث مهم و اساسی در بحث تقطیع تصویر، آشکارسازی خط می باشد. در خصوص آشکار سازی خط در تصاویر به تفضیل در بخش های آینده صحبت خواهیم نمود. اما نکته ای باید قبل از ورود به بحث آشکار سازی خط در نظر گرفت مرحله ای است که در صورت پیاده صحیح آن، فرآیند تشخیص خط را به آسانی و کمترین خطا به سرانجام خواهد رساند.

در سیستم های OCR یکی از ویژگیهای مهم و تأثیرگذار در شناسایی حروف و کلمات، خطوطی است که جملات بر پایه آن تعریف می شوند. این خطوط که به خطوط کرسی نیز معروفند گاه به صورت آشکار و گاه به صورت پنهان (در اسناد چاپی) قابل استخراج می باشد و براساس آن می توان الگوریتم های قطعه بندی شناسایی را تعریف کرد.

در زبان موسیقی این خطوط به خطوط حامل معروفند و در همه حال بایستی در یک صفحه از قطعه موسیقی (چه دستنویس و چه چاپی) وجود داشته باشند زیرا همانگونه که شرح کامل آن و کاربردهای آن در بخش ضمیمه آورده شده است، خطوط حامل از ارکان اصلی در نحوه شناسایی نت های موسیقی می باشند.

از آنجا که موقعیت خطوط حامل و نیز خطوط فرضی حامل نقش تعیین کننده ای در نحوه شناسایی نت ها و علائم موسیقی دارد، هر عاملی که باعث افت کیفیت یا بروز اعوجاج در این خطوط گردند، می تواند کل فرآیند شناسایی و اجرا را تحت الشعاع قرار داده و به نتایج نامطلوبی منجر شود.

در بسیاری از سیستم های OCR و نیز شناسایی متون از جمله زبان موسیقی ما علاقمند به تبدیل تصویر به یک تصویر باینری می باشیم و نیز علاوه بر آن در برخی کاربردها نیاز به برخی عملیات از جمله عملیات ریخت شناسی نیز است.

در چنین حالاتی در صورتی که زاویه تصویر ورودی به هر علتی دارای زاویه ناصحیح باشد کل سیستم را با مشکل مواجه نموده و باعث اختلال در نتیجه نهایی می گردد.

دو راه حل برای این موضوع می توان تصور کرد:

اعمال عملیاتی در حین پردازش اصلی جهت تصحیح زاویه (پردازشهای محلی)

ابتدا قبل از هر پردازشی، زاویه تصویر ورودی را اصلاح و سپس به عملیات اصلی بپردازیم.

در مورد راه حل اول, با توجه بار محاسباتی بالا در حین پردازش اصلی, پیشنهاد نمی گردد، مگر در مواردی که نیاز به اعمال پردازشهای محلی در تصویر باشد.

اما در خصوص راه حل دوم، در صورتی که سیستم شناسایی کننده نسبت به زاویه تصویر ورودی حساس باشد، یک بار در ابتدای کار و در مرحله پیش پردازش اقدام به تصحیح زاویه ای نموده و هزینه محاسباتی آن را به یک باره پرداخت می نمائیم.

در بسیاری موارد در زمان دریافت تصاویر ورودی با کمی خطای زاویه ای نسبت به محور x مواجه می شویم. که حتی چشم انسان نیز از درک آن عاجز است. لذا می توان با پیاده سازی الگوریتم ابتدای زاویه تصویر را با حساسیت بسیار بالا تصحیح نمود و سپس به پردازش اصلی پرداخت.

۳-۲-۱ تبدیل رادون[1]

در این بخش به یکی از روشهایی که توانایی بالایی در تصحیح زاویه تصویر ورودی براساس اشیاء موجود در تصویر ورودی دارد می پردازیم این روش به روش تبدیل رادون معروف بوده که تئوری اولیه آن در سال ۱۹۷۷ ارائه گردیده است.

[1] - Radon Trasnform.

کاربرد روش مذکور در تصحیح تصاویر، در صورتی که تصویر ورودی و اشیاء داخل آن دارای چگالی کوچکتر از ۷۰٪ نسبت به پس زمینه باشد بسیار مطلوب بوده و در غیر اینصورت برای تصاویر با چگالی بالاتر از ۷۰٪ با مشکلاتی مواجه می شود. تصاویری مانند مناظر با تراکم اشیاء زیاد و بدون نظم مانند تصویر جنگل که معمولاً زوایا در همه تصاویر پراکنده بوده و درهم تنیدگی اجسام در تصویر بسیار زیاد باشد.

۳-۲-۱-۱- تاریخچه تئوری تبدیل رادون

تئوری تومرگرافی محاسبه شده (CT) توسط جان رادون[1][۱]، ریاضیدان اهل وین در سال ۱۹۱۷ ارائه گردید که روشی برای یک شی دو بعدی دو پرتو نور موازی ارائه نمود. به این روش تبدیل رادون گفته می شود. ۴۵ سال بعد آلن ام کورمارک[2][۱] فیزیکدان دانشگاه توفتنه، بخشی از این مفاهیم را در CT به کار برد. کورماک یافته های اولیه خود را در سالهای ۱۹۶۳ـ ۱۹۶۴ به چاپ رساند و نشان داد که چگونه می توان از آنها برای بازسازی برش عرض بدن از تصاویر اشعه x که درجات متفاوت گرفته می شود استفاده کرد. یک مهندس برق به نام گادوی ان هانسفیلد[3] [۱] نیز به کمک همکارانش و به کمک این روش اولین دستگاه CT پزشکی را ساختند.

نسل اول روبشگرهای CT یک مدار اشعه x و یک آشکار ساز دارند. به ازای زاویه خاص چرخش، زوج منبع و آشکار ساز حول جهت خطی نمایش داده شده, می چرخند. افکنش R مانند افکنش P با اندازه گیری خروجی آشکار ساز در هر مرحله انتقال, تولید می شود. پس از یک انتقال خطی کامل مجموعه منبع و آشکارساز چرخانده شده و این فرآیند به منظور تولید یک افکنش دیگر در زاویه دیگر تکرار می شود. این

---

[1] - John Radon.
[2] - Allan M. Cormack.
[3] - Godfery N. Hounsfield.

روند برای تمام زوایای مطلوب در بازه $۰^\circ,۱۸۰^\circ$ به منظور تولید مجموعه ای کامل از افکنش ها تکرار می شود تا از آنها یک تصویر توسط پس افکنش تولید گردد. با کنار هم قرار دادن این برشها یک حجم سه بعدی از بخشی از بدن بدست می آید. از زمان تولید اولین نسل از اینگونه روبشگرها، تا کنون ۷ نسل ایجاد شده که در نسل هفتم از اینگونه روبشگرها تصاویری با جزئیات بیشتر تولید می شود.

۳-۲-۱-۲- اصول تبدیل رادون و روابط حاکم بر آن

یک خط مستقیم در دستگاه مختصات را می توان با دانستن شیب، به شکل $y=ax+b$ یا با نمایش نرمال نمایش داد.

$$a\cos\theta + y\sin\theta = P$$

رابطه ۳-۱

افکنش پرتوی از اشعه های موازی را می توان با استفاده از چندین خطوط مدل سازی کرد. یک نقطه دلخواه در سیگنال تأکیده شده را می توان از جمع شعاع ها در امتداد خط $x\cos\theta_k + y\sin\theta_k = P_j$ یافت.

با پیوسته در نظر گرفتن کمیتها، یک انتگرال خطی به صورت زیر است:

$$g\left(P_j,\theta_k\right) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f\left(x,y\right)\delta\left(x\cos\theta_k + y\sin\theta_k - P_j\right)\partial x\partial j$$

رابطه ۳-۲

که در آن از خاصیت ضربه استفاده شده است. به عبارت دیگر، سمت راست معادله ( ) صفر است مگر اینکه آرگومان صفر باشد. بنابراین انتگرال تنها در امتداد خط $x\cos\theta_k + y\sin\theta_k = P_j$ محاسبه می شود. اگر همه مقادیر θ, P را در نظر بگیریم، معادله فوق در حالت کل تبدیل به معادله زیر می شود.

$$P(P,\theta)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta\left(x\cos\theta_k + y\sin\theta_k - P_j\right)\partial x\partial j$$

رابطه ۳-۳

تبدیل رادون، براساس بازسازی از پرتوهای تابیده شده در حالت گسسته، معادله ۳-۳ به صورت زیر است:

$$P(P,\theta)=\sum_{a=0}^{m-1}\sum_{y=0}^{N-1} f(x,y)\delta\left(x\cos\theta + y\sin\theta - P\right)$$

رابطه ۳-۴

که در آن θ,P,y,x متغیرهای گسسته می باشند. اگر θ ثابت بوده و P تغییر کند، خواهیم دید که معادله ۳-۴ در حقیقت جمع پیکسلهای $f(x,y)$ در امتداد خطی هستند که با مقادیر مشخص این دو پارامتر مشخص می شوند. پوشش دادن تمام مقادیر ممکن در تصویر، یک افکنش را به ما می دهد، تغییر پی در پی زاویه θ و تکرار فرآیند فوق منجر به ایجاد افکنشهای دیگر می شود.

۳-۲-۱-۳- استفاده از تبدیل رادون جهت تصحیح کجی تصویر

تصاویر دیجیتال معمولاً به صورت ریاضیاتی و به کمک تابعی در حوزه مکان[1] و به صورت $f(x,y)$ نمایش داده می شوند. $f(x,y)$ بیانگر مقدار تابع در نقطه x,y است.

عملیات انتقال بیانگر نوعی نمایش ریاضیاتی تصویر در حوزه ای غیر از حوزه مکانی می باشد. تبدیل رادون در تصاویر دو بعدی یک نگاشت[2] از فضای کارتزین به یک فضای قطبی است.

تبدیل رادون روی یک تصویر، از مجموع تبدیل رادون روی هر پیکسل از تصویر بدست می آید. تابع رادون در راستای زاویه ای خاص ماتریس تبدیل رادون را ایجاد می کند. ماتریس تبدیل رادون تابع دو

[1] - Spatial domain.
[2] - Map.

شعبه ی $f(x,y)$ مجموعه است از انتگرال های خطی، که در راستای مختلف، زوایای مختلف از تصویر دو بعدی دریافت می شود.

محاسبه تبدیل رادون در زوایای مختلف با چرخش تصویر حول مرکز آن بدست می آید. در نهایت زاویه انحنای تصویر براساس ماکزیمم مقداری که در تبدیل رادون از زوایای مختلف بدست می آید، محاسبه می شود.

در عمل محاسبه تبدیل رادون برای تصویر ورودی با محاسبه انتگرالی خطی که با چرخش تصویر حول مرکز و در زوایای مختلف بدست می آید محاسبه می شود. زاویه های تصویر می تواند در بازه $[۰^\circ, ۱۷۹^\circ]$ درجه محاسبه گردد. رابطه ۳-۵ نحوه کار تبدیل رادون را نمایش می دهد.

$$R \oplus (x^{/}) = \int_{-\infty}^{+\infty} f\left(x^{/}\cos\theta - y^{/}\sin\theta, x^{/}\sin\theta + y^{/}\cos\theta\right) dy^{/}$$

رابطه ۳-۵

به طوری که:

$$\begin{bmatrix} x^{/} \\ y^{/} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \sin\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

رابطه ۳-۶

به طور کلی تبدیل رادون جهت تخمین زاویه تصاویر عملکرد مناسبی از خود نشان می دهد این کارآیی در مورد تصاویری که حاوی اطلاعات متنی باشند نیز به خوبی عمل می نمایند. آزمایشات مختلف دقت ۱۰۰٪ این روش در تصاویری که پس از عملیات باینری پس زمینه از پیش زمینه به خوبی جدا گردیده اند را نشان می دهند.

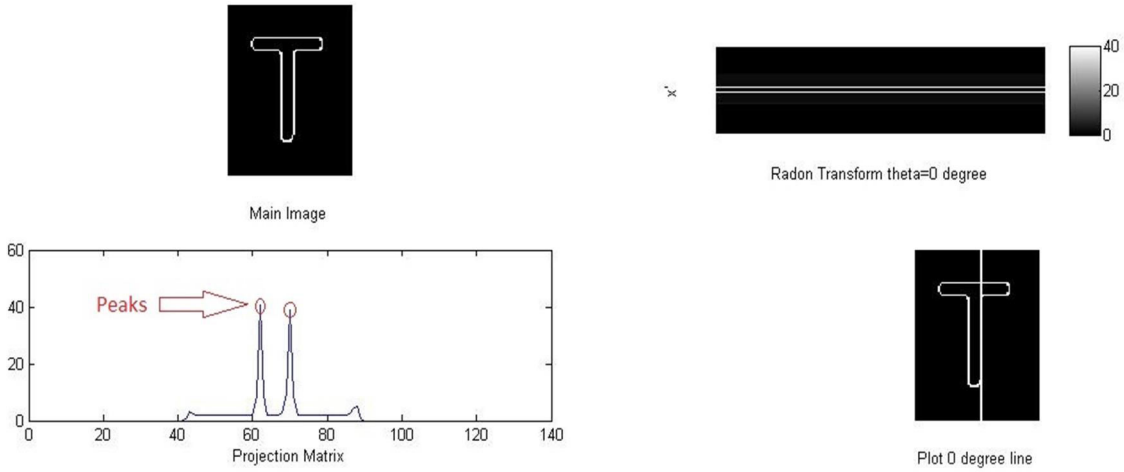
۳-۳- به کار گیری تبدیل رادون در شناسایی خط

شناسایی خطوط در بسیاری کاربردها پایه و اساس سیستم شناسایی کننده می باشد و در واقع نه شناسایی مرز بلکه خطوط موجود در تصویر مورد بررسی قرار می گیرد. به عنوان مثال در سیستم شناسایی کننده پلاک خودروها یکی از اساسی ترین چالشها , یافتن خط با طول مشخص نسبت به کل شئی شناسایی شده (در اینجا خودرو) می باشد.

در سیستم های OCR و OMR نیز شناسایی خط از ارکان اصلی در سیستم شناسایی کننده محسوب می شود . در فصل ۲ به برخی روش های شناسایی خط اشاره شد. یکی از این روش ها استفاده از روش لبه یابی Canny است . اما باید در نظر داشت که روش Canny در مواردی که خط در تصویر دارای شکستگی باشد به خوبی عمل نکرده و هر قطعه خط را به صورت جداگانه شناسایی می نماید. از دیگر مشکلات روش Canny عملکرد نا مناسب آن در تشخیص خطوط در زوایای مختلف است. گاهی در برخی کاربردها نیاز به شناسایی خطوط با زوایای مختلف می باشد. در صورتی که عموم روشها از ماسک گذاری استفاده کرده و در این خصوص عملکرد ضعیفی از خود نشان می دهند.

یکی ار برجسته ترین روش ها در یافتن خط, استفاده از تبدیل رادون می باشد که تاریخچه و شرح عملکرد آن در ابتدای فصل آورده شد. همانگونه که قبلا به آن اشاره شد, تبدیل رادون در راستای زاویه خاص , ماتریس تبدیل رادون را ایجاد می نماید . در واقع این مانریس حاصل محاسبه انتگرال خطی در زوایای مختلف می باشد. با این توضیح حال فرض کنید ماتریس تبدیل رادون را فقط برای زاویه خاص ایجاد کنیم. به عنوان مثال زاویه ۰ و یا ۹۰ درجه . با توجه به رابطه ۳-۴ تبدیل رادون با ورودی $\theta=90^\circ$ و یا هر زاویه دیگری به سیستم اعمال می گردد و نتیجه ایجاد ماتریسی است که بیشترین مقدار موجود در آن وجود خط در آن زاویه را نشان می دهد.

در شکل زیر یک نمونه تصویر که حاوی حرف T نمایش داده شده است . پس از اعمال تبدیل رادون , ماتریس تبدیل رادون ایجاد شده به صورت طیف رنگ های مختلف نمایش داده شده است. همچنین ماتریس R به صورت یک سیگنال نیز برای درک بهتر آورده شده است . همانگونه مشاهده می شود ۲ نقطه دارای بیشترین مقدار می باشد که در واقع نشانگر ۲ خط عمودی با زاویه قائم در تصویر است . در نهایت یکی از این دو مقدار در تصویر علامت گذاری شده است.



شکل ۳-۱: تبدیل رادون جهت شناسایی خطوط

۳-۴- قطعه بندی تصویر

قطعه بندی فرآیندی است که در آن اجزای تشکیل دهنده تصویر از هم تفکیک شده و اجزائی که ویژگی های مشترکی دارند برچسب یکسان به آنها داده می شود. گاه این برچسب گذاری جهت جداسازی پس زمینه از پیش زمینه بوده و گاه بمنظور جداسازی چند جزء از تصویر که دارای مشخصات متمایز از یکدیگر می باشند بکار گرفته می شوند. در واقع آستانه ای برای تفکیک خصوصیات ۲ یا چند جزء در تصویر یافت شده و براساس مقدار آستانه آن را از هم جدا می سازد.

۳-۴-۱- آستانه گذاری به روش Otsu:

روش Otsu براساس بیشینه کردن واریانس بین دو قطعه از تصویر، عمل قطعه بندی را انجام می دهد. به عبارت دیگر این روش براساس پیدا کردن دو ناحیه با واریانس بالا نسبت یکدیگر عمل نموده و مقدار آستانه را بگونه ای انتخاب می کند که بیشترین واریانس در آن محقق گردد. در بسیاری کاربردها به بیش از ۱ مقدار آستانه نیاز است زیرا لازم است تصویر به بیش از ۲ قطعه تقسیم شود. و یا با توجه به طیف نور تابیده شده به تصویر امکان قطعه بندی سراسری با در نظر گرفتن یک سطح آستانه میسر نمی باشد. در این مواقع می توان از آستانه گذاری متغیر استفاده نمود.

جهت استفاده از این روش, تصویر ورودی به بلوکهایی تقسیم می شوند و عمل آستانه گذاری روی هر بلوک که نسبت به بلوک های دیگر تصویر غیر هم پوشان می باشند، صورت می گیرد. عملکرد این روش در مقایسه با آستانه گذاری سراسری بسیار بهتر بوده و نتایج بهتری از آن دریافت می شود.

به طور کلی روش Otsu در مسائلی که ورودی دو ناحیه قابل تمایز باشند به خوبی عمل کرده و نتایج مطلوبی را حاصل می نمایند و رابطه آن به صورت زیر می باشد:

$$\sigma^2 = P_1(M_1 - M_I)^2 + P_2(M_2 - M_I)^2$$ رابطه ۳-۷

در رابطه بالا M_1 میانگین مقادیر سطح خاکستری قطعه ۱، M_2 میانگین مقادیر سطح خاکستری قطعه ۲، M_I میانگین کل مقادیر خاکستری تصویر، P_1، احتمال کلاس ۱، P_2 احتمال کلاس ۲ و σ^2 واریانس بین دو قطعه می باشند. پارامترهای موجود در رابطه () برای مقدار آستانه T به صورت روابط زیر تعریف می گردند.

$$P_1 = \sum_{i=1}^{T} P_i$$

رابطه ۳ـ ۸

$$P_2 = \sum_{i=T+1}^{L-1} P_i = 1 - P_1$$

رابطه ۳ـ۹

$$M_1 = \frac{1}{P_1} \sum_{i=0}^{T} iP_i$$

رابطه ۳ـ ۱۰

$$M_2 = \frac{1}{P_2} \sum_{i=T+1}^{L-1} iP_i$$

رابطه ۳ـ ۱۱

$$M_I = \sum_{i=0}^{L-1} iP_i = P_1 M_2 + P_2 M_2$$

رابطه ۳ـ ۱۲

و در نهایت رابطه واریانس بین دو ناحیه را به صورت رابطه ۳ـ ۷ می توان بازنویسی کرد:

$$\sigma^2 = P_1 P_2 (M_1 - M_2)^2 = \frac{(M_I P_1 - M)^2}{P_1(1 - P_1)}$$

رابطه ۳ـ ۱۳

که M میانگین تحمیلی تا آستانه T است که صورت رابطه ۳ـ ۸ می توان آن را محاسبه کرد:

$$M = \sum_{i=0}^{T} iP_i$$

رابطه ۳ـ ۱۴

۳-۴-۲- قطعه بندی المانهای تصویر

همانگونه که بیان شد, قطعه بندی مرحله ای بسیار مهم است که در آن اجزاء تصویر که باید به مرحله شناسایی کننده ارسال شود. هیچ روش عمومی برای قطعه بندی المان های نت های موسیقی موجود نیست و روش و تکنیکهای مختلفی برای آن ایجاد شده است.

قطعه بندی در قطعات موسیقی در مرحله اول بر مبنای خطوط حامل شناسایی شده و در مرحله بعد بر مبنای فواصل بین قطعات موسیقی صورت می پذیرد. این مرحله علی رغم قطعه بندی ولی دارای نواقصی

است. درست مانند اغلب سیستم های OCR که فرآیند قطعه بندی از شناسایی خط، شناسایی کلمات و شناسایی کاراکترها تشکیل شده است؛ در قطعات موسیقی نیز همین روال وجود دارد با این تفاوت که شناسایی خطوط حامل به دلیل نقش مهم و کاملاً کلیدی که در قطعه موسیقی و در نهایت نتیجه سیستم های OMR دارد, بایستی به درستی و کاملاً دقیق محاسبه گردند. همچنین زبان موسیقی نیز مانند زبان محاوره (فارسی و عربی) دارای اتصالاتی بین المانهای آن می باشند که بایستی جهت شناسایی, هر قطعه کلی، به قطعات تشکیل دهنده تجزیه شوند. لذا ابتدا می بایست قطعات کلی شناسایی شده و سپس هر یک در صورت قابل تجزیه بودن به قطعات و المانهای تشکیل دهنده تجزیه گردند.

از نظر فواصل بین المان ها، فاصله ها از قانون خاصی پیروی نمی کنند و گاه در قطعه ای فاصله ها بسیار کم و گاه در قطعات ساده تر فاصله های زیادی بین قطعات وجود دارد. لذا الگوریتم جدا کننده بایستی به درستی و با دقت نسبت به جداسازی آنها عمل نماید.

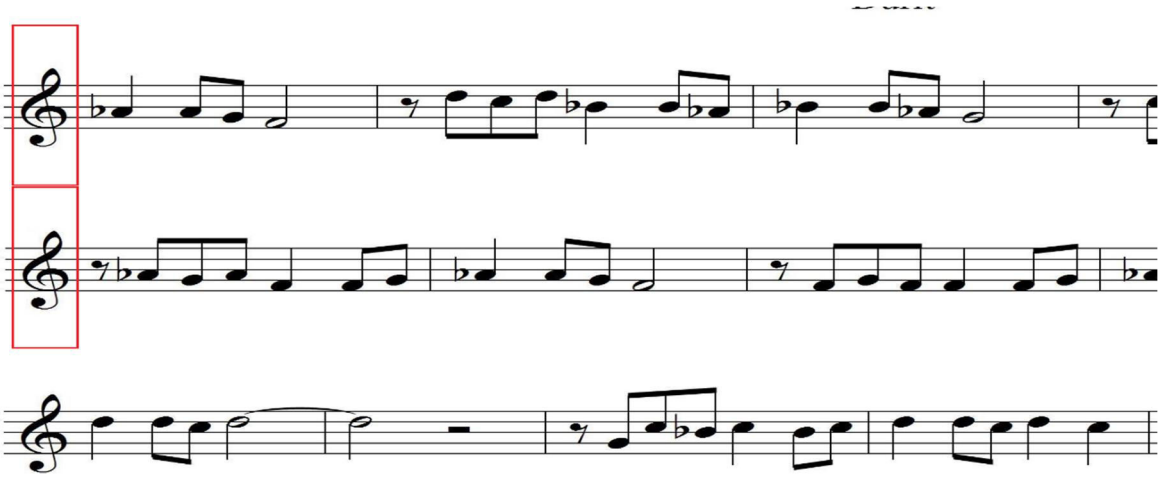
از طرفی وجود برخی علائم در زبان موسیقی مانند خطوط اتحاد و اتصال که برای نشان دادن اشتراکات زمان بین دو نت و یا گروهی از نت ها استفاده می گردند در بسیاری موارد دارای چسبندگی با نت بوده که این موضوع نیز یکی از چالشهای پیش رو در قطعه بندی می باشد.

همچنین وجود خط اتحاد و اتصال لزوماً در پایین و یا بالای گروهی از نت ها قطعه بندی را با پیچیدگی هایی مواجه می سازد بدلیل اینکه سیستم قطعه بندی علاوه بر قطعه بندی در راستای عمودی، بایستی در راستای افقی نیز اقدام به قطعه بندی نماید.

شکل ۳-۲ نمونه خط اتحاد چسبیده به نت

۳-۴-۳- قطعه بندی کل تصویر

براساس داده های بدست آمده از محل دقیق خطوط حامل در تصویر و فاصله هر دسته از خطوط حامل از یکدیگر فاصله هر گروه از خطوط حامل محاسبه می گردد، فاصله مذکور تقسیم به ۲ شده و هر بخش به گروهی از خطوط تعلق پیدا می کند. این مرز بندی از هر دو جهت بالا و پایین خطوط حامل رعایت شده و برای تعیین مرز خطوط حامل اول و آخر از میانگین فاصله هایی که برای خطوط حامل میانی بدست آمده استفاده می شود به شرط آنکه مرز از ابعاد تصویر تجاوز ننماید.



شکل ۳-۳ قطعه بندی کلی قطعات موسیقی

پس برای هر دسته از خطوط حامل، اقدام به حرکت روی خطوط مرز از قبل تعریف شده می نمائیم. در هر نقطه که مجموع نقاط عمودی مرز، بزرگتر از صفر بود، به عنوان نقطه آغاز قطعه کلی ذخیره شده و تا زمانی که مجموع پیکسلهای ستون بعدی بزرگتر از صفر بود، اقدام به حرکت رو به جلو می نمائیم تا جایی که مجموع مقادیر پیکسلهای ستونی صفر گردد. پس از آن اقدام به ذخیره قطعه نموده و مراحل تا آخرین ستون تصویر تکرار می شود. بدین ترتیب در هر گروه از خطوط حامل قطعات شناسایی می گردند.

۳-۴-۴- قطعه بندی المانها از قطعات کلی با استفاده از معادلات تفاضلی

همانگونه که قبلاً نیز اشاره شد، فرآیند قطعه بندی المانهای موسیقی نیز به طور کلی مشابه فرآیند قطعه بندی کلمات (در زبان فارسی و عربی) می باشد و از سه بخش تشکیل شده است.

ـ شناسایی خط

ـ شناسایی کلمات

ـ شناسایی کاراکتر

در زبان موسیقی نیز همین روال معتبر است به این صورت که ابتدا خطوط حامل شناسایی می گردد سپس قطعات کلی و در نهایت المانهای تشکیل دهنده که هر کدام موجودیتی مستقل و دارای معنی می باشند از قطعات کلی شناسایی می گردد.

۳-۴-۴-۱ بکارگیری مشتق مرتبه اول در جداسازی المانهای موسیقی از قطعات کلی

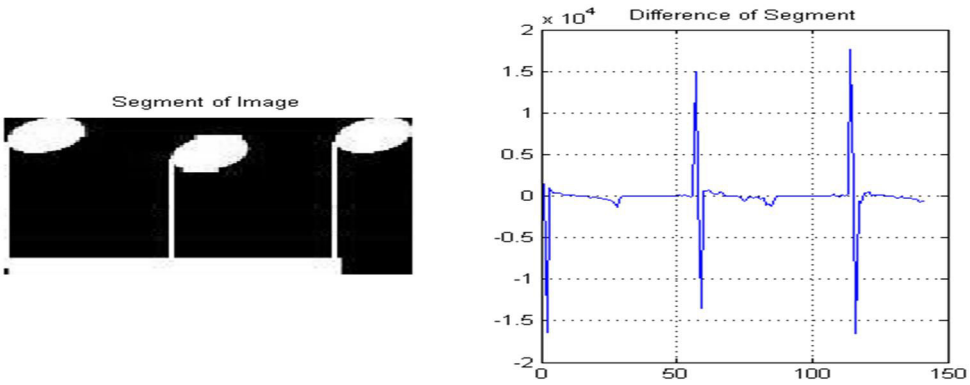
به طور کلی مشتقات یک تابع دیجیتال برحسب تفاضل بیان می شوند و دارای خواصی است که به کمک این خواص می توان اقدام به جداسازی و تفکیک المانها از قطعات کلی نمود.

تعریف پایه مشتق مرتبه اول یک تابع یک بعدی $f(x)$ به صورت تفاضل زیر است:

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

رابطه ۳ـ ۱۵

در یک سیگنال که شامل پروفایل یک قطعه است با اعمال مشتق مرتبه اول روی آن در نقاطی که شدت روشنایی ثابت است مقدار مشتق صفر خواهد بود و در مکانهایی که شدت روشنایی تغییر داشته باشد مقدار مشتق غیر صفر (یا حول عدد صفر) خواهد بود.



شکل ۳-۴ نمونه قطعه موسیقی و مشتق مرتبه اول آن

از این خاصیت می توان برای یافتن نقاط انفصال در قطعات کلی استفاده نمود. به این صورت که ابتدا مجموع ستونهای قطعه کلی محاسبه و به صورت یک سیگنال تک بعدی ذخیره می گردد، سپس با اعمال مشتق مرتبه اول نقاطی که مشتق محاسبه شد، در آن نقاط صفر می باشد به عنوان نقطه انقطاع انتخاب می شود. با توجه به اینکه این سیگنال صرفاً قطعه را از دید عمودی بررسی می کند در انتها هر بخش که به وسیله مشتق گیری تفکیک گردیده را جهت کنترل اینکه آیا از نظر افقی نیز قابل تفکیک می باشد به صورت افقی نیز مورد بررسی قرار می دهیم.

۳-۴-۵ شناسایی و رفع مشکل چسبندگی برخی المانها:

در بخش قبلی اشاره شده که قطعات کلی علاوه بر اعمال مشتق روی مجموع داده هایی که از پویش عمودی بدست می آید، مشتق روی داده هایی که از پویش افقی بدست می آید نیز اعمال می شود.
بدین ترتیب در صورتی که هر زیر قطعه دارای دو المان مجزا باشد یافت می گردد. در اغلب نمونه ها این مشکل مربوط به چسبندگی انتهای خطوط اتحاد و یا اتصال با نت ها می باشد. این چسبندگی علاوه بر اینکه باعث اخلال در روند جداسازی می گردد، و سیستم را برای شناسایی تک تک المانهای موسیقی دچار مشکل می سازد، موجب اختلال در شناسایی خود خطوط اتصال یا اتحاد نیز می گردد، و از آنجا که خطوط اتحاد نقش بسیار کلیدی در اجزای یک قطعه موسیقی دارد، در نتیجه تأثیر نامطلوب خواهد گذاشت.

۳-۴-۵-۱ شناسایی چسبندگی

در زمان قطعه بندی المانهای ممکن است به مشکلی برخورد کنیم که پس از جداسازی در آزمونی که برای آن در نظر گرفته شده است مردود اعلام شود و دلیل آن این است که یک قطعه جدید خود دارای دو یا بیش از دو ناحیه باشند. در این صورت نتیجه می گیریم این قطعه دارای ارتباطی با نت کناری و یا چند نت کناری است. همانگونه که گفته شد در اغلب نمونه ها این مشکل در زمانی رخ می دهد که خطوط اتحاد و یا اتصال با فاصله بسیار نزدیکی از سر نت ها قرار گرفته باشند و در فرآیند قطعه بندی کلی تصویر و یا فرآیند باینری کردن این اتصال همچنان برقرار مانده است.

۳-۴-۵-۲ روش پیشنهادی جهت رفع مشکل

نمایش عمومی نت ها همچنان که در بخش ضمیمه و در بحث معرفی و آشنایی با مبانی موسیقی آورده شده است بسیار مشابه هم بود و اساس خطوط اتحاد و یا اتصال نیز, ارتباط دادن دو یا بیش از دو نت با یکیدگر می باشد. از این رو می توان با مقایسه قطعات با یکدیگر و حذف تدریجی سلولهای ماتریس و آزمون شباهت با المان حذف شده به نقطه مطلوب جهت جداسازی رسید.

در اینجا به معرفی و بررسی معیار شباهت که توسط جناب آقای دکتر حمید حسن پور و همکارشان ارائه شده می پردازیم و نحوه دقیق پیاده سازی را به فصل ارجاع می دهیم.

معیارهای شباهت بسیاری جهت اندازه گیری میزان دو پدیده ابداع شده است. مانند معیارهای شباهت کوساین (Cosine)، کورلیشن (Correlation) و یا معیار جنسن (Jensen) که هر یک دارای نقاط قوت و ضعف می باشند. اما در این پایان نامه روشی معرفی گردیده که تا حدودی بسیاری نقاط ضعف روشهای موجود را پوشش می دهد. روش مذکور به کمک تلفیق دو معیار شباهت Cosine، Correlation بدست آمده است.

از جمله خصوصیات آن می توان مقاوم بودن در مقابل شیفت در زمان، شیفت در دامنه، مقیاس دامنه و تأثیر در فاز می باشد. الگوریتم ارائه شده به صورت زیر می باشد:

$$\mu_y = \frac{1}{m}\sum_{i=1}^{m}\gamma_i$$

رابطه ۳ـ۱۶

$$X_k = Circshift\left(x,\left(-k+1\right)\right), \mu_{x_k} = \frac{1}{m}\sum_{i=1}^{m}X_{k_i}$$

رابطه ۳-۱۷

$$k = 1,2,\ldots,n$$

$$s_k = \frac{\sum_{i=1} \left(x_{k_i} - \mu_{x_k}\right)\left(\gamma_1 - \mu_y\right)}{\sqrt{\sum_{i=1}^{m} \left(x_{x_i} - \mu_{x_k}\right)^2 \sum_{i=1}^{m} \left(\gamma_i - \mu_y\right)^2}}$$

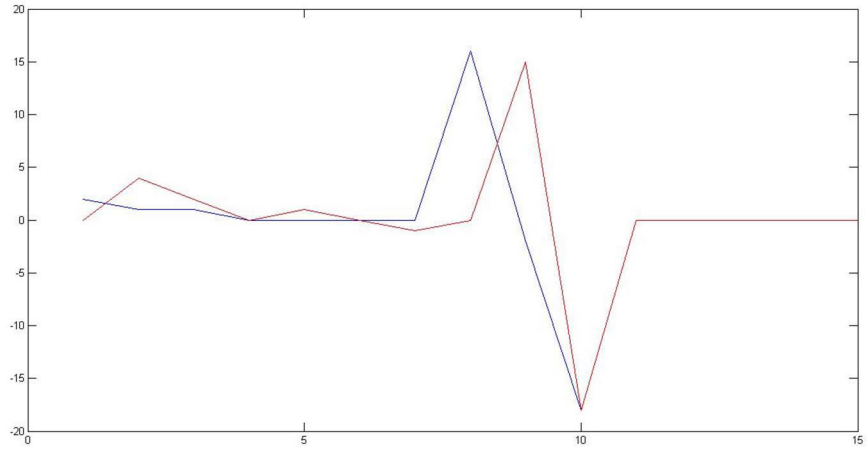
رابطه ۳ـ۱۷

$$similority(x, y) = \underset{1 \le i \le n}{Max}\left[s(i)\right]$$

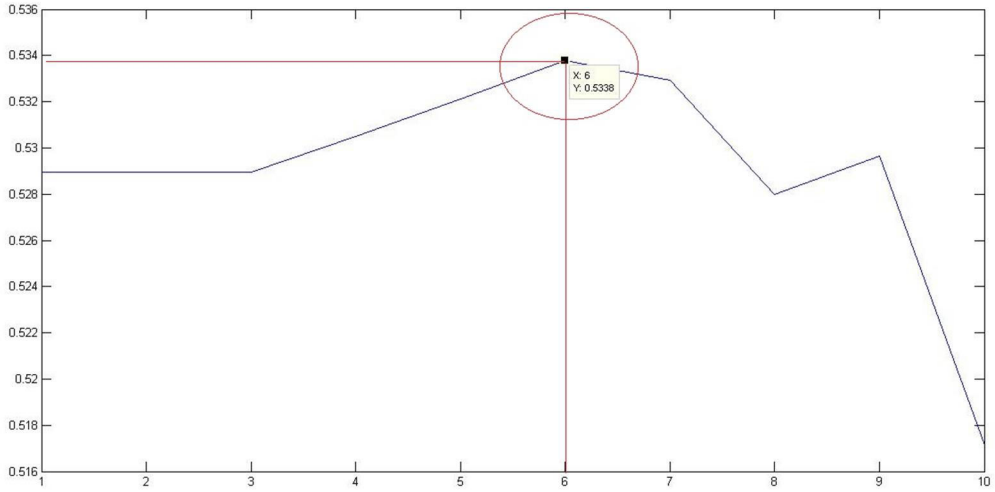
رابطه ۳ـ۱۸

با توجه به مشکل چسبندگی و راه حلی که مطرح شد و معیار شباهت پیشنهادی، جهت پیدا کردن بهترین نقطه جهت جداسازی، ابتدا سیگنال هر دو قطعه را به صورت مشتق مجموع روی ستونهای آن بدست می آوریم و به عنوان ۲ ورودی به الگوریتم اندازه گیری میزان شباهت می دهیم و در هر قدم خانه های آخر قطعه متصله که قصد پیدا کردن مناسب ترین نقطه جداسازی در آن را داریم را صفر می نمائیم و بزرگترین مقدار بدست آمده از الگوریتم شباهت را در برداری ذخیره می کنیم در نهایت با بررسی بردار بدست آمده نقطه ای که بزرگترین شباهت در آن اتفاق افتاده است را به عنوان نقطه مطلوب برمی گزینیم.

شکل ۳-۵ نمونه المان قطعه بندی شده و خط اتحاد چسبیده به آن



نمودار ۳-۱ دو نمونه پروفایل از دو نت هم شکل جهت یافتن نقطه انفصال



نمودار ۳-۲ نمودار مقادیر بدست آمده از مراحل معیار شباهت

شکل ۳-۶ نقطه انفصال مطابق مقدار ماگزیمم یافت شده در نمودار ۳-۲

۳-۵-شبکه های عصبی مصنوعی

سرعت و قدرت پردازش بسیار بالای مغز انسان، به ارتباط های بسیار انبوهی بر می گردد که در میان سلول های سازنده مغز وجود دارد و اساسا، بدون وجود این لینک های ارتباطی، مغز انسان هم به یک سیستم معمول کاهش می یابد، که قطعا توانایی های فعلی را نخواهد داشت.

گذشته از همه این ها، عملکرد عالی مغز در حل انواع مسائل و کارایی بالای آن، شبیه سازی مغز و قابلیت های آن را به مهم ترین آرمان معماران سخت افزار و نرم افزار تبدیل کرده است. در واقع، اگر روزی فرا برسد که ما بتوانیم کامپیوتری در حد و اندازه های مغز انسان را بسازیم، قطعا یک انقلاب بزرگ در علم، صنعت و االبته زندگی انسان ها، رخ خواهد داد.

در راستای شبیه سازی رفتار محاسباتی مغز انسان، از چند دهه گذشته، که کامپیوترها امکان پیاده سازی الگوریتم های محاسباتی را فراهم نمودند، کارهای پژوهشی توسط متخصصین علوم کامپیوتر، مهندسین و ریاضی دان ها شروع شده است، که ما حصل کار آن ها، در شاخه ای از علم هوش مصنوعی، و در زیر شاخه هوش محاسباتی، تحت عنوان موضوع «شبکه های عصبی مصنوعی طبقه بندی شده است. در مبحث شبکه های عصبی مصنوعی، مدل های ریاضی و نرم افزاری متعددی با الهام گرفتن از مغز انسان

پیشنهاد شده اند، که برای حل گستره وسیعی از مسائل علمی، مهندسی و کاربردی، در حوزه های مختلف کاربرد دارند.

۳-۵-۱ شبکه عصبی هاپفیلد[1]

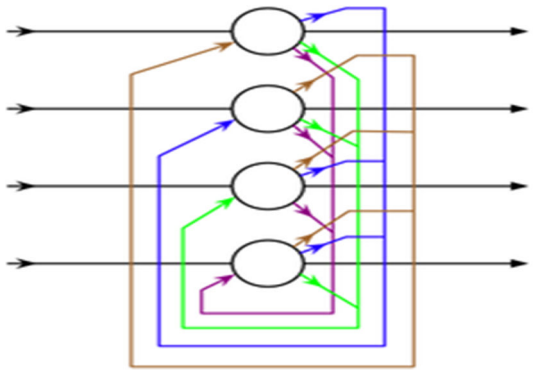
این نوع شبکه عصبی، بیشتر دارای ماهیتی شبیه به یک سیستم دینامیکی است، که دو یا چند نقطه تعادل پایدار دارد. این سیستم با شروع از یک شرایط اولیه، نهایتا به یکی از نقاط تعادلش همگرا می شود. همگرایی به هر نقطه تعادل، به عنوان تشخیصی است که شبکه عصبی آن را ایجاد کرده است و در واقع می تواند به عنوان یک رویکرد برای حل مسائل طبقه بندی استفاده شود.

این سیستم، یکی از قدیمی ترین انواع شبکه های عصبی است، که دارای ساختار بازگشتی است و در ساختار آن، فیدبک های داخلی وجود دارند.

ارتباطات بین نورونها با وزن مثبت بین دو نرون، نرون دیگر را تحریک می نماید و ارتباط نرون با وزن منفی بین دو نرون، نرون فعال دیگری را غیر فعال می نماید. شبکه هاپفیلد با انتخاب یک نرون تصادفی کار خود را شروع می کند. در صورتی که یکی یا بیشتر از نرونهای همسایه های آن نرون فعال باشند، مجموع وزندار نرونهای منتهی به آن نرون محاسبه می شود. سپس مجدداً یک نرون دیگر به تصادف انتخاب شده و همین عملیات آنقدر ادامه پیدا می کند تا شبکه به یک حالت پایدار برسد.

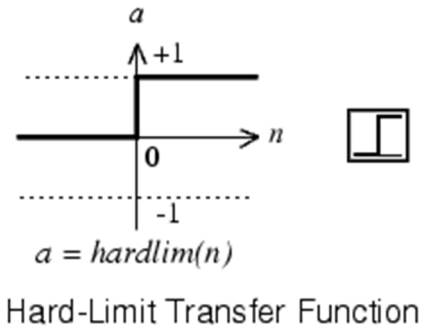
شبکه هاپفیلد شبکه ای پویا، بازگشتی و شامل چندین نرون با واحدهای تأثیری است که سامانه پس خور چند حلقه ای را تشکیل می دهند، یعنی خروجی تک تک نرون ها به سایر نرونها غیر از خودش پس خور دارد. خاصیت تأخیری به این معناست که سیستم در هر لحظه براساس وضعیت لحظه قبل عمل نموده و از همین رو همانگونه که به آن اشاره شد, شبکه های هاپفیلد یک سیستم دینامیکی است.

[1] - Hopfield Neural Network.



شکل ۳-۷ شمای کلی یک شبکه عصبی هاپفیلد

همانگونه که در شکل ۳-۷ مشاهده می شود در این نوع شبکه عصبی ساختار شبکه دارای یک لایه نورون می باشد و خروجی هر نورون به نورونهای دیگر متصل می گردد. تابع فعال ساز در این نوع شبکه از نوع Hardlim بوده که به صورت ذیر تعریف می گردد.



شکل ۳-۸ نمونه تابع ***Hard lim***

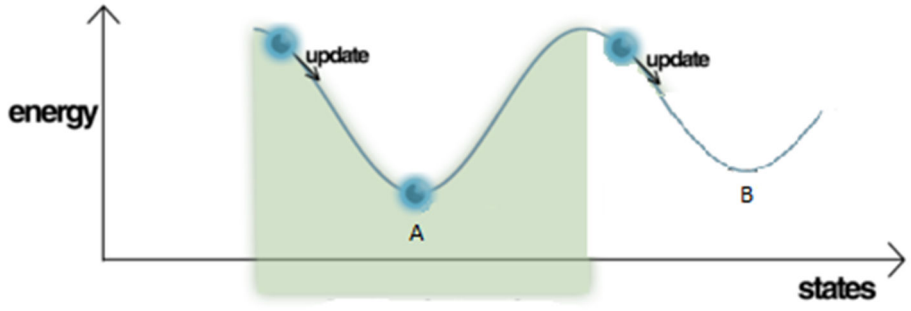
یک شبکه عصبی هاپفیلد در حال کلی شامل چندین معادله دیفرنس (معادلا تفاضلی) می گردد چرا که عملکرد آن همواره به صورت عمل بر اساس حالت گذشته سیستم می باشد.

در شبکه های عصبی هاپفیلد هر نرون که در شبکه وجود دارد مشخص کننده یک نقطه تعادل می باشد یا به عبارت دیگر تعداد نرون طراحی شده در شبکه عصبی درجه سیستم را نشان می دهد.

۳-۵-۱-۱ مفهوم انرژی در شبکه عصبی هاپفیلد

در طراحی شبکه عصبی هاپفیلد مفهومی به نام انرژی تعریف می گردد و سعی ما در این است, نقاط تعادل بگونه ای تعریف گردند که در آن کمترین انرژی را داشته باشیم.

به شکل زیر توجه کنید. اگر نقطه A را بعنوان نقطه تعادل یک سیستم در نظر بگیریم, در همان نقطه کمترین انرژی برای جسمی که که در قله قرار دارد و به هر دلیلی به سمت پایین سقوط می کند داریم. در مقابل نقطه B را در نظر بگیریم که به لحاظ میزان انرژی , نقطه تعادل آن دارای انرژی بیشتری نسبت به نقطه A می باشد.در واقع ناحیه جذب بین نقاط تعادل نیز میتواند متفاوت باشد.



شکل ۳-۹ دو نقطه تعادل نواحی جذب آنها

با توضحات ارائه شده در واقع درزمان طراحی یک شبکه عصبی هاپفیلد ما نقاط تعادل مختلف را برای سیستم تعریف می کنیم . به عنوان مثال خصوصیات کاراکتر های زبان فارسی یا اعداد و یا علائم و المان های موسیقی و زمانی که سیستم از یک حالتی که در اینجا ورودی جدید است, به این نقاط تعادل

نزدیک باشد در N مرحله به سمت نقاط تعادل همگرا می شود در این حالت می توان مفوم انرژی را در شبکه های عصبی هاپفیلد تعریف نود که به صورت زیر تعریف می شود.

$$E = -\frac{1}{2} \sum_i \sum_j W_{ij} \ X_i \ X_j - \sum_i b_i \ X_i$$ رابطه ۳-۱۹

با فرض در نظر گرفتن Wها و b ها به صورت بردا رابطه به صورت زیر نوشته می شود.

$$E = -\frac{1}{2} X^T \ WX - b^T X$$ رابطه ۳-۲۰

در واقع وظیفه شبکه هپفیلد همگرایی به نقاط مینیمم انرژی است که به ازای نقاط تعادل مطلوب در سیستم طراحی شده است. منظور از مینمم انرژی لزوما مینیمم مطلق نیست و در بسیاری سیتم ها همگرایی به نقاط مینیمم محلی نیز ما را به پاسخ صحیح راهنمایی می نماید.

۳-۶ جعبه ابزار MIDI

MIDI مخفف Musical Instrument Digital Interface است و در واقع رابطی است سخت افزاری ، نرم افزاری برای صحبت کردن نرم افزار های موسیقی و دستگاه ها ی موسیقی دیجیتال با یکدیگر, همانگونه که یک نوازنده موسیقی با نگاه کردن به نت موسیقی از روی علائم تشخیص میدهد کدام نت ها را با چه حسی باید اجرا کند ، لازم است زبان مشترکی وجود داشته باشد تا وسائل دیجیتال موسیقی نیز بتوانند با یکدیگر در زمینه موسیقی تبادل اطلاعات کنند.

از منظر نرم افزاری فایل MIDI فایلی است که شامل نت های موسیقی می گردد. حجم فایل های MIDI عموما بسیار پایین است و دلیل آن نیز به نحوه ذخیره سازی نت های موسیقی بر میگردد. نت های موسیقی در فایل های MIDI به صورت یک فایل نوشتاری ذخیره می گردد.

۳-۶-۱ خواندن , ایجاد و تغییردادن فایلهای MIDI

کتابخانه ای در نرم افزار متلب جهت خواندن و برخی عملیات روی فایل های MIDI ایجاد گردیده است[۱] که در ادامه به معرفی و نحوه کار برخی از آنها می پردازیم . با توضیحات مختصری که در ادامه آورده شده ، می توان عملیات مختلفی را روی فایل های MIDI انجام داد.

۳-۶-۱-۱ خواندن , ایجاد و تغییردادن فایلهای MIDI

بمنظور خواندن یک فایل از نوع MIDI از تابع ()Readmidi استفاده می شود . به کمک تابع مذکور ماتریسی از نت های موجود در فایل MIDI ایجاد می گردد:

Nmat= readmidi (FileName.mid);

ماتریس نت ها به ماتریسی اشاره دارد که حاوی اطلاعات نت های فایل مربوطه می باشد که در جدیل زیر بخشی از محتویات ماتریس nmat را مشاهده می کنیم:

```
nmat =
0         0.9000    1.0000   64.0000   82.0000    0         0.5516
1.0000    0.9000    1.0000   71.0000   89.0000    0.6122    0.5516
2.0000    0.4500    1.0000   71.0000   82.0000    1.2245    0.2755
2.5000    0.4500    1.0000   69.0000   70.0000    1.5306    0.2755
3.0000    0.4528    1.0000   67.0000   72.0000    1.8367    0.2772
3.5000    0.4528    1.0000   66.0000   72.0000    2.1429    0.2772
4.0000    0.9000    1.0000   64.0000   70.0000    2.4490    0.5516
5.0000    0.9000    1.0000   66.0000   79.0000    3.0612    0.5516
5.0000    0.9000    1.0000   67.0000   85.0000    3.6735    0.5516
7.0000    1.7500    1.0000   66.0000   72.0000    4.2857    1.0714
```

شکل ۳-۹ نمونه محتوای یک فایل ***MIDI***

۱- کتابخانه توسط دپارتمان موسیقی دانشگاه یوواسکیلا کشور فنلاند ایجاد و توسعه داده شده است.

همانگونه که مشاهده می شود ماتریس nmat شامل ۱۰ سطر و ۷ ستون می باشد که با اعدادی پر شده است . ستون ها به انواع گوناگون اطلاعات مانند تن صدا [1] و کانال MIDI اشاره دارد. ردیف ها به رویداد های مربوط به یک نت اشاره دارد .

نام گذاری مربوط به هر ستون به صورت زیر می باشد:

ONSET BEATS)	DURATION (BEATS)	MIDI channel	MIDI PITCH	VELOCITY	ONSET (SEC)	DURATION (SEC)

شکل ۳-۱۰ نام هر یک از ستون های فایل ***MIDI***

ONSET(BEATS): اولین ستون به شروع نت در واحد ضرب دلالت دارد.(بر اساس میزان تیک در یک چهارم نت)

DURATION(BEATS): دومین ستون به طول اجرای نت در واحد ضرب اشاره دارد.

MIDI channel: سوین ستون به کانال های MIDI اشاره دارد. (۱۶-۱)

MIDI PITCH: چهارمین ستون تن صدای MIDI را نشان میدهد.

VELOCITY: پنجمین ستون سرعت نواختن نت را نشان می دهد و یا به عبارت دیگر به چه مقدار صدا می بایستی اجرا شود.(۱۲۷-۱)

ستون های ششم و هفتم مشابه ستون اول و دوم می باشد با این تفاوت که بجای واحد ضرب (beat) از واحد ثانیه استفاده می کند.

[1] Pitch

اغلب نیاز است فقط به اطلاعات Pitch یا Duration دسترسی داشته باشیم . برای وضوح و سادگی این ستون ها به کمک صدا زدن برخی توابع انتخاب کننده فراخوانی می شوند تا فقط به یک ستون خاص دسترسی داشته باشیم . این توابع شامل onset، dur ، channel ، pitch وvelocity می باشند.

بعنوان مثال Pitch(nmat) فقط مقادیر نت ها را بر میگرداند و Onset(nmat) فقط مقادیر زمانی شروع (در مقیاس Beat) نت ها را در ماتریس نت ها باز می گرداند.

۳-۶-۱-۲ ایجاد یک رشته از نت ها

اگر بخواهیم از جعبه ابزار در اختیار برای ایجاد یک رشته از نت ها استفاده کنیم ، از تابع ()createnmat بهره می بریم.از این تابع برای ایجاد ماتریس نتها در فواصل منظم و مساوی استفاده می شود.شکل کلی این تابع به صورت زیر می باشد.

```
nmat = createnmat(notes,<dur>,<vel>,<ch>);
```

معرفی آرگومان های ورودی:

Notes: بردار نت ها

Dur : طول اجرای نت ها در واحد ثانیه (به طور پیش فرض 0.25)

Vel: سرعت نواختن نت ها (۱۲۷-۰)

Ch : کانالی که نت ها در آن قرار خواهند گرفت (به طور پیش فرض ۱)

با توجه به اینکه در بخش ضمیمه تئوری موسیقی نت ها را با نام های لاتین معرفی نمودیم در اینجا نام نت و عدد متناظر هر نت را در فایل های MIDI معرفی می کنیم:

سبک	نوع	نخست		دومی		سومین	چهارمین		پنجمین		ششمین		هفتمین
	Natural	C		D		E	F		G		A		B
نام انگلیسی	Sharp		C SHARP		D sharp			F sharp		G sharp		A sharp	
	Flat		D FLAT		E flat			G flat		A flat		B flat	
نماد	Sharp		C♯		D♯			F♯		G♯		A♯	
	Flat		D ♭		E ♭			G ♭		A ♭		B ♭	
لاتین، ایتالیایی؛ ایران	Natural	***DO***		***RE***		***MI***	***FA***		***SOL***		***LA***		***SI***
	Sharp		Do diesis		Re diesis			Fa diesis		Sol diesis		La diesis	
	Flat		Re bemolle		Mi bemolle			Sol bemolle		La bemolle		Si bemolle	
فرکانس تقریبی		**262**	**277**	**294**	**311**	**330**	**349**	**370**	**392**	**415**	**440**	**466**	**494**
شماره نت در MIDI		**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**

جدول ۳-۱ نام گذاری نت های موسیقی در زبان های مختلف و نیز اطلاعات فرکانسی و شماره نت در فایل نت

۳-۶-۱-۳ تغییر دادن ماتریس نت ها :

به کمک تابع shift می توانیم مقادیر ماتریس نت ها را تغییر دهیم . شکل کلی تابع مذکور به شکل زیر می باشد:

nm = shift(nmat,dim,amount,<timetype>);

معرفی آرگومان های ورودی :

Nmat: ماتریس نت ها

Dim: معیاری که خواستار تغییر در آن می باشیم ('vel' , 'chan' , 'pitch' , 'dur' , 'onset')

Amount : مقدار و میزان جابجایی.

Timetype: نوع زمان نشان داده شده 'beat' ' یا ضرب و 'sec' یا واحد ثانیه.

۳-۷ نتیجه گیری

در فصل جاری به تشریح روش ها و مکانیزم عملکرد بکار گرفته شده در این پایان نامه پرداخته شد. روش ها شامل روشهایی که جهت پیش پردازش روی تصاویر, قطعه بندی, و اجرای نهایی صدا می گردد. در ادامه و در فصل بعدی به تشریح دقیق روش ها بروی مثالها و ورودی های مختلف خواهیم پرداخت و عملکرد روش های بکار گرفته شده را مورد آزمون و بررسی قرار خواهیم داد و خروجی نهایی که همانا اجرای موسیقی است را نمایش خواهیم داد.

فصل چهارم:

تشریح روش های بکار رفته در پیاده سازی

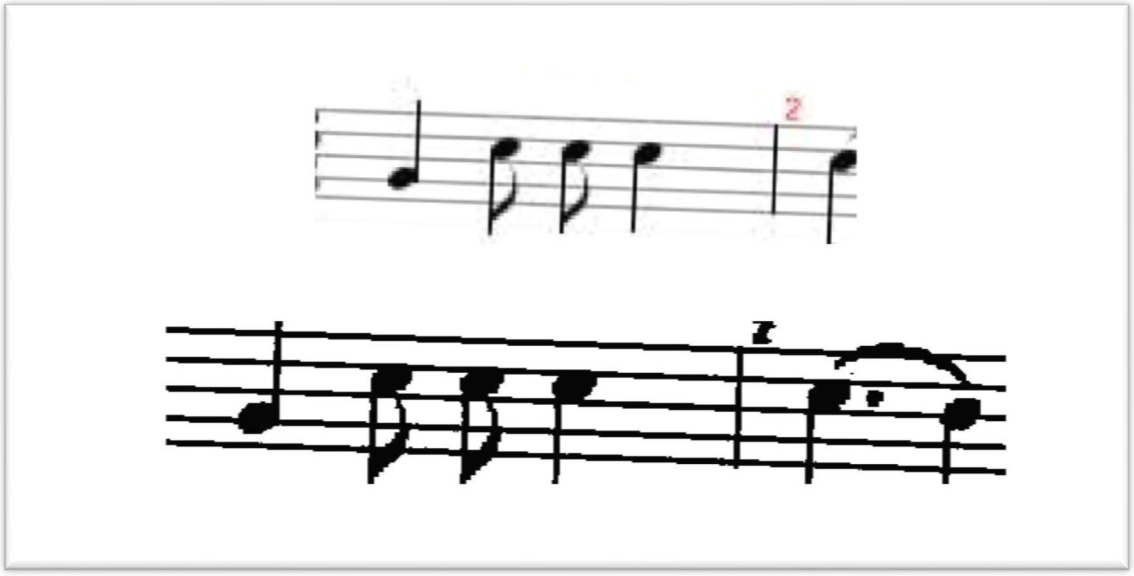
۴ـ ۱ـ مقدمه

در فصل گذشته مباحثی به تشریح تکنیک هایی پرداخته شدکه به کمک آن می توان کلیه عملیات مربوط به شناسایی نت های پایه ای موسیقی و اجرا را به انجام رساند. در این فصل به تشریح هر یک از روشها در عمل خواهیم پرداخت و تأثیر و کاربرد هر یک از روشها را به تفصیل بیان می نمائیم. بسیاری از اصول بکار رفته در سیستم OMR مشابه سیستم های شناسایی کاراکتر بوده لذا بررسی آنها را در فصل دوم صورت داده و در این فصل به تشریح عملیات کلیدی در فرآیند شناسایی و اجرا می پردازیم. تشریح روشها همراه با مثال و نمونه های آن بیان خواهد شد.

۴ـ ۲ـ اصلاح زاویه تصویر:

اصلاح زاویه تصویر ورودی همواره ما را در جهت بهبود نتیجه سوق می دهد و یکی از پر کاربردترین روشها در زمان پیش پردازش بوده خصوصاً سیستم های شناسایی کننده های متنی می باشد. با توجه به قوانین حاکم بر هنر موسیقی و زبان موسیقی تحت هر شرایطی یک قطعه موسیقی بایستی دارای خطوط حامل بوده و اساساً وجود خطوط حامل کاملاً الزامی است. از طرفی پردازشهایی که روی علائم موسیقی صورت می گیرد کاملاً وابسته به محل قرارگیری آنها روی خطوط حامل می باشد.

در بسیاری مواقع به جهت سهولت در انجام، نیاز به باینری کردن تصویر بوده لذا در صورت کج بودن تصویر ورودی شاهد پله پله شدن خطوط موجود در تصویر پس از باینری کردن می باشیم.



شکل ۴-۱ـ نمونه خط غیرمستقیم که در زمان باینری شدن پله پله شده است.

شناخت موقعیت علائم موسیقی, رو و یا ما بین خطوط, عامل کلیدی در شناسایی می باشد پس قبل از هر اقدام, ابتدا بایستی صحت تصویر ورودی را از حیث زاویه ورودی مورد توجه قرار داده و با دقت بالا آن را اصلاح نمائیم. برای این منظور ۲ مرحله بایستی صورت پذیرد.

- بدست آوردن زاویه تصویر
- اصلاح زاویه تصویر

۴ـ ۲ـ ۱ـ محاسبه زاویه تصویر ورودی

همانگونه که در فصل سوم و در بخش ۳ـ ۲ به معرفی تکنیک تبدیل رادون و قابلیت های بالای این روش در اصلاح تصاویر پرداختیم. حال نحوه کارکرد آن را به صورت عملی بیان می نماییم.

برای اصلاح تصویر، ابتدا می بایست زاویه دقیق تصویر ورودی را محاسبه نمائیم تا براساس آن و اختلاف آن با زاویه قائم 90° بتوانیم تصویر را به لحاظ کجی اصلاح نمائیم.

مطابق رابطه ۳ـ ۴ که به نمایش گسسته تبدیل رادون می پردازد, تبدیل رادون در واقع، جمع مقادیر نقاط یک تصویر در امتداد خطوط موازی، عمود بر محوری است که با محور افقی زاویه θ می سازد. لذا جهت محاسبه کلید مقادیر تغییرات زاویه از $\Delta\theta = \phi \ldots ۱۸۰^{\circ}$ در نظر گرفته شده است.

تابع زیر در نرم افزار مقلب جهت محاسبه تبدیل رادون به کار گرفته می شود.

$$[R, xP] = radon(I, theta)$$

در صورتی که θ به صورت عدد به تابع داده شود، R شامل یک ماتریس ستونی حاوی مقادیر محاسبه شده در زاویه θ خواهد بود اما در صورتی که θ به صورت بردار به ورودی تابع داده شود. R ماتریسی عمودی است که هر ستون نمایانگر مقدار محاسبه شده تابع رادون در هر زاویه θ می باشد.

بردار xp نیز حاوی اطلاعات مختصات شعاعی متناظر با هر ردیف از بردار R می باشد. مختصات شعاعی موجود در بردار xP شامل مقادیری است که در امتداد محور xها قرار دارند به طوری که برخلاف عقربه های ساعت و براساس محور xها روی زاویه های مختلف θ گرایش دارد. با اعمال تابع تبدیل رادون روی تصویر حاوی قطعه موسیقی ماتریس xP, R تشکیل می شود.

از آنجایی که خطوط حامل همواره در راستای محور x بوده و با محور y زاویه قائم می سازد لذا در صورت کاملاً صاف بودن تصویر ورودی بیشترین مقدار بایستی مکان $\theta = 90^{\circ}$ یافت شود. از آنجایی که تصویر ورودی حاوی خطوط حامل است و خطوط حامل در امتداد محور xها و به دفعات تکرار شده است لذا همواره بیشترین مقدار حول زاویه ۹۰° نوسان دارد.

دقت این روش بسیار وابسته به گامهای است که در بردار θ وجود دارد. گامها می تواند به صورت اعداد صحیح و یا حقیقی ظاهر شود. هر چه گامها کوچکتر انتخاب شود دقت الگوریتم نیز بیشتر خواهد بود. لذا در این پروژه گامهای 0.05 برای یافتن زاویه تصویر ورودی انتخاب شده است.

$$\theta = 0, 0.05, 0.1, 0.15, \ldots, 180^{\circ}$$

پس از محاسبه تبدیل رادون با مقادیر θ ماتریس R تشکیل می شود. با توجه به گامهای انتخاب شده ماتریس R شامل 3601 ستون می باشد(به تعداد عناصر θ). ستون متناظر با بزرگترین مقدار موجود در ماتریس R، زاویه تصویر ورودی را به ما خواهد داد.

$$COLIdx = \max(\max(R))$$ رابطه ۴ـ ۱

حال جهت پیدا کردن اختلاف زاویه تصویر ورودی با زاویه 90° از رابطه ۴ـ ۲ استفاده می نمائیم.

$$\mathrm{Im}age - Angle = \frac{COLIdx}{\frac{1}{\theta steps}}$$ رابطه ۴-۲

۴ـ ۲ـ ۲ـ اصلاح زاویه تصویر

در بخش گذشته اقدام به محاسبه زاویه تصویر ورودی براساس اختلافی که اندیس ستون بزرگترین مقدار ماتریس R در آن یافت شد نمودیم حال به کمک رابطه ۴ـ ۳ اقدام به یافتن اختلاف زاویه تصویر ورودی با زاویه 90° می نمائیم.

$$Difference - Angle = \frac{90 - COLIdx}{\frac{1}{\theta steps}}$$ رابطه ۴-۳

با توجه به هزینه محاسباتی بالا چرخش تصویر، اصلاح تصویر را طی شرایطی انجام می دهیم.

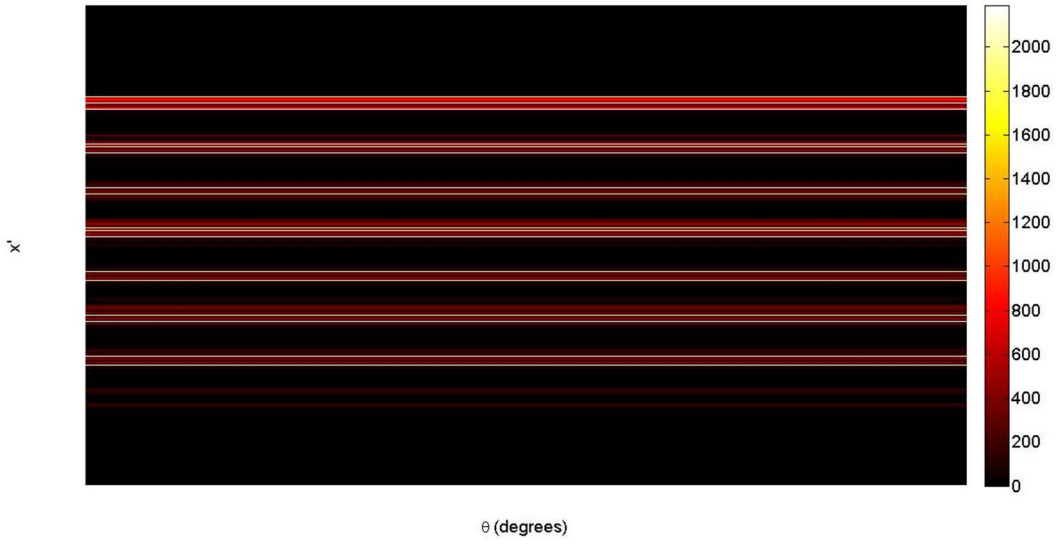
$$Difference-Angle\begin{cases}\geq 0.1 & rotate\ image\\ otherwise & don't\ rotate\ image\end{cases}$$

۴ـ ۳ـ پیدا کردن خطوط حامل

پیشتر در بخش ۳ـ ۳ به تشریح کاربرد تبدیل رادون در پیدا کردن خطوط در تصویر پرداختیم. حال به تفصیل به بیان روش فوق می پردازیم.

همانگونه که بیان شد تمامی قطعات موسیقی بدون استثناء حاوی خطوط حامل می باشد، لذا یافتن این خطوط و محل دقیق قرارگیری آنها بسیار مهم و حائز اهمیت می باشد. از طرفی ضخامت خطوط حامل در تصاویر مختلف متفاوت بوده و در بسیاری از موارد ضخامت خطوط حامل در یک تصویر نیز دارای مقادیر متفاوت می باشد.

با توجه به اصلاح زاویه در ابتدای کار، در این مرحله تبدیل رادون را صرفاً در زاویه $\theta = 90^{\circ}$ محاسبه می کنیم. از آنجایی که تصویر ورودی به این مرحله دارای زاویه صحیح می باشد و خطوط حامل در راستای محور xها بوده لذا طبیعتا بیشترین مقادیر در ماتریس تک ستونی R، معرف وجود خطوط حامل بوده و بایستی نقاط ماکزیمم ماتریس تک ستونی R را استخراج نمائیم.



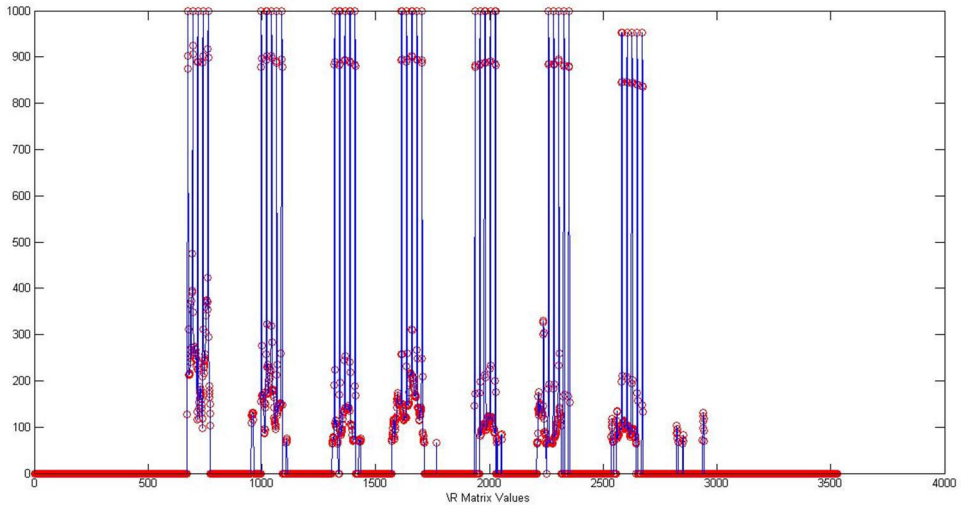
شکل ۴ـ ۳ـ نمودار ماتریس تک ستونی R.

همانگونه که در شکل ۴ـ ۳ مشاهده می شود تصویر ورودی شامل تعدادی نقاط روشن نسبت به دیگر نقاط می باشد که در دسته های مختلف قرار دارند. تعداد دسته ها در واقع بیان کننده تعداد خطوط حامل می باشند[1]. با توجه به مقادیر متفاوت در ماتریس R ابتدا به کمک رابطه ۴ـ ۴ اقدام به استاندارد سازی مقادیر می نمائیم.

$$s\tan dard - Values = \frac{x - \min(R)}{\max(R) - \min(R)} \times 1000$$ رابطه ۴ـ ۴

پس از استاندارد سازی داده ها میانگین آنها را نیز محاسبه و ذخیره می نمائیم. سپس کلیه مقادیر کوچکتر از مقدار میانگین را در ماتریس R صفر می نمائیم.

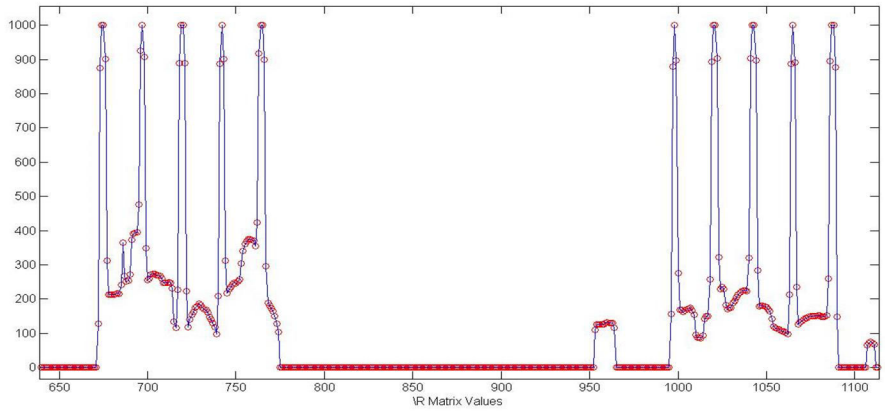
[1] هر مجموعه از خطوط حامل شامل پنج خط می باشد که توضیح آن در بخش ضمائم آورده شده است.



شکل ۴ـ ۵ـ نمودار میله ای به همراه نقاط علامت گذاری شده ماتریس $\boldsymbol{R}$ پس از استاندارد سازی

همانطور که در شکل ۴ـ ۵ مشاهده می شود برای تصویر ورودی تعداد ۷ سگمنت نمایش داده شده است که این ۷ سگمنت بیانگر وجود ۷ مجموعه از خطوط حامل که هر دسته از خطوط طبق قواعد و قوانین حاکم بر زبان موسیقی حاوی ۵ خط است، می باشد.

در صورت بزرگنمایی هر یک از سگمنت ها تعداد نقاط بیش از یک در قله هر سگمنت مشاهده می شود.



شکل ۴ـ ۶ بزرگنمایی سگمنت اول و دوم ماتریس $\boldsymbol{R}$

همچنین در نقاطی نزدیک به رأس, نقاطی نیز مشاهده می شود که خود بیانگر وجود خط در آن ناحیه می باشد. همانگونه که پیش تر نیز توضیح داده شد خطوط موجود در قطعه موسیقی با ضخامت های متفاوت وجود دارند و از آنجایی که استراتژی ما بر حذف خطوط حامل بنا شده است، لذا بایستی کلید خطوط موجود در تصویر با هر ضخامتی شناسایی شده و در مرحله بعد حذف گردند. پس در نتیجه علاوه بر نقاط ماگزیمم, سایر نقاط که بیانگر وجود خطوط دیگر می باشند نیز بایستی شناسایی و مختصات آن ذخیره گردند. برای این منظور ابتدا میانگین و انحراف معیار مقادیر ماتریس R را محاسبه می نمائیم.

مطابق رابطه ۵ـ ۴ و ۶ـ ۴ میانگین و انحراف معیار را محاسبه می نمائیم.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

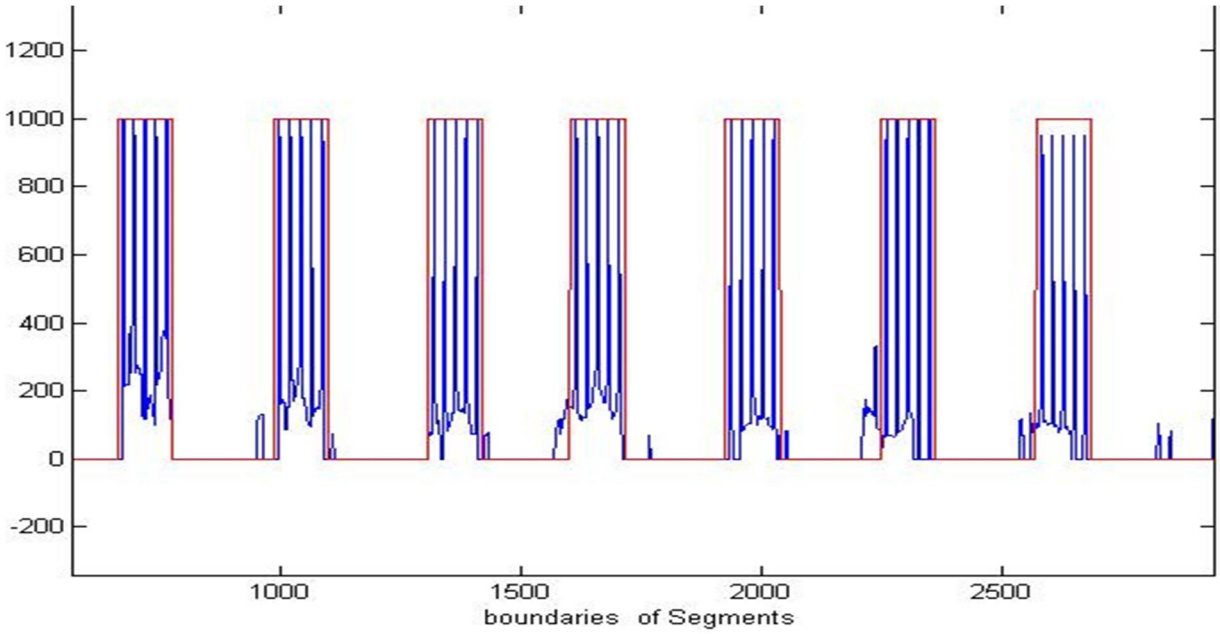
رابطه ۵ـ ۴

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{N-1}}$$

رابطه ۶ـ ۴

پس کلیه مقادیر موجود در ماتریس R که مقادیر آنها بزرگتر از مجموع میانگین داده به علاوه انحراف معیار می باشند را برابر با بزرگترین مقدار موجود در ماتریس R پس از استاندارد نمودن آن می نمائیم. برای اطمینان از اینکه مقادیر نزدیک به پیک هر سگمنت ماکزیمم گردند ضریب 1.5 را نیز در محاسبه دخالت می دهیم. کلید مقادیر بدست آمده را در ماتریسی که هم اندازه ماتریس R می باشد و کلیه مقادیر اولیه آن برابر صفر است ذخیره می نمائیم و پس از آن کلیه مقادیری که شرایط ما را دارا بودند را به ماکزیمم مقدار تغییر می دهیم.

به این ترتیب ماتریس جدید مرزهای بین هر سگمنت را مشخص می نماید. شکل ۴ـ ۷ نمونه بزرگنمایی شده ماتریس جدید را نمایش می دهد.



شکل ۴ـ ۷ـ نمودار بزرگنمایی شده $R2$ که مرز بین هر سگمنت را مشخص نموده است.

حال می توانیم در هر سگمنت به دنبال بیشترین مقادیر جهت بدست آوردن مختصات خطوط باشیم.

جدول ۴ـ ۱ مقادیر مرزی هر سگمنت به همراه طول آن را نمایش می دهد.

نقطه شروع	نقطه پایان	فاصله بین هر سگمنت
۶۶۳	۷۷۷	۱۱۴
۹۸۷	۱۱۰۰	۱۱۳
۱۳۰۹	۱۴۲۲	۱۱۳
۱۶۰۳	۱۷۱۶	۱۱۳
۱۹۲۶	۲۰۳۹	۱۱۳
۲۲۴۹	۲۳۶۲	۱۱۳
۲۵۷۱	۲۶۸۴	۱۱۳
میانگین فاصله		۱۱۳

جدول ۴ـ ۱ مقادیر شروع و پایان هر سگمنت به همراه طول آن

بایستی توجه داشت مختصات مرزی بدست آمده اگر چه دقیق است اما براساس خروجی الگوریتم تبدیل رادون محاسبه شده است لذا بایستی این مقادیر به مقادیر واقعی که بیانگر مختصات واقعی هر دسته از خطوط در تصویر می باشد تبدیل گردد، برای این منظور کلیه مقادیر را با کوچکترین مقدار موجود در ماتریس xp جمع نموده و از تعداد ستون های تصویر تقسیم بر ۲ کم می کنیم.

با این روش مقادیر مرزی بدست آمده کاملاً منطبق با مختصات مرزی در تصویر ورودی خواهند بود.

۴ـ ۳ـ ۱ـ محاسبه خطوط اصلی و خطوط در سایه:

همانگونه که قبلاً نیز اشاره شد ضخامت خطوط حامل متفاوت است. گاه خطی دارای ضخامت ۵ پیکسل بوده و گاه ضخامت ۲ و حتی ۱ پیکسل را دارا می باشد. اما نکته ای که در این میان وجود دارد یافتن خطی است که بیشترین مقدار را داشته تا مختصات آن در زمان شناسایی علائم مورد استفاده قرار گیرد. از طرفی یافتن کلیه خطوط (خط اصلی و خطوط سایه) جهت حذف آنها از تصویر مورد نیاز می باشد.

۴ـ ۳ـ ۱ـ ۱ـ محاسبه مکان خطوط اصلی و خطوط حامل

به منظور پیدا کردن خطوط اصلی (با ضخامت یک پیکسل) در هر سگمنت اقدام به جستجوی ۵ بزرگترین مقدار می نمائیم. برای این منظور ابتدا نقطه ماکزیمم را بدست آورده و پس از ذخیره مشخصات آن مقدار آن را صفر نموده و به دنبال دومین مقدار ماکزیمم می نمائیم و این روش تا ۵ مرحله ادامه می دهیم (کلیه خطوط حامل شامل پنج خط می باشند).

۴ـ ۳ـ ۱ـ ۲ـ محاسبه مکان خطوط در سایه

پس از محاسبه مکان های پیک در نمودار R و ذخیره آنها به عنوان ۵ خط اصلی خطوط حامل، بایستی اقدام به یافتن خطوط در سایه نمود. همانگونه که در شکل ۴-۶ مشاهده نمودیم تعدادی نقطه حول نقطه ماکزیمم وجود دارد که همان خطوط سایه می باشند. مجموع این خطوط با خط اصلی، ضخامت خط را معین می نماید. لذا برای این منظور به دنبال نقاط نزدیک به نقطه پیک می گردیم. با فرض اینکه x شامل نقاط پیک در هر سگمنت باشد با محاسبه مشتق مرتبه اول، تفاضل بین نقاط پیک را محاسبه می نمائیم.

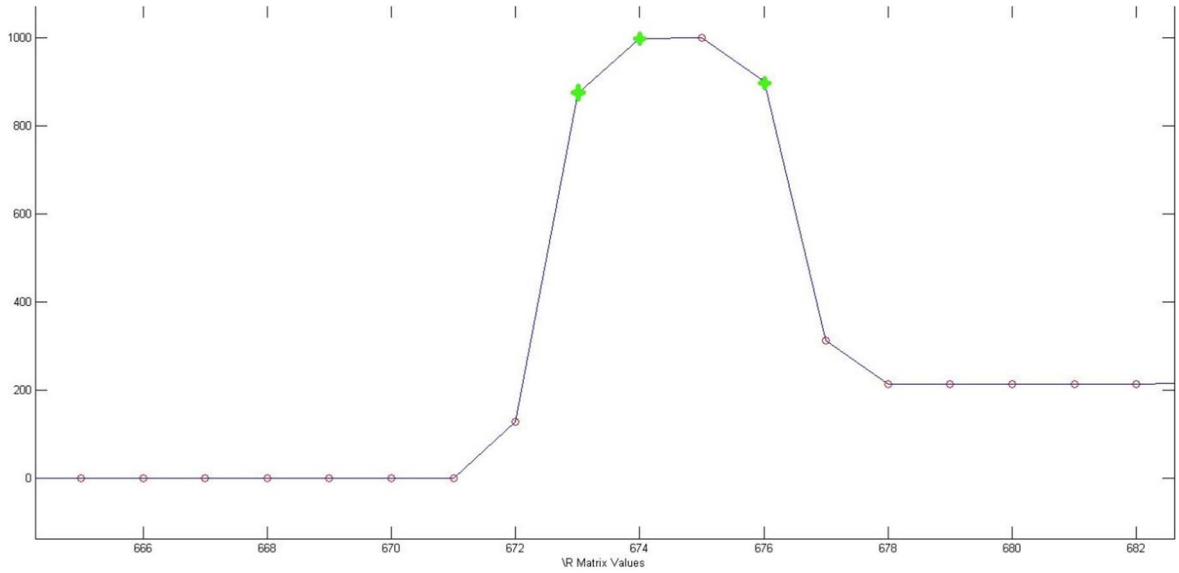
به عنوان مثال:

$$x = 675,697,720,742,764$$

رابطه ۶-۵

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

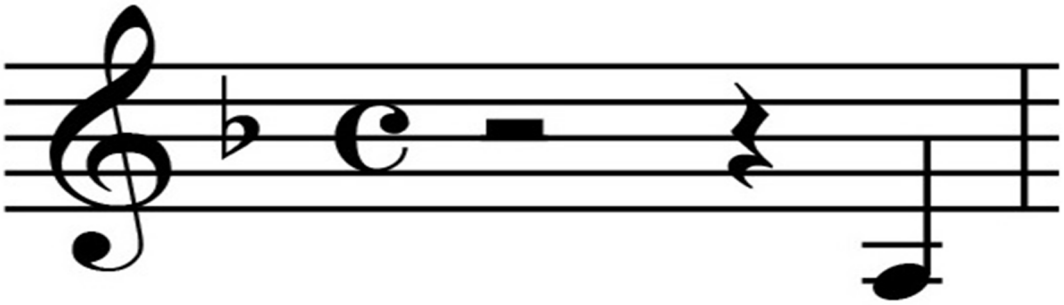
با میانگین گرفتن از حاصل مشتق، فاصله بین هر نقطه پیک محاسبه می شود. حال محدوده مورد نیاز حول هر نقطه پیک بدست آمده را به جستجوی بزرگترین نقاط که از میانگین داده ها بزرگترند می گردیم.



شکل ۴ـ ۸ـ نقاط یافت شده حول هر نقطه پیک با شرایط اشاره شده

۴ـ ۳ـ ۲ـ محاسبه مکان خطوط غیراصلی

همانگونه که در بخش ضمائم و در معرفی زبان موسیقی به آن اشاره شده است، علائم و نت های موسیقی گاه روی خطوط غیر از خطوط اصلی قرار می گیرند.



۴ـ ۹ـ نمونه نت موسیقی که در محلی غیر از خطوط حامل واقع شده است.

خطوط غیر اصلی هم در قسمت بالای خطوط حامل و هم در قسمت پایین خطوط حامل ظاهر می شوند. برای یافتن مکان این خطوط کافی است میانگین فاصله بین خطوط حامل را محاسبه و براساس آن مختصات خطوط غیراصلی را ذخیره نمود. در اینجا تعداد خطوط غیر اصلی در نظر گرفته شده ۳ عدد در بالا و ۳ عدد در پایین خطوط حامل می باشد. خطوط غیر اصلی را هنگان ذخیره سازی با علامت ۱- علامت گذاری می کنیم.

۴ـ ۳ـ ۳ـ حذف خطوط میزان۱

خطوط حامل به وسیله خطوط عمودی در مکانهای مختلف قطع می گردند. هر قسمت را میزان و به خطوطی که میزان ها را تشکیل می دهند، خط میزان گفته می شود.[۲]

حذف خطوط میزان علاوه بر تأثیرگذاری روی عملیات قطعه بندی، باعث کاهش حجم محاسبات نیز می گردد.

در بخش ۴ـ ۳ـ ۱ مکان خط اصلی و خطوط سایه را بدست آوردیم. حال با حرکت روی خط حامل و محاسبه تعداد پیکسل های روشن از کوچکترین مختصات خط سایه تا بزرگترین آن در مکانهایی که تعداد پیکسلهای روشن برابر با اختلاف بزرگترین و کوچکترین مختصات باشد را به عنوان مکان وجود خط میزان کاندید می کنیم.

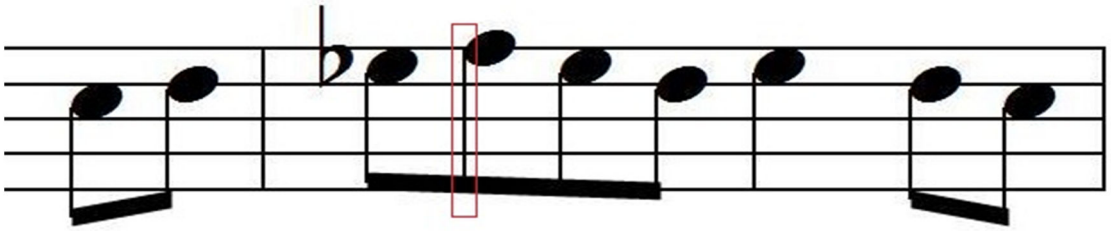
از آنجایی که مکان یافت شده ممکن است در اثر وجود دنباله یک نت باشد که برحسب اتفاق تمامی عرض خطوط حامل را در برگرفته باشد، نقطه کاندید را جهت بررسی به تابعی به این منظور ارسال می نمائیم تا از حذف المان های اصلی پرهیز کنیم.

---

[1] Bar Lines

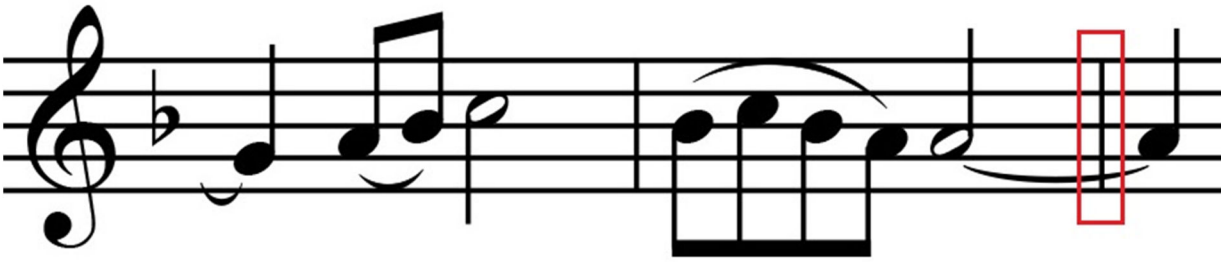
[2] رجوع به بخش ضمائم

همانگونه که با اشکال نت ها در بخش ضمیمه آشنایی داریم در صورتی که خطی که حامل را قطع نموده باشد متعلق به یک نت باشد پس قطعاً در بالا و پایین خط شاهد وجود تعدادی پیکسل که نشانگر سر نت است می باشیم لذا در صورتی که مجموع نقاط بالا و پائین خط برابر صفر باشد، از خط میزان بودن آن اطمینان حاصل می نمائیم در غیر اینصورت آن مکان به عنوان مکانی که المان موسیقی در آن واقع شده است را نگه می داریم.



شکل ۴ـ ۱۰ـ نمونه نتی که ممکن است به اشتباه خط میزان فرض شود

همچنین ممکن است علامت خطوط اتحاد و یا اتصال در امتداد دو میزان وجود داشته باشد لذا پس از حذف خط میزان، از عملیات اتساع[1] استفاده می نمائیم. این عمل علاوه بر حفظ ساختارهای موجود در تصویر که جزء المانهای ضروری می باشند خطوط حامل را نیز حفظ می نماید.



شکل ۴ـ ۱۱ـ نمونه خط اتحاد که خط میزان را قطع نموده است.

[1] Dilate

۴ـ ۳ـ ۴ـ حذف خطوط حامل

یکی از مهمترین پیش پردازشها قبل از اقدام به قطعه بندی المانهای موجود در قطعه موسیقی, حذف خطوط حامل از تصویر ورودی است. پیش از این نحوه استخراج مکان خطوط حامل اصلی و غیر اصلی را بیان کردیم. حال با داشتن مختصات خطوط اقدام به حذف آن می نمائیم.

در زمان حذف خطوط حامل باید توجه داشت که با توجه به اینکه کلیه المانهای موسیقی روی خطوط حامل قرار دارند با حذف این خطوط طبیعتاً المانها نیز دچار تقطیع گردیده که این موضوع باعث بروز مشکلات آتی در زمان قطعه بندی جزئی خواهد شد.

از طرفی وجود خطوط اتحاد و خطوط اتصال در تصویر که به صورت کمان ظاهر می شود عملیات ترمیم را با مشکل مواجه خواهد کرد.

شکل ۴ـ ۱۲ـ نمونه خط حامل حذف شده از تصویر

به منظور بازسازی المانهای قطع شده ابتدا ضخامت خط حامل را به دو نیمه تقسیم می نمائیم. پس از آن با در نظر گرفتن یک ردیف بالاتر و یک ردیف پایین تر از خط حامل جمع مقادیر یافت شده در عرض تصویر را محاسبه و در ۲ ماتریس به نام های D_2, D_1 ذخیره می نمائیم. (D_1 مقادیر مثبت و D_2 مقادیر منفی) مطابق نمودار ۴ـ ۱ با تلفیق این دو بردار مکانهایی که در آن اشتراک وجود دارد را پیدا کرده به عنوان مکانی که المانی در آن مکان وجود داشته و به واسطه حذف خط حامل، منقطع گردیده را در نظر

می گیریم و پس از آن به کمک عملیات اتساع در مکانهایی که اشتراک وجود دارد اقدام به پر کردن پیکسل های مربوط می نمائیم.

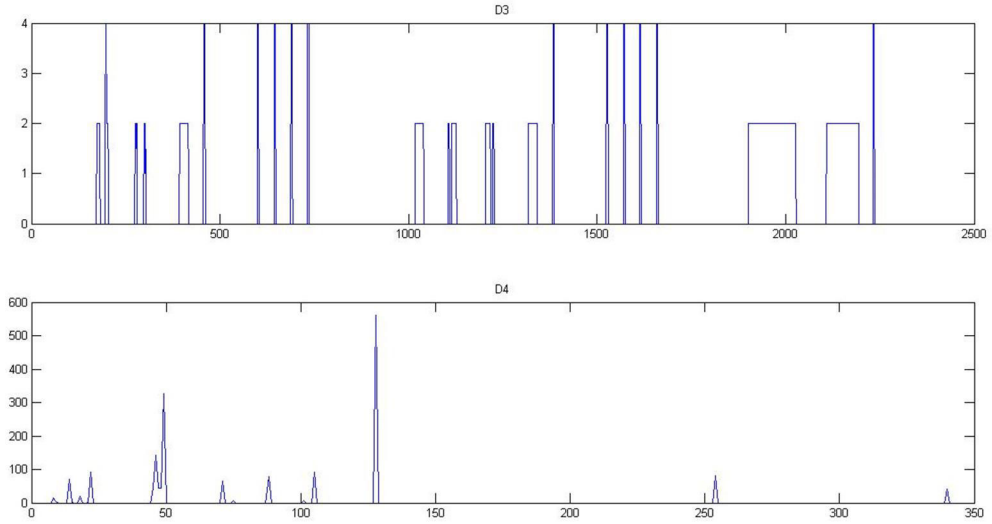
شکل ۴ـ ۱۳ـ نمونه خط حذف شده و عملیات بازسازی المانهای منقطع شده

اما در برخی شرایط در شرایطی که خط حامل از روی خطوط اتحاد و یا اتصال عبور نموده باشد، روش فوق روش کارآمدی نخواهد بود. شکل ۴ـ ۱۴ به وضوح دلیل این مشکل را بیان می نماید.

شکل ۴ـ ۱۴ـ حذف قسمتهایی از خط اتحاد به دلیل حذف خط حامل

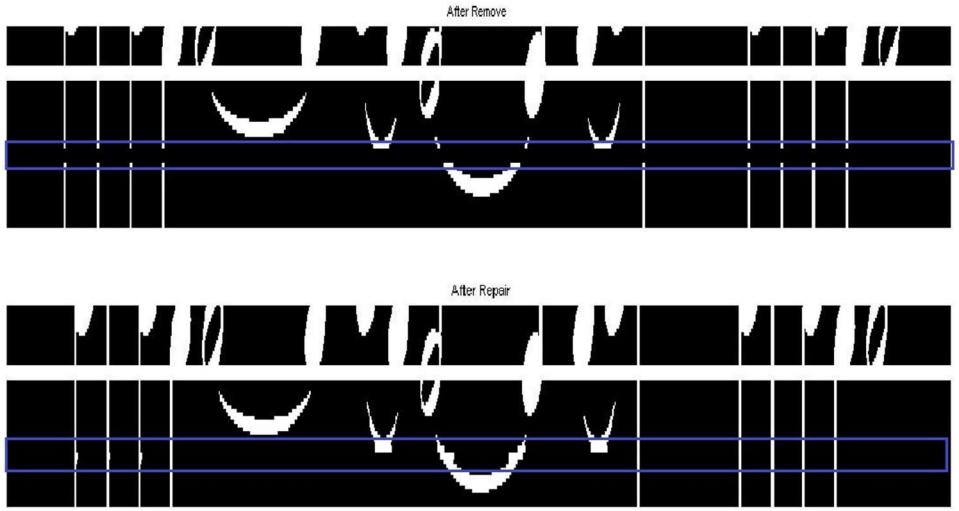
به دلیل کمانی شکل بودن خطوط اتحاد و اتصال، پس از حذف خط حامل عبور کننده از آنها در نمودار تشکیل شده در مکانی که خط اتحاد حذف شده هیچ گونه اشتراکی وجود ندارد لذا به منظور ترمیم آن ابتدا تفاضل دو بردار D_2, D_1 را محاسبه می نمایم و در بردار $D3$ ذخیره می کنیم. سپس اندیس مکانهایی که مقدار غیر صفر دارند را در بردار D_3 بدست می آوریم و پس از آن روی مقادیر اندیس ها مشتق مرتبه اول را محاسبه نموده و در بردار D_4 ذخیره می نمائیم. با توجه به اینکه اندیس های ذخیره

در اغلب مکانها دنباله ای مرتب از اعداد می باشند لذا پس از عملیات مشتق گیری تفاضل بسیاری از آنها برابر با عدد ۱ خواهد بود از این رو کلیه سلولهای با مقدار ۱ را به مقدار صفر تغییر می دهیم.



نمودار ۴ـ ۲ـ نمودار مربوط به متغیر D_3 و متغیر D_4 پس از عملیات مشتق گیری

همانگونه که در نمودار ۴ـ ۲ و در نمودار مربوط به D_4 مشاهده می شود در مکانهایی که در ان مقادیر بسیار بزرگ وجود دارد نشان وجود فاصله بسیار زیادی است که بین دو المان وجود داشته و خط حامل بین آن حذف گردیده است اما در مکانهایی که مقدار D_4 کوچکتر از مقدار ضخامت خط حامل بوده نشان از مکانی دارد که به احتمال غریب به یقین کمانی وجود داشته است و با توجه به انحناهای مختلف کمان خطوط اتحاد و اتصال برای اطمینان ضخامت خط حامل را به علاوه ۲ کرده تا در صورتی که انحنای کمان بسیار مماس با خط حامل بوده باشد آن را دربرگیرد. لذا پس از یافتن نقاط اشاره شده اقدام به ترمیم نقطه می نمائیم.



شکل ۴ـ ۱۵ـ نمونه خط اتحاد منقطع شده و تصویر پس از ترمیم قسمت منقطع شده

به این ترتیب مشکل قطع خطوط اتحاد و اتصال پس از حذف خطوط حامل عبور کننده از روی آنها حل شده و این امر کمک شایانی به قطعه بندی دقیق تصویر خواهد کرد.

۴ـ ۴ـ استخراج مرزهای هر دسته از خطوط حامل

پس از حذف خطوط میزان و نیز خطوط حامل و نیز محاسبه خطوط فرعی؛ برای عملیات قطعه بندی کلی تصویر نیاز به محاسبه محدوده هر یک از خطوط میزان داریم. پیشتر به کمک محاسبه انتقال رادون محل دقیق هر یک خطوط موجود روی خطوط حامل استخراج گردید و بر مبنای میانگین خطوط اصلی، خطوط فرعی نیز محاسبه و ذخیره گردید. مطابق شکل ۴ـ۷ فاصله بین هر دسته از خطوط حامل نیز قابل محاسبه و استخراج می باشد. برای بدست آوردن فاصله بین هر دسته ابتدا فاصله بین سگمنت ها را محاسبه می کنیم. مبنای این محاسبه فاصله آخرین خط از سگمنت قبلی و اولین خط از سگمنت بعدی خواهد بود. این محاسبه برای کلیه سگمنتها به استثنای سگمنت اول و سگمنت آخر صورت می پذیرد

برای محاسبه فاصله بالا و پایین سگمنت های اول و آخر از میانگین بدست آمده از سگمنت های میانی استفاده می نمائیم البته با این شرط که محدوده سگمنت از اندازه تصویر ورودی تجاوز ننماید.

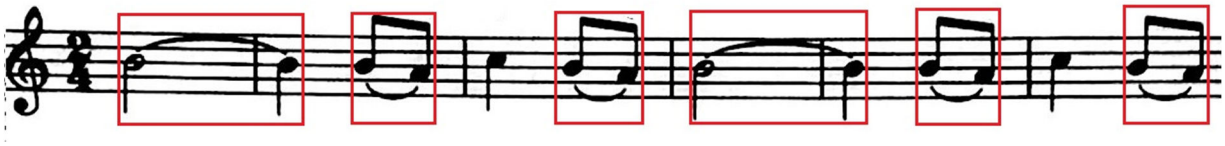
جدول ۴ـ ۲ نمونه مقادیر بدست آمده و میانگین فاصله بین هر سگمنت را نشان می دهد.

نقطه شروع	نقطه پایان	ارتفاع در قطعه
۲۳۲	۵۴۹	۳۱۷
۵۵۵	۸۷۲	۳۱۷
۸۷۸	۱۱۹۵	۳۱۷
۱۲۰۰	۱۵۱۷	۳۱۷
۱۴۹۴	۱۸۱۲	۳۱۸
۱۸۱۷	۲۱۳۴	۳۱۷
۲۱۴۰	۲۴۵۷	۳۱۷
میانگین		۳۱۷

جدول ۴ـ ۲ـ فاصله های بین هر سگمنت

۴ـ ۵ـ قطعه بندی کلی تصویر:

قبل از قطعه بندی اجزای تشکیل دهنده قطعه موسیقی ابتدا بایستی اقدام به قطعه بندی کلی آنها نمائیم. دلیل نیاز به قطعه بندی کلی هر دسته از خطوط حامل, وجود علائم و نشانه های است که از نظر نقطه مکانی با دیگر المانها دارای نقاط اشتراک می باشند. شکل ۴ـ ۱۶ به وضوح دلیل نیاز به قطعه بندی کلی را نشان می دهد.



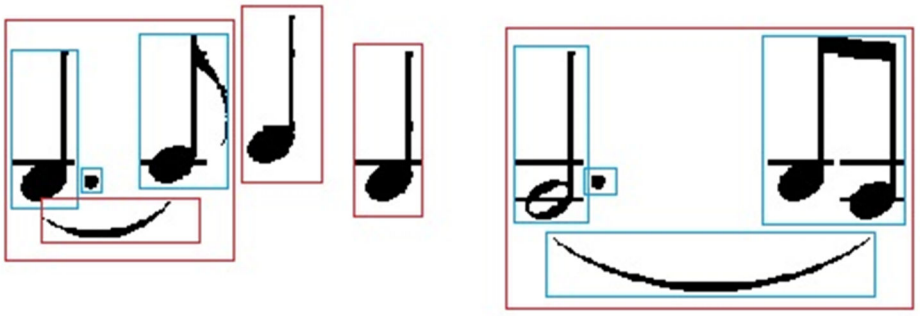
شکل ۴ـ ۱۶ـ وجود علائم و نشانه هایی که از منظر عمودی دارای نقطه مشترک با دیگر علائم می باشند.

این مرحله از کار به کمک نقاط مرزی که در بخش ۴ـ ۴ تشریح شد صورت می پذیرد. روش کار به این صورت است که در هر دسته از خطوط حامل حرکت کرده و هر مکانی که مجموع پیکسلهای آن بزرگتر از صفر باشد را به عنوان نقطه شروع علامت گذاری می کنیم. پس با حرکت رو به جلو (حرکت از چپ به راست) در هر مکانی که مجموع پیکسلها صفر شد، آن مکان را به عنوان نقطه پایانی قطعه در نظر می گیریم. در بحث قطعه بندی کلی دو نکته بسیار مهم حائز اهمیت است.

۱ـ ذخیره و نگهداری مختصات قطعه در تصویر

۲ـ شماره گذاری هر قطعه

قطعات کلی بسته به نوع المان بکار رفته در آن ممکن است خود دارای یک و یا بیش از یک المان باشند. از این رو در قطعه بندی کلی فارغ از جداسازی تک تک المانهای تشکیل دهنده، اقدام به قطعه بندی می نمائیم. شکل ۴ـ ۱۷ نمونه ۴ قطعه را که دارای یک المان و بیش از یک المان می باشند را نمایش می دهد.



شکل ۴ـ ۱۷ـ نمونه ۲ قطعه دارای یک و بیش از یک المان

۴ـ ۶ـ قطعه بندی جزئی جهت استخراج المانهای موسیقی

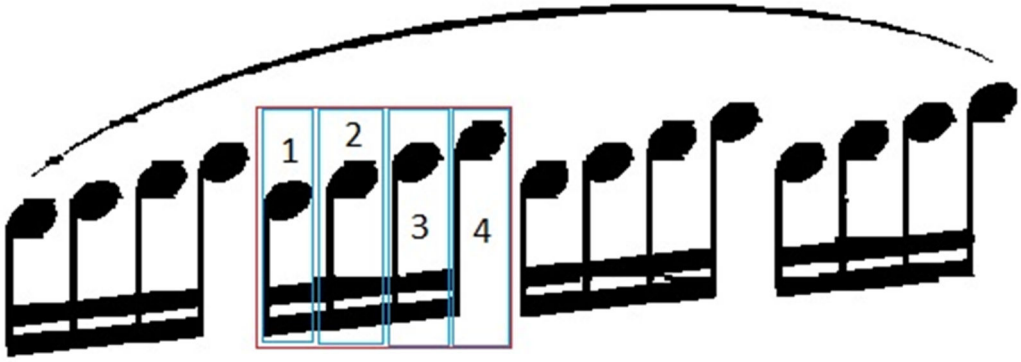
یکی از مهم ترین بخش کار قبل شناسایی و اجرا، یافتن صحیح المانهای تشکیل دهنده از هر قطعه می باشد. اهمیت بالای این موضوع در نتیجه نهای است زیرا در صورتیکه این بخش دارای مشکل باشد، هر فردی که آشنایی کمی نیز با هنر موسیقی داشته باشد, براحتی آن را شناسایی خواهد کرد. حالتهای بسیار پیچیده و استثنائی در هر قطعه موسیقی کاملاً نمایان است که پیدا کردن الگوریتم کلی جهت قطعه بندی جزئی را با مشکلات فراوانی مواجه خواهد کرد.

نت های چنگ به هم متصل، خطوط اتحاد و اتصال، چسبندگی هایی که گاه به دلیل مشکلات چاپی به وجود می آیند، فاصله های عرضی متفاوت بین هر المان از جمله مشکلاتی است که سیستم قطعه بندی کننده بایستی بتواند بخوبی آنها را تشخیص و در صورت نیاز اصلاحاتی را نیز روی آن انجام دهد.

از طرفی تنوع بسیار در ظاهر نت ها نیز خود دلیلی است که دقت بسیار بالای قطعه بندی را طلب می کند تا سیستم شناسایی کننده بتواند آن را به درستی شناسایی نماید. در این بخش به تشریح هر یک از مشکلات عنوان شده می پردازیم و راهکار پیشنهادی برای هر یک از مسائل را به تفصیل بیان خواهیم نمود.

۴ـ ۶ـ ۱ـ جداسازی هر یک المانهای موجود در قطعه کلی

همانگونه که در شکل ۴ـ ۱۷ مشاهده شد اغلب قطعات کلی دارای بیش از یک المان در درون خود می باشند و نیز در بسیاری موارد خود المانهای داخل یک قطعه کلی خود به چند المان قابل تفکیک می باشند. شکل ۴ـ ۱۸ نمونه یک قطعه کلی و اجزای تشکیل دهنده را نمایش می دهد.



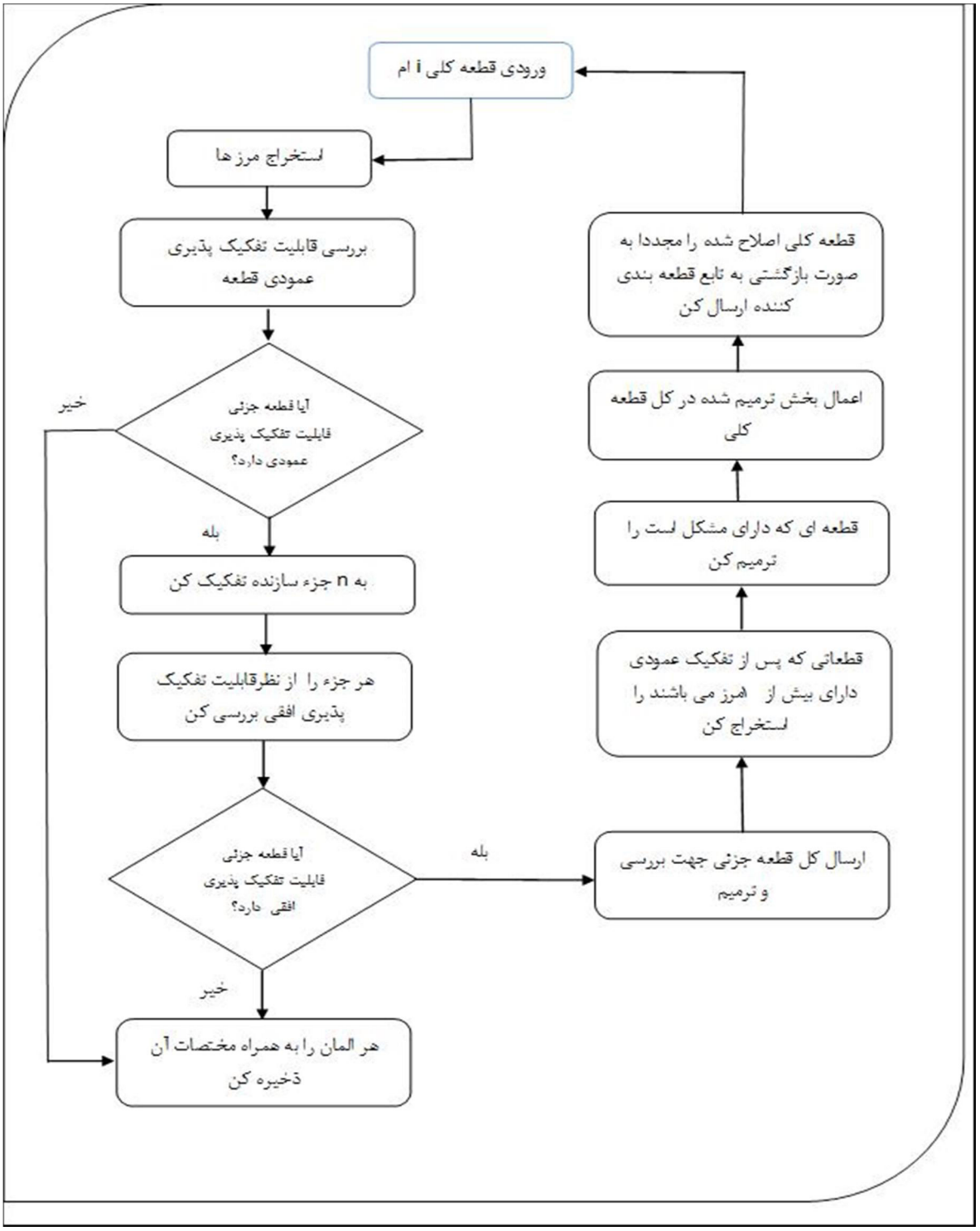
شکل ۴ـ ۱۸ـ نمونه یک قطعه کلی و اجزای تشکیل دهنده هر یک

همانگونه که در شکل ۴ـ ۱۸ مشاهده می نمائیم, یک قطعه پس از یک بار تفکیک به المانهای سازنده اش، در سطح دوم نیز قابل تفکیک بوده که بایستی به نحوی از یکدیگر تفکیک گردند.

به دلیل وجود خطوط میزان فاصله های هر یک از المانها نیز در هر بخش متفاوت بوده و گاه این فاصله بسیار زیاد و گاه بسیار کم مشاهده می شود.

لذا سیستم قطعه بندی کننده علاوه بر شناسایی هر جزء از قطعه کلی، بایستی در صورت نیاز اجزائی که قابل تفکیک بوده را نیز شناسایی نموده و در هر مرحله آنها را ذخیره نماید.

الگوریتم کلی قطعه بندی جزئی به شرح زیر می باشد.



شکل ۴ـ ۱۹ـ الگوریتم کلی قطعه بندی جزئی المانهای موجود در قطعه های کلی

یکی از مباحث مهم در زمینه پردازش تصاویر دیجیتال، کاربرد الگوریتم های مبتنی بر ریخت شناسایی[1] می باشد. یکی از کاربردهای اصلی ریخت شناسی، استخراج اجزای تصاویر است که در نمایش و توصیف اشکال مفید می باشند. به طور خاص الگوریتم های مبتنی بر ریخت شناسی جهت استخراج مرزها، قطعات متصل استفاده می گردند.

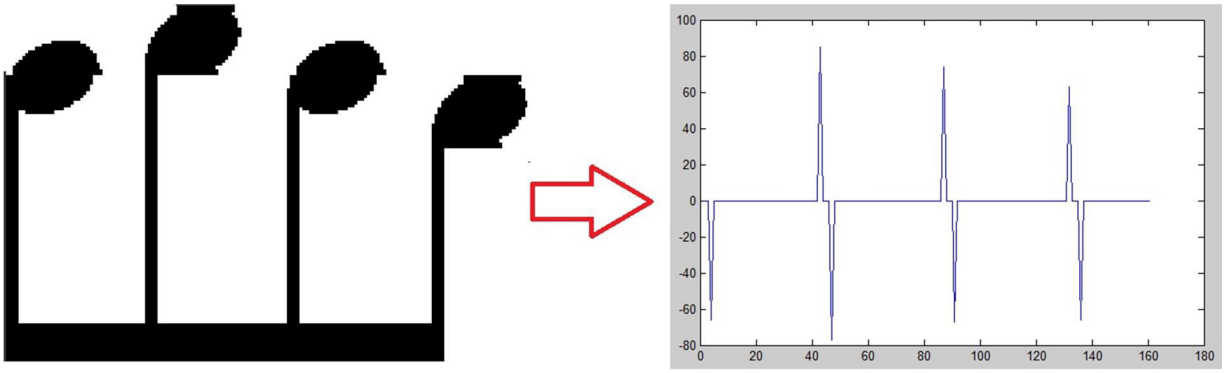
تابع bwboundaries () در نرم افزار متلب قابلیت استخراج مرزها در تصویر را داراست. خروجی این تابع برداری است $n \times 2$ که حاوی مختصات نقاط مرزی تمامی اشکال متصل در تصویر می باشد. پس از اعمال این تابع روی قطعات کلی تصویر، اشکال متصل به صورت آرایه سلولی ذخیره می گردند.

۴ـ ۶ـ ۲ـ بررسی قابلیت تفکیک عمودی هر قطعه جزئی

پس از استخراج مرزهای قطعات, هر زیر قطعه را جهت بررسی اینکه آیا از نظر عمودی قابل تفکیک است به تابعی که برای این منظور طراحی کرده ایم ارسال می نمائیم. اساس کار این تابع اولاً محاسبه تعداد المانهای موجود در قطعه و ثانیاً در صورت وجود بیش از یک المان محاسبه نقطه مناسب انفصال در هر قطعه جزئی است.

برای این منظور ابتدا مجموع نقاط قطعه را به صورت ستونی ذخیره نموده و از روی آن مشتق رتبه اول را محاسبه و ذخیره می نمائیم.

[1] Morphogical Image processing

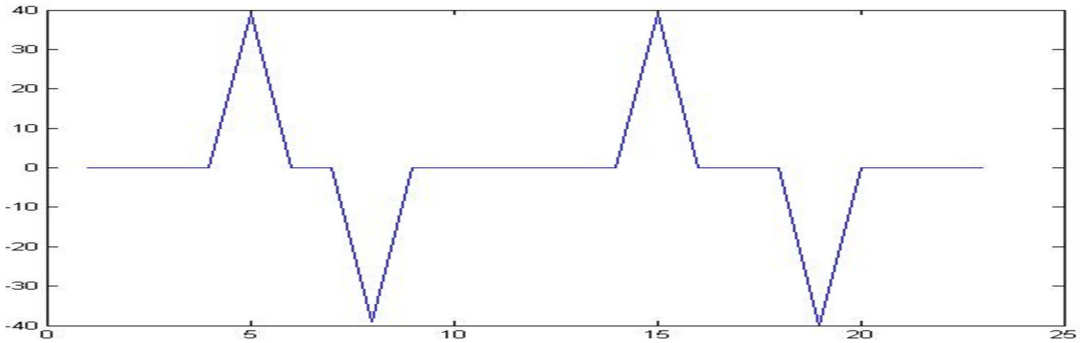


شکل ۴ـ ۲۰ـ قطعه استخراج شده و نمودار محاسبه مشتق مرتبه اول

نکته بسیار مهم در اینجا نوع قرارگیری المانها از حیث جهت می باشد تشخیص اینکه از چه سمتی بایستی نقاط انفصال را یافت نیاز به دقت مضاعف دارد. از این رو با بررسی ماتریس حاصل از مشتق گیری، ابتدا بررسی می نمائیم تغییرات از کدام سمت شروع شده است.

لذا فاصله تغییرات محسوس در هر سمت نمودار بررسی می شود و عملیات از سمتی آغاز می گردد که فاصله تغییرات مشاهده شده در آن قسمت از نقاط ابتدایی یا انتهایی بیشترین فاصله باشد.

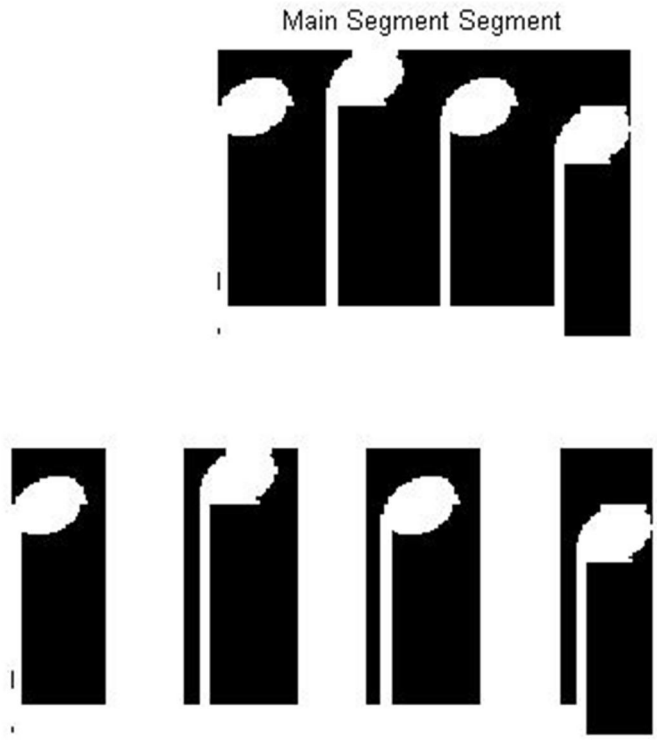
به دلیل وجود برخی علائم در زبان موسیقی گاه ممکن است این الگوریتم به اشتباه آن علائم را تفکیک نمایید. علائمی مانند دِیز #. نمودار حاصل از مشتق علائم مذکور به صورت زیر می باشد.



شکل ۴ـ ۲۱ـ نمودار حاصل از مشتق مرتبه اول روی علامت دیز

با بررسی روی نمودار حاصل از مشتق بروی این المان و مشتق روی نت های موسیقی درمی یابیم به دلیل تقارن سمت چپ و راست نمودار می توان این نکته را به عنوان عامل کنترل کننده جهت جلوگیری از تفکیک اشتباه استفاده کرد زیرا تنها در مواردی قطعه جزئی قابل تفکیک است که در آن نت وجود داشته باشد و نه علامت دیگری.

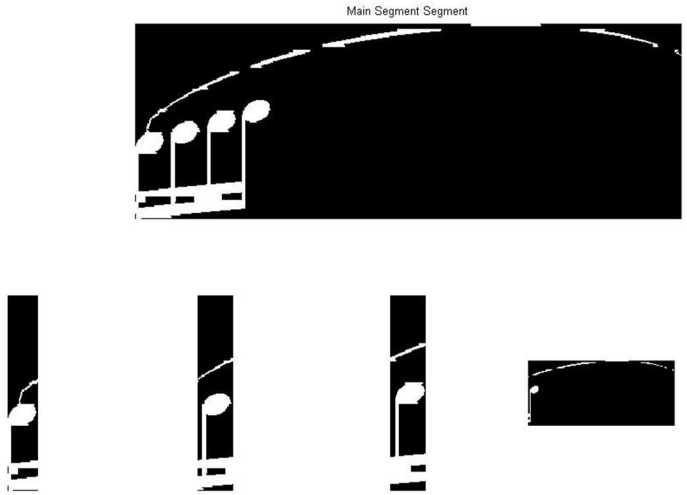
همچنین با توجه به فواصل گوناگون در هر بخش از قطعه فاصله ای که برای محاسبه نقطه انفصال در نظر گرفته ایم $\frac{1}{4}$ فاصله هر پیک در نمودار حاصل از مشتق می باشد. شکل ۴ـ ۲۲ نمونه قطعه کلی و تصاویر مربوط به تفکیک هر المان را نمایش می دهد.



شکل ۴ـ ۲۲ـ نمونه قطعه کلی و المانهای تفکیک شده

۴ـ ۶ـ ۳ـ بررسی تفکیک پذیری افقی هر زیر قطعه

پس از آنکه قطعه از منظر تفکیک پذیری عمودی مورد بررسی قرار گرفت بایستی از نظر افقی نیز مورد ارزیابی قرار گیرد. شکل ۴ـ ۲۳ نمونه قطعه ای را نشان می دهد که پس از تفکیک نمودن عمودی المانهای آن ولی همچنان دارای مشکل می باشد.



شکل ۴ـ ۲۳ـ نمونه سگمنت که به خوبی المانهای آن تجزیه نگردیده است

همانگونه که در شکل ۴ـ ۲۳ مشاهده می شود به دلیل چسبندگی خط اتصال به سر یکی از نت ها، عملیات قطعه بندی به خوبی صورت نگرفته است.

اگر هر یک از قطعات را به طور مجزا به کمک روش استخراج مرزها مورد بررسی قرار دهیم، خواهیم دید یک یا بیش از یک زیر قطعه دارای بیش از یک جزء متصل می باشند. لذا ابتدا بایستی زیر قطعه ای که مسبب ایجاد این اختلال شده است را پیدا کرده و پس از اصلاح آن؛ مجدداً جهت قطعه بندی به تابع قطعه بندی کننده ارسال نمائیم.

۴ـ ۶ـ ۳ـ ۱ـ یافتن قطعه معیوب

روشی که برای یافتن قطعه معیوب بکار برده ایم به این صورت است که ابتدا ماتریس به تعداد زیر قطعات تشکیل داده و سپس تعداد اجزاء متصل هر قطعه را در آن ذخیره می کنیم به عنوان نمونه داده های ماتریس M برای به صورت زیر می باشد:

M=[0 , 1 , 1 , 1]

اگر مجموع داده های ذخیره شده در ماتریس M برابر صفر باشد، به این معنی است که کلیه قطعات دارای یک جزء متصل می باشند. اما در صورتی که مجموع داده های ماتریس M بیش از ۱ شد، آنگاه بایستی سلولهای ماتریس که دارای مقدار صفر می باشند را مورد بررسی قرار داد.

همانگونه که در شکل ۴ـ ۲۳ نشان داده شده است قطعه ای که دارای بیش از ۱ جزء متصل باشد قطعاً با اندازه های مختلف قابل تمیز می باشند لذا جهت یافتن اینکه کدام قطعه نیاز به اصلاح دارد. ابتدا بایستی به وسیله اطلاعات درون سایر قطعات, جهت قطعه معیوب را محاسبه نمود. برای این امر اتصال اجزاء کوچک قطعات را از نظر اینکه با چه میزان عرضی در قطعه قرار گرفته اند را مورد بررسی قرار دهیم.

۳ حالت کلی برای اجزاء کوچک می توان در نظر گرفت:

۱ـ امتداد از سمت راست قطعه به چپ (کل عرض قطعه را دربرگرفته است).

۲ـ امتداد از چپ به قسمت های میانی قطعه.

۳ـ امتداد از راست به قسمت های میانی قطعه.

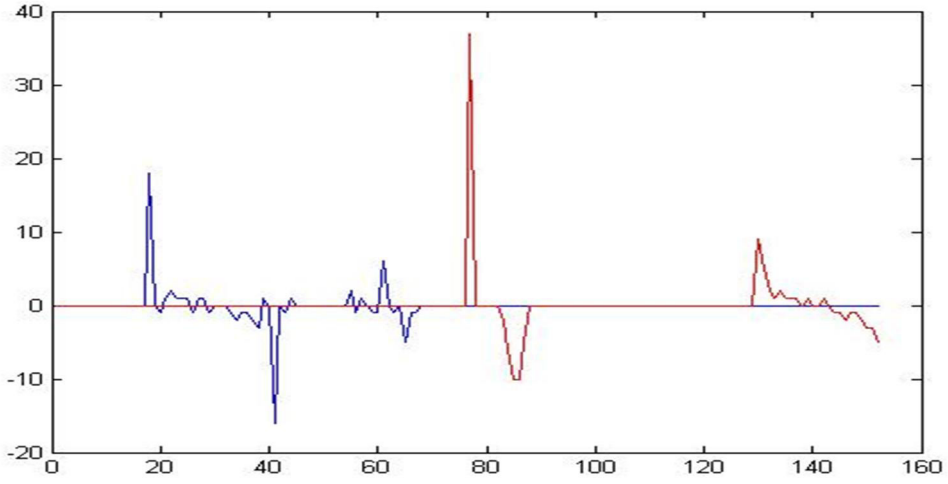
برای بدست آوردن این مقادیر کافیست ستون های اول و آخر هر قطعه را صرفاً در محدوده جزء کوچک مورد بررسی قرار دهیم.

در این روش به دنبال قطعه تک جزئی می باشیم که از سمت قطعه کناری دارای اتصال باشد. پس از یافتن آن اقدام به اصلاح می نمائیم.

۴ـ ۶ـ ۳ـ ۲ـ اصلاح قطعه دارای چسبندگی:

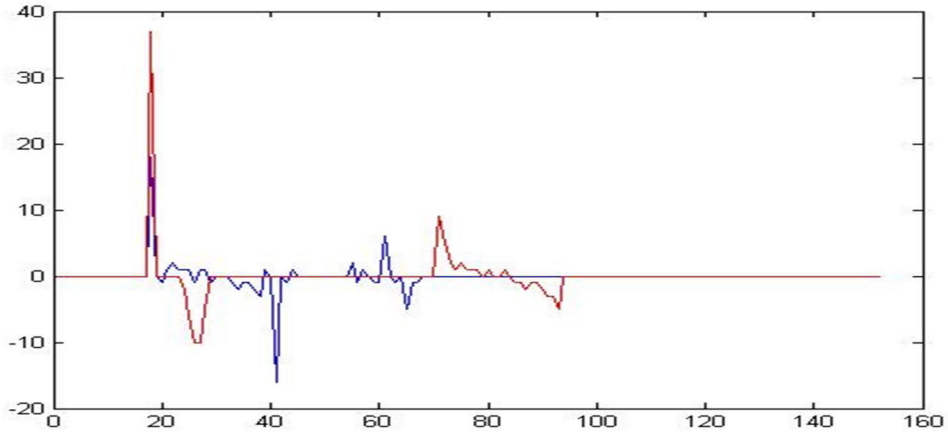
از آنجا که اغلب مشکلات ناشی از چسبندگی المانهای یک قطعه به دلیل وجود خطوط اتحاد و اتصال می باشد و این خطوط صرفاً به سر نت ها می چسبند از روش اندازه گیری میزان شباهت استفاده نموده ایم. شرح کامل روش معیار شباهت به تفصیل در بخش ۳ـ ۴ـ ۵ـ ۲ بیان شده است.

به این صورت که ابتدا سیگنال مشتق مرتبه اول روی مجموع قطعات را محاسبه و جهت بررسی میزان شباهت به تابع مربوط ارسال می نمائیم.



شکل ۴ـ ۲۴ـ نمونه دو سیگنال حامل مشتق گیری روی مجموع قطعات

تقریباً اغلب معیارهای شباهت در شرایط مقیاس زمان[1] نتیجه مطلوبی ارائه نمی نمایند این روش در زمانی که خط اتحاد بین دو نت وجود دارد به خوبی عمل کرده ولی در زمانی که دلیل چسبندگی خط اتصال باشد به دلیل تفاوت در مقیاس زمان دو سیگنال، عملکرد ضعیفی از خود نشان می دهند. از این رو قبل از محاسبه معیار شباهت ابتدا به کمک الگوریتمی که براساس نقاط ماکزیمم هر سیگنال عمل می نمایددو سیگنال را تا حد زیادی بر هم منطبق کرده و سپس اقدام به محاسبه میزان شباهت آنها می نمائیم. نمونه دو سیگنال و نتیجه شیفت در زمان را در شکل ۴ـ ۲۶ مشاهده می نمائیم.



شکل ۴ـ ۲۵ سیگنال مربوط به شکل ۴-۲۴ پس از شیفت و انطباق

همانگونه که در شکل ۴-۲۵ مشاهده می شود دو سیگنال بر اساس نقاط ماگزیمم بر یکدیگر منطبق گردیده . دلیل استفاده از نقطه ماگزیمم جهت منطبق نمودن, شباهت نت های موجود در قطعه می باشد در غیر اینصورت این روش روش کارآمدی جهت حل مشکل اختلاف مقیاس زمان نمی باشد.

[1] Time Scale

در ادامه و در هر مرحله یک ردیف از قطعه دارای چسبندگی را حذف نموده و نتیجه معیار شباهت را در ماتریسی ذخیره می نمائیم. در نهایت مقدار ماکزیمم را جهت مقایسه با دیگر ردیف های حذف شده محاسبه و در نهایت ماکزیمم کل مقادیر را به عنوان نقطه انفصال انتخاب می نمائیم.

پس اصلاح نقطه چسبندگی و اصلاح آن در قطعه جزئی، قطعه را به صورت بازگشتی به تابع قطعه بندی کننده ارسال می نمائیم.

۴ـ ۶ـ ۴ـ ذخیره سازی المانهای تفکیک شده

پس از اطمینان از غیرقابل تفکیک بودن قطعات، در مرحله آخر آنها را به همراه برخی خصوصیاتش ذخیره می نمائیم. خصوصیات فوق شامل:

۱ـ شماره قطعه جزئی.

۲ـ شماره قطعه کلی

۳ـ مختصات گوشه چپ بالا $(x_1 y)$

۴ـ مختصات گوشه راست پایین $(x_2 y)$

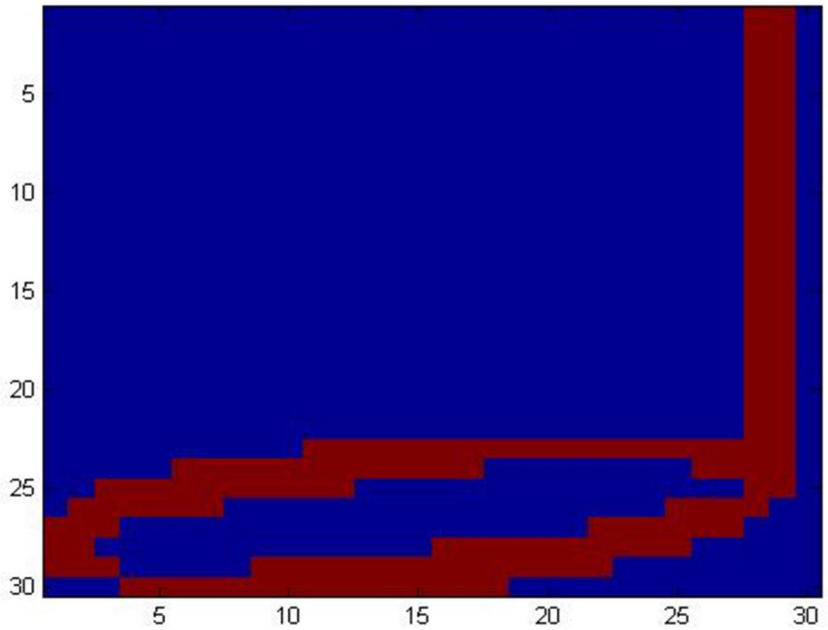
خصوصیات بیان شده در مراحل بعدی تکمیل خواهند شد.

۴ـ ۵ـ شناسایی کننده المانها

به منظور شناسایی المانها، نیاز به یک سیستم شناسایی کننده داریم. شناسایی کننده ای که در اینجا از آن استفاده می نمائیم، شناسایی کنند براساس شبکه های عصبی مصنوعی است. نوع شبکه عصبی به کار

رفته شبکه های هاپفیلد می باشد. شبکه های هاپفیلد به دلیل توانایی های بالایی که در شناسایی الگوها دارند به عنوان شناسایی کننده انتخاب گردیده اند.

ابتدا کلیه اشکال المانهای موسیقی را که به صورت یک ماتریس 30×30 است ایجاد شده است و از پیش ذخیره داریم را به کمک تابع Reshape () در نرم افزار متلب به یک ماتریس تک ستونی تبدیل می نمائیم. سپس شبکه عصبی را به کمک این ماتریس های تک ستونی آموزش داده و نتیجه را ذخیره می نمائیم.

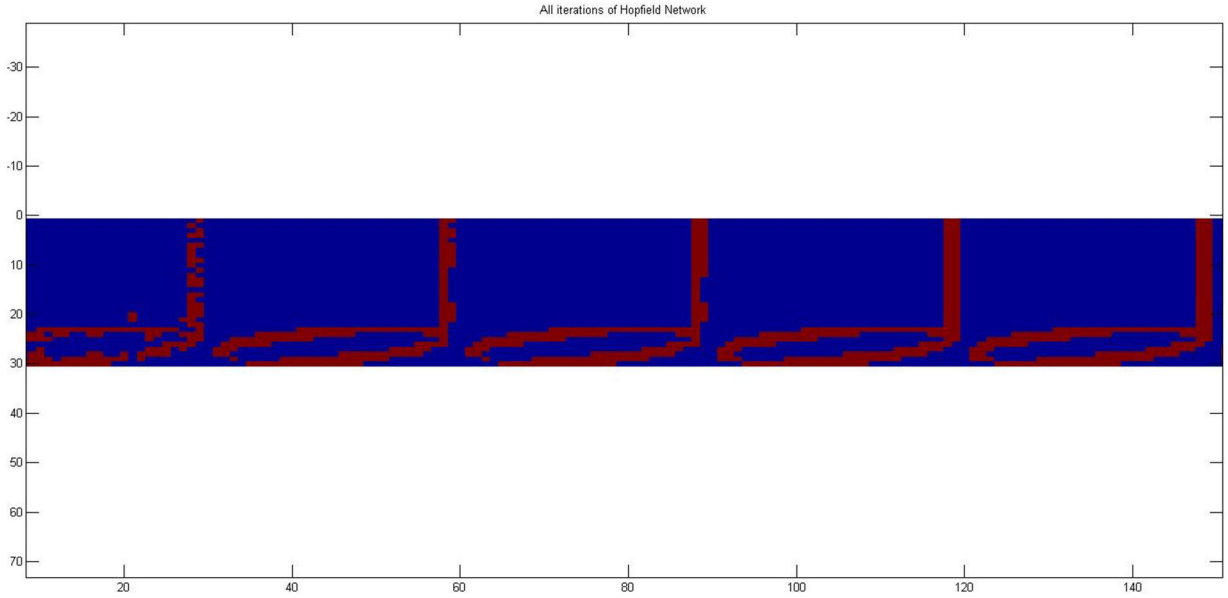


شکل۴ـ۲۶ـ نمونه یک المان ذخیره شده در ماتریس 30×30 (المان مربوط به نت سفید)

سپس المانهای حاصل از قطعه بندی را پس از reshape نمودن به عنوان ورودی، و به کمک تابع sim () به شبکه عصبی می دهیم. در نهایت حاصل شناسایی شده را به وسیله معیار فاصله همینگ[1] با الگو

[1] Hamming Distance.

ورودی مقایسه می نمایئم. در صورت منطبق بودن حاصل شناسایی شبکه عصبی با بردار ورودی, تابع همنیگ مقدار صفر را باز خواهد گرداند. شکل ۴-۲۷ نمونه ورودی را پس از تکرار ۲۰ مرحله ای نشان میدهد. همانگونه که مشاهده شبکه در همان مراحل اولیه به خوبی همگرا گردیده است.



شکل۴ـ۲۷ـ نمونه خروجی شبکه هاپفیلد(المان مربوط به نت سفید)

پس از شناسایی اقدام به ذخیره خصوصیات بیشتر از المان شناسایی شده نموده و اطلاعات آن را کامل می نمائیم. اطلاعات تکمیلی شامل نوع المان شناسایی شده به عنوان مثال المان شناسایی شده میتواند از نوع تغییر دهنده زمان اجرای نت باشد, مانند علامت نقطه و یا ممکن است المان شناسایی شده جزء المان های صدا داار باشد, مانند نت های موسیقی و سایر انواع علائم که در بخش ضمیمه به آنها اشاره شده است.

۴ـ ۶ـ اجرای موسیقی

پیش تر در بخش ۳ـ ۶ با جعبه ابزار MIDI در نرم افزار متلب مطالبی بیان شد. حال با توجه به شناسایی علائم موسیقی؛ براساس نوع علائم؛ تفسیر آنها را از حیث طنین و دیرند و فاصله، به ماتریس nmat وارد می نمائیم. ماتریس nmat حاوی اطلاعات نت، کشش زمانی نت، قدرت نواختن و فاصله زمانی که با نت بعدی ایجاد می شود، می باشد. گاه لازم است در برخی موارد ماتریس تشکیل شده در میانه تولید، بروز رسانی گردد. در مواردی که علائم تکرار شونده مشاهده می شود لازم است قسمتی از ماتریس مجدداً تکرار گردد. یا در برخی موارد که خطوط اتحاد و یا اتصال شناسایی می گردند. با توجه به تأثیر سیستم روی فاصله و کشش نت ها، لازم است به عقب برگشته و داده های قبلی را بروز رسانی نمائیم.

همچنین علائم تغییر دهنده کشش نت ها مانند «نقطه» که طبق قاعده سمت راست نت قرار می گیرند، پس از شناسایی, میزان کشش نت قبلی را مورد تغییر قرار می دهد که لازم است برای هر یک از آنها قواعدی گرامر گونه طراحی گردد.

۴ـ ۷ـ خلاصه و نتیجه گیری

در این فصل مراحل پیش پردازش که شامل اصلاح زاویه تصویر، محاسبه مکان دقیق خطوط حامل، حذف خطوط حامل، محاسبه خطوط میزان، حذف خطوط میزان، قطعه بندی کلی روی خطوط حامل و نیز قطعه بندی جزئی و بررسی اشکالات احتمالی و اصلاح آنها تشریح شد. همچنین جهت شناسایی از شبکه عصبی هاپنیله که توانایی بالایی در شناسایی الگو دارند نیز بهره بردیم و در نهایت به کمک ابزار MIDI و تشکیل ماتریس nmat اقدام به اجرای علائم شناسایی کننده کردیم.

فصل پنجم:

خلاصه و نتیجه گیری

۵ـ ۱ـ خلاصه پایان نامه

در طی فصول گذشته با پیشینه تحقیقات صورت گرفته در سیستم های شناسایی نت های موسیقی یا OMR تا حدودی آشنا شدیم. به دلیل قرابت موضوع OMR با سیستم های OCR بسیاری از روشها، خصوصاً در مبحث پیش پردازش از روشهای مذکور برداشت شده است. بسیاری از آنها با کمی تغییر قابل پیاده سازی در سیستم های OMR بوده از این رو منابعی که در خصوص سیستم های تولید شده شناسایی نوری کاراکتر کار شده بود بهره برداری شده است. همانگونه که در فصل سوم معماری کلی این سیستم معرفی شد بیشترین حساسیت و چالش در بحث پیش پردازش وجود داشت. به نظر اینجانب ضمانت اجرای صحیح سیستم های نوری شناسایی کننده در گرو طراحی بخش های مربوط به پیش پردازش کننده می باشد از این رو بیشترین زمان در طول اجرا این سیستم صرف این مقوله گردید.

در زبان موسیقی, فواصل، ضرب ها، شدت ها، همه و همه در قالب تعداد محدودی علامت قابل پیداسازی است در این پایان نامه ابتدا تمرکز روی صحت تصویر ورودی از حیث زاویه تصویر بنا نهاده شد. بدلیل اینکه خط جزء جدا ناپذیر در زبان موسیقی است وجود کجی در تصویر در همان ابتدای مراحل انجام کار، مراحل بعدی را با مشکل مواجه می ساخت از طرفی محل قرارگیری دقیق خطوط حامل نیز جهت شناسایی نت که روی این خطوط قرار می گیرند از درجه اهمیت بسیار زیادی برخوردار است.

بسیاری از پردازشهای صورت گرفته روی تصویر در واقع بر روی نسخه باینری شده تصویر پیاده سازی گردیده اند، از این رو خطای زاویه تصویر باعث ایجاد پدیده پله پله شدن در خطوط می گردد. یکی از روشهای دقیق جهت محاسبه زاویه تصویر و تصحیح آن، استفاده از روش تبدیل رادون می باشد. همچنین یکی دیگر از کاربردهای این روش پیدا خط در تصویر در هر زاویه دلخواه است به همین جهت هم در بحث تصحیح زاویه و هم یافتن خطوط از این روش بهره جستیم.

خطوط موجود در تصویر با ضخامت های متفاوت وجود داشته، لذا با توجه به لزوم قطعه بندی نیاز بود تا همه خطوط با هر ضخامتی از تصویر حذف و صرفاً آدرس آنها جهت شناسایی المانها ذخیره گردند.

همچنین یکی از مراحل پیش پردازش حذف خطوط میزان که به صورت عمود بر خطوط حامل و بین خطوط حامل ظاهر می شوند، می باشد. به دلیل شباهت خطوط میزان با دُم نت ها (شاهین) باید کنترل دقیقی روی مکانهایی که به عنوان خطوط میزان شناسایی شده اند، صورت می پذیرفت.

خطوط اتحاد و اتصال از جمله علائمی می باشند نقش بسیار مهم در زیبایی موسیقی داشته لذا در زمان حذف خطوط حامل و خطوط میزان بایستی این امر به نحوی صورت می پذیرفت تا کمترین آسیب را به خطوط اتحاد و اتصال وارد نماید از این رو در بحث حذف خطوط حامل تدابیری اندیشیده شد تا در مکان هایی که حذف صورت گرفته و فاصله دو نقطه حذف شده برابر با ضخامت خط حامل رووی آن نقطه می باشد عملیات ترمیم صورت پذیرد.

یکی دیگر از مسائل چالش برانگیز قطعه بندی جزئی از قطعه های کلی می باشد. تا زمانی که قطعات به اجزای تشکیل دهنده خود که همان المانهای تشکیل دهنده زبان موسیقی می باشند، تفکیک نگردند, امکان شناسایی نیز وجود نخواهد داشت. جهت تفکیک و تجزیه المانها از درون قطعات, ابتدا از روش های مبنتی بر عملیات ریخت شناسی استفاده کرده و با استخراج مرزها هر قطعه را از حیث تفکیک پذیری عمودی و سپس افقی مورد بررسی قرار دادیم.

روشهای نسبتاً کارا و مطمئن برای این منظور بکار گرفته شد، که براساس استخراج ویژگی از مشتق حاصل از مجموع داده های درون قطعه اقدام به قطعه بندی می نماید.

در برخی شرایط برخی قطعات به دلیل مشکلات چاپی با مشکل چسبندگی مواجه می گردند. این مشکل به کمک محاسبه معیار شباهت که توسط آقای دکتر حمید حسن پور و همکارشان طراحی گردیده بود

استفاده شد. چسبندگی ها اغلب در شرایطی پیش می آید که خط اتحاد یا اتصال وجود داشته باشد. معیار شباهت بکار رفته روی خط اتحاد به خوبی عمل کرده ولی روی خطوط اتصال نیاز به کمی تغییر در سیگنال وجود داشت تا مشکل اختلاف دو سیگنال در Time Scale را مرتفع سازد. تغییر ایجاد شده به صورت مطلوبی نقطه انفصال را محاسبه کرده و قطعه را اصلاح می نماید.

در ادامه هر یک از المانهای استخراج شده را به صورت آرایه ای مرتب ذخیره و به سیستم شناسایی کننده ارسال نمودیم.

سیستم شناسایی کننده استفاده شده در این پایان نامه، شناسایی کننده شبکه عصبی مصنوعی و با شبکه های هاپفیلد صورت پذیرفت. شبکه های عصبی هاپفیلد جزء شبکه های پویا دسته بندی گردیده اند زیرا عملکرد هر لحظه از شبکه به لحظه قبل وابسته است. همچنین توانایی بالای این شبکه در شناسایی الگو باعث انتخاب آن به عنوان سیستم شناسایی کننده گردید.

ماحصل هر سیستمی، خروجی است که از آن دریافت می شود. قطعه نت شناسایی شده تا زمانی که نواخته نشود ارزش زیادی نخواهد داشت. جعبه ابزاری توسط محقیق دانشگاه .... کشور فنلاند در نرم افزار متلب ایجاد گردیده است اگر چه بسیاری از توابع آن نیاز به وجود فایل پردازش شده دارد (فایلهای mex) اما با کمی تغییر در برخی از آنها امکان استفاده از آن نیز فراهم شد. اساس کار آن بر مبنای قالب های MIDI بنا شده که امکان خواندن و اجرای فایل های MIDI را فراهم می سازد.

فایلهای MIDI دارای ساختاری ماتریس گونه می باشند که اطلاعات هر نت به همراه شدت، کشش، ضرب و فاصله هر نت در آن ذخیره می گردند. توابع موجود در جعبه ابزار صداهایی با فرکانس مطابق با ساز پیانو تولید می نمایند و هر ۸ اکتاو موسیقی را پشتیبانی می نمایند.

بایستی در نظر داشت که اگر چه از نظر کیفیت اجرا نرم افزار متلب و این جعبه ابزار کیفیت مطلوبی ندارد اما فایل های MIDI ساخته شده را می توان به عنوان ورودی به سازهایی که Master می باشند ارسال نمود و خروجی با طنین دلنشین (صدایی به مراتب نزدیک به خروجی هر نوع ساز)از آن دریافت کرد. این به معنی اجرای ناصحیح نت توسط توابع موجود نمی باشد زیرا توابع اجرای هر نت را طبق قواعد موجود اجرا می نمایند.

۵ـ ۲ـ پیشنهاد ادامه کار

در این پایان نامه کلیه نت هایی که تولید صدا می نمایند شناسایی و اجرا گردیده اند غیر از نت ها, سایر علائم نیز وجود دارد که نقش مهمی در اجرا داشته و بسیاری از آنها خاصیت عمومی داشته و در تمامی کشورها با فرهنگ های موسیقی مختلف کاربرد دارند.

اما برخی علائم تغییر دهنده نیز وجود دارند که در برخی فرهنگها وجود داشته و یا در برخی آلات موسیقی معنی پیدا می نمایند که شناسایی آن باعث کامل شدن هر چه بیشتر سیستم شناسایی کننده خواهد شد.

همچنین در بسیاری قطعات موسیقی لزوم استفاده از ۲ دسته از خطوط حامل جهت اجرا لازم است. مانند ساز پیانو و سنتور.

قطعات مربوط به پیانو دارای دو خط حامل مرتبط به هم می باشند که با کلیدهای متفاوت نواخته می شود. هر دو دسته از خطوط حامل به کمک علامت کروشه «}» با یکدیگر مرتبط می گردند البته در بسیاری از قطعات علامت } وجود نداشته و نوازنده با توجه به سازی که نت برای آن نوشته شده و تجربه خود اقدام به نواختن هم زمان آن می نماید. قاعدتاً قطعه ای که با هر دو دست نواخته می شود دارای

صداهای هم پوشان می باشد که زیبایی اجرا را بیشتر می نماید. از این رو به علاقه مندان به سیستم های OMR پیشنهاد می شود قطعات موسیقی از این دست را هم مورد بررسی قرار داده تا زیبایی بیتشری به خروجی سیستم داده شود.

یکی از نکات بسیار مهم در زمان اجرا, شناخت ضرب در هر میزان است که معرفی آن به صورت عددی کسری در ابتدای قطعه آورده می شود. ضرب های ۲ ضربی، ۳ ضربی، ... در موسیقی باعث زیبایی چند برابر آن خواهد شد. لذا برای اجرای آن نیاز به سیستمی است که قواعد را به صورت گرامر در خود ذخیره کرده و برای اجرا از آن بهره گیرد. لذا تمرکز روی ضرب هر قطعه نیز خود مقوله ای است جالب که در صورت علاقه می توان روی آن تحقیق نمود.

شناسایی قطعات دستنویس نیز از جمله تحقیقاتی است که می توان مورد توجه علاقه مندان به هنر موسیقی و سیستم های OMR قرار گیرد.

منابع

[1] Rafael C.Gonzalez & Rechard E. Wood, Digital Image Processing, PEARSON, 2008.

[2] H.Hasanpour & A.Darvish , “NEW SIMILARITY MEASURE FOR SIGNAL PROCESSING
APPLICTIONS”

[3] حمید حسن زاده و سکینه اسدی امیری, مفاهیم جامع پردازش تصویر دیجیتال بهمراه پیاده سازی الگوریتم ها با متلب, شاهرود: دانشگاه صنعتی شاهرود, ۱۳۹۴ .

[4] Jinal Patel, Anup Shah, Dr. Hetal Patel, "Skew Angle Detection and Correction using Radon Transform," *International Journal of Electronics, Electrical and Computational System,* vol. 4, 2015.

[5] Rafael Grompone and Jérémie Jakubowicz and Jean-Michel Morel and Gregory Randall, "On Straight Line Segment Detection," *Springer Science+Business Media,* 2008.

[6] T. C. Shehan Fernando, "ROTATION INVARIANT IMAGE REGISTRATION WITH RADON TRANSFORM," *SAITM Research Symposium on Engineering Advancements,* 2011.

[7] Toft, Peter Aundal; Sørensen, John Aasted, "The Radon Transform - Theory and Implementation," Technical University of Denmark, 1996.

[8] Carsten Høilund, "The Radon Transform," Aalborg University, 2007.

[9] فرح ترکمنی آذرو مجید زنجانی, "کاربرد تبدیل رادن در تشخیص زاویه دوران اجسام و اسناد چند زبانه," دانشکدهٔ مهندسی برق و کامپیوتر، دانشگاه صنعتی اصفهان, ۱۳۸۱.

[10] Tuomas Eerola , Petri Toiviainen, "MATLAB Tools for Music Reaserch," Department of Music, University of Jyväskylä, Finland, 2004.

[11] پ. منصوری, تئوری بنیادی موسیقی, تهران: نشر کارنامه, ۱۳۷۰ .

[12] س. م. کیا, محاسبات نرم در MATLAB, تهران: انتشارات دانشگاهی کیان, ۱۳۹۳ .

[13] ا. م. ف. پ. فکری, "مقایسه و بررسی برخی روشهای یادگیری در دسته بندی حروف دستنویس گسسته فارسی in ",

همایش ملی مهندسی رایانه و مدیریت فناوری اطلاعات .

[14] م. ا. ش. لادن لاله پرور, "مروری بر بازشناسی بروز خط حروف دستنویس فارسی in ",همایش ملی مهندسی برق باشگاه پژوهشگران جوانو نخبگان دانشگاه آزاد ,شیراز, ۱۳۹۵ .

[15] م. ا. م. و. ع. ا. فرد, "ارائه و مقایسه چندین روش تعبیه برچسب جهت بازیابی و بازشناسی کلمات in ",نهمین کنفرانس ماشین بینایی و پردازش تصویر ایران، دانشگاه شهید بهشتی ,تهران, ۱۳۹۴ .

[16] Pokorny, Jiri Novotny and Jaroslav, "Introduction to Optical Music Recognition : Overview and Practical Challenges," 2015.

[17] Kia NG , Alex McLean , Alan Marsden, "Big data. Music corpus analysis. Music information retrieval," 2014.

[18] F. R. ,. S.-I. B. ,. M. B. ,. D. M. ,. S. P. ,. V. A. ,. M. M. ,. I. R. DAN GÂLEA, "A REVIEW ON PRINTED MUSIC RECOGNITION SYSTEM DEVELOPED IN INSTITUTE OF COMPUTER SCIENCE IASI," 2014.

[19] DAVID BAINBRIDGE and TIM BELL, "The Challenge of Optical Music Recognition," *Computers and the Humanities,* p. 28, 2001.

[20] I. F. ,. F. P. ,. R. M. ,. C. G. ,. J. S. C. Ana Rebelo, "Optical Music Recognition - State-of-the-Art and Open Issues," *Int J Multimed Info Ret,* p. 22, 2012.

[21] R. G. U. D. A. a. S. P. C. Prakash K Aithal, "A Fast and Novel Skew Estimation Approach using Radon Transform," *International Journal of Computer Information Systems and Industrial Management Applications,* p. 8, 2013.

ضمیمه :

معرفی تئوری بنیادین موسیقی

۱- تعریف موسیقی

تعاریف زیادی از موسیقی و منابع موسیقی غربی و شرقی وجود دارد که مطابق فرهنگ‌ها و مولفه هایی برجسته یک جامعه ارائه شده است که با مطالعه آن‌ها می‌توان به این نتیجه رسید که تعریف پایه موسیقی از دید تقریباً تمامی فرهنگها و جوامع دارای نقاط مشترک بسیاری است. در بین تعاریف بسیاری که از موسیقی توسط موسیقیدانان وجود دارد، تعاریف جالب توجهی نیز به چشم می‌خورد که توسط فلاسفه و دانشمندن و اندیشمندان قرون گذشته ارائه گردیده است که البته اکثر تعاریف موجب ارتباط موسیقی با علوم, به خصوص علوم طبیعی و ریاضی است.

در اینجا با توجه به عنوان فصل به تعریف عمومی در رابطه با موسیقی می‌پردازیم.

مجموعه‌ای به هم تنیده از صداهای خوشایند و ناخوشایند که با نظمی خاص شنیده شدند، موسیقی نام دارد. با توجه به تعریف بالا نیاز است تا تعریفی نیز از صدا، صداهای خوشایند و صداهای ناخوشاینده داشته باشیم.

۱-۱-صدا

صدا بر اثر لرزش هوا بوجود می‌آید. بنابراین هر وسیله یا عاملی که بتواند با معیاری خاص هوا را مرتعش کند، یک تولیدکننده‌ی صدا به حساب می‌آید.

هر صدایی دارای لرزش هایی است که بر حسب تعداد لرزش در ثانیه شناخته می‌شود که به آن بسامد گفته می‌شود.

۱-۱-۱- بسامد[1]

تعداد لرزش های هر صدایی در یک ثانیه را، بسامد می‌گویند. بسامدهای قابل دریافت با گوش انسان از ۲۰ بار در ثانیه آغاز می‌شود و تا ۲۰۰۰۰ بار در ثانیه هم می‌رسد. البته این محدوده مربوط به یک گوش کاملاً سالم می‌باشد. واحد شمارش بسامد به احترام هرتز دانشمند آلمانی با نام وی بیان می‌شود و با حروف اختصاری Hz نشان می‌دهند. اگر بسامد صدایی پایین باشد، می‌گوییم آن صدا بم و اگر بسامدش بالا باشد، می‌گوییم آن صدا زیر است.

۱-۲- صداهای خوشایند

اگر صدایی دارای ارتعاش منظم و قانونمندی باشد، آن صدا جزو اصوات خوشایند به حساب می‌آید و به همین دلیل برای گوش انسان جذاب است. نظیر این صداها را در طبیعت و همچنین زندگی روزمره فراوان داریم. صدای پرندگان خوش‌آواز، سازهای موسیقی و از همه مهم تر صدای آواز انسان نمونه‌های عالی از اصوات خوشایند هستند.

۱-۳- صداهای ناخوشایند

چنانچه لرزش و ارتعاشات یک صدا بی‌نظم و قاعده باشد، اثر آن بر گوش انسان ناخوشایند خواهد بود. نمونه چنین صداهایی، پیرامون زندگی امروزی بسیار است. نظیر صدای برهم خوردن فلزات، شکستن شیشه، صدای ارّه و چکش از جمله اصوات ناخوشایند به حساب می‌آیند.

[1] -frequency

نکته: با توجه به عنوان فصل جاری «تئوری موسیقی» تمامی تعاریف را بر اساس تعاریف بیان شده از حوزه هنر موسیقی بیان می‌نماییم و چه بسا به عنوان مثال در فصول بعدی تعریف ریاضیاتی تری از بسامد یا صدا آورده شود.

۲- نت موسیقی

نُت اصطلاحی است در موسیقی، که دو معنی عمده دارد:

- نشانه‌ای است در یادداشت‌های موزیکال برای نشان دادن نسبی مدت و زمان اجرای یک صوت.
- یک صدای معین، به‌تنهایی.

کلمه نت اشاره شده در این فصل ، به تعریف دوم از نت موسیقی اشاره دارد. تجربه نشان داده است که بیشتر موسیقی‌های رایج فقط از هفت صدای مختلف تشکیل شده است. و با تکرارهای گوناگون همین هفت صدا، انواع موسیقی به وجود آمده است. به هر یک از این صداهای هفتگانه، نت می‌گویند.[1]

۳- الفبای موسیقی

برای نوشتن موسیقی همانند زبانهای رایج بشری احتیاج به نمادها و یا الفبای مخصوص داریم. به عنوان در زبان فارسی ما از ۳۲ حرف به تناوب استفاده می‌کنیم و همینطور در نگارش زبان انگلیسی از ۲۶ حرف لاتین بهره می‌گیریم. به همین ترتیب می‌توان در نوشتن موسیقی از حروف و نمادهایی

[1] - موسيقي هايي وجود دارد كه از ۶ يا ۵ نت تشكيل شده اند. ولي غالب موسيقي‌هاي رايج در همان چهارچوب ۷ نتي شكل مي گيرند.

مخصوص کمک گرفت. البته تعداد حروف مورد استفاده در زبان موسیقایی فقط هفت مورد است که به شکلهای متفاوتی مورد استفاده قرار می‌گیرند.

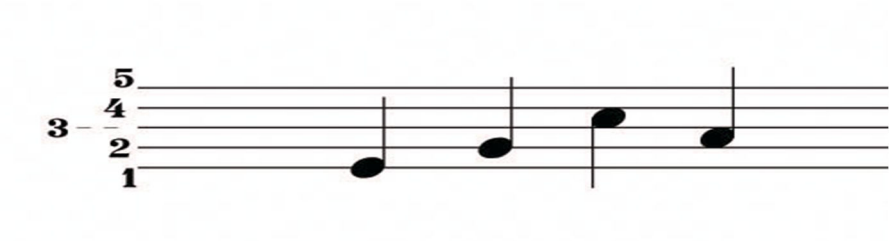
همانند حروف زبان فارسی که هر یک اسمی خاص دارند، الفبای موسیقی نیز هر یک اسم جداگانه‌ای دارند و به همین عنوان شناخته می‌شوند. اسامی نت‌های هفتگانه بدین ترتیب است:

دو ـ ر ـ می ـ فا ـ سُل ـ لا ـ سی

بر خلاف حروف زبان‌های فارسی و انگلیسی که علاوه بر اسم خاص، دارای شکل خاص و منحصر به فردی هستند، نت‌ها یا همان الفبای موسیقی فقط دارای اسم خاص‌اند و از نظر شکل و شمایل، قیافه مشترکی دارند. یعنی تمام نت های هفتگانه با یک نماد شناخته می‌شوند. حال سوالی که در این جا مطرح می‌شود این است که با توجه به این که زبان نوشتاری موسیقی دارای شکل یکسانی هستند، چگونه می‌توان نت‌های هفتگانه را از یکدیگر تشخیص داد؟

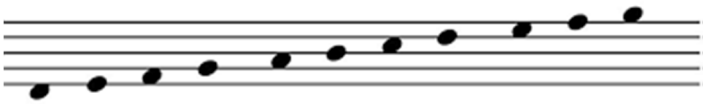
۳-۱- خطوط حامل

چنانکه می‌دانیم، در نوشتن زبانهای فارسی و انگلیسی و... همواره حروف بر روی یک خط (در زبان فارسی) و یا دو خط (در زبان انگلیسی) پشت سر هم نوشته می‌شوند اما نت‌های موسیقی هم شکلند و اگر بخواهیم روی یک خط نت نویسی کنیم، آنگاه نمی‌توان آنها را از هم تشخیص داد. به همین جهت خط مورد استفاده در موسیقی از پنج خط موازی و با فاصله‌ی کم تشکیل می‌شود که به خطوط حامل معروف هستند. خطوط حامل به ترتیب از پایین به بالا شماره‌گذاری می‌شوند بدین صورت خط اول، خط دوم، خط سوم، خط چهارم و خط پنجم.



مطابق شکل بالا ملاحظه می‌کنید که همه نت‌ها یک شکل هستند درحالیکه هیچ‌یک از آنها هم‌نام نیستند و هریک شخصیت جداگانه‌ای دارند. این تشخیص به خاطر مکان قرار گرفتن نت‌ها در رو، زیر یا بالای خطوط حامل است.

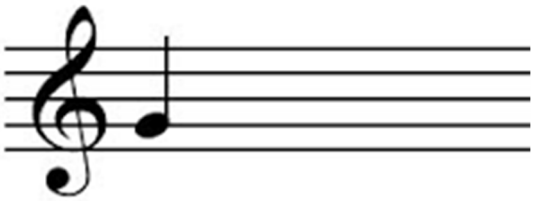
بنابراین چهار نت شکل بالا با اینکه قیافه‌ی یکسانی دارند ولی هریک نماینده‌ی یک صدای منحصر به فرد هستند، به عبارتی ما با چهار نت مختلف روبرو هستیم. در خطوط حامل علاوه بر اینکه روی پنج خط می‌توانیم نت بنویسیم، در چهار مکان دیگر هم در لابلای خطوط می‌توان نت نوشت. همچنین دو مکان دیگر یکی در پایین خط اول و دیگری در بالای خط پنجم بدین منظور وجود دارد. پس در یک خط حامل پنج خطی، تعداد یازده نت می‌توان شناسایی کرد.



۳-۲- کلید

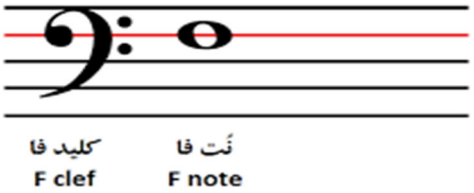
در بخش قبلی این سوال مطرح می‌شود که با وجود بالا و پایین رفتن نت‌ها در خطوط حامل، نام هر یک از آن‌ها در مکان‌های مختلف چگونه شناسایی می‌شود؟ پاسخ آن نت کلید است.

کلید نشانه‌ای است که در ابتدای حامل قرار می‌گیرد و باعث شناسایی یک نت در روی یکی از خطوط حامل می‌شود. وقتی چنین اتفاقی پیش آمد، آنگاه شناسایی سایر نت‌ها آسان است. زیرا نت‌ها با ترتیب خاصی پشت سر هم قرار می‌گیرند. در خط موسیقی کلیدهای متفاوتی به کار می‌رود. یکی از آنها کلید سُل نام دارد، زیرا معرف نام همین نت در خطوط حامل است. این کلید، همیشه نت سُل را بر روی خط دوم مشخص می‌کند و به همین دلیل به کلید سُل خط دوم نیز معروف است.



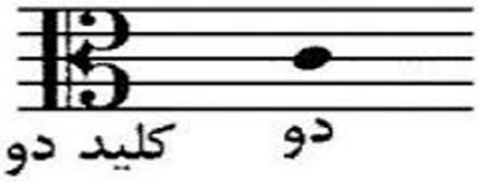
با توجه به شکل بالا اکنون می‌توانیم هر یک از نت‌ها را با نام خودش بخوانیم. لازم به ذکر است زبان موسیقی نیز مانند زبان انگلیسی از چپ به راست خوانده می‌شود. بنابراین نام نت‌های شکل بالا چنین خواهد بود: می ـ سل ـ دو ـ لا.

یکی دیگر از کلیدهای مهم و کاربردی، کلید «فا»ی خط چهارم است.



با قرار گرفتن این کلید بر روی حامل، خط چهارم به نت «فا» اختصاص می‌یابد و سایر نت‌ها نسبت به آن شناسایی می‌شوند. بنابراین در شکل بالا نت‌خوانی به ترتیب خواهد بود، فا ـ دو ـ سل ـ لا.

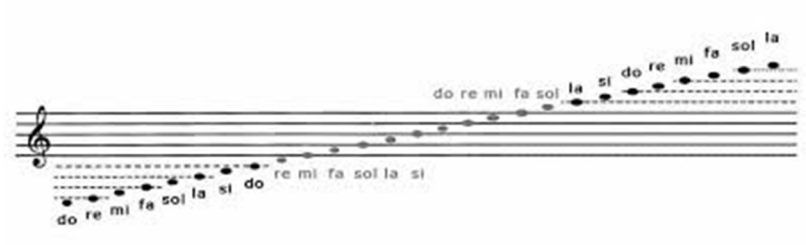
کلید دیگری هم در موسیقی ایرانی کم و بیش کاربرد دارد و آن کلید «دو»ی خط سوم است.



در ابتدای خطوط حامل، خط سوم مکان قرارگیری نت «دو» خواهد بود و نام سایر نت‌ها بر مبنای آن شناسایی می‌شوند. بنابراین نت‌خوانی در شکل بالا بدین صورت خواهد بود، دو ـ سی ـ دو ـ فا.

۳-۳- خطوط اضافه

هر نت موسیقی شخصیت منحصر به فردی دارد و آن هم فقط با بسامد آن مشخص می‌شود. یعنی وقتی نت «لا» با بسامد ۴۴۰ هرتز را در نظر بگیریم، تنها و تنها یک نت در جهان هستی وجود دارد که بسامد آن ۴۴۰ هرتز است و هرگز دو تا نیست و حتی اگر بسامدش دوبرابر شود یعنی «لا»ی ۸۸۰ هرتز و یا نصف شود مثل «لا»ی ۲۲۰ هرتز باز هم در هنگام شنیدن ملاحظه می‌کنیم با وجود شباهت ظاهری آنها، تفاوت قابل تشخیصی وجود دارد. بنابراین با قرار گرفتن نت‌ها بر روی خطوط حامل، در عمل، بسامد آنها نیز مشخص می‌شود، با این توضیح حال می‌توانیم درک خوبی از چرایی وجود کلیدهای مختلف بر روی خط حامل داشته باشیم. تعداد اصوات موسیقایی قابل شنیدن با گوش انسان بسیار بیشتر از ظرفیت یک خط حامل پنج تایی است. یعنی اگر تمام نت‌ها را بخواهیم در روی فقط یک حامل و آن هم با یک کلید بنویسیم تعداد بسیار زیادی از نت‌ها، بالاتر و یا پایین‌تر از خطوط حامل قرار می گیرند. در چنین حالتی برای این که بتوان نت های خارج از پنج خاط حامل را شناسایی کرد، لازم است از خطوط اضافه کمک بگیریم. تا زمانی که خطوط اضافه قابل شمارش باشند، مفید هستند ولی بدیهی است خواندن و نوشتن نت‌های خیلی دور از خطوط حامل که خط های اضافی زیادی دارند، دشوار خواهد بود.



چنانکه در شکل بالا می‌بینید نت‌های بیرون از خطوط حامل با خط های اضافی همراه اند و هرچه نت از حامل دورتر می‌شود، خطوط اضافه بیشتری به آن تعلق می گیرد، به طوری که خواندن آنها گاهی اوقات بسیار وقت‌گیر و آزاردهنده است.

به همین دلیل و به جهت سهولت نت‌خوانی و نت‌نویسی، کلید لازم برای حامل را با توجه به گستره صدایی هر آهنگ انتخاب می‌کنند. صداهای زیر یا بسامد بالا معمولاًبا کلید سل خط دوم نوشته می‌شوند، برای مثال نت‌نویسی صدای سازهای نی، کمانچه، سرنا و بخشی از صدای سنتور، تار و سه‌تار با این کلید نوشته می‌شود. همچنین آوانویسی صدای آوازی زنان و برخی مردان نیز با استفاده از همین کلید صورت می‌گیرد. برای نت‌نویسی صداهای متوسط از کلید «دو»ی خط سوم استفاده می‌کنند. بخش زیادی از صداهای تار و سه‌تار، کمانچه آلتو، رباب و کرنا قابل نگارش با این کلید است. صدای آوازی بیشتر مردان و برخی از زنان بم‌خوان نیز با این کلید نوشته می‌شود و بالاخره صداهای بم را با کلید «فا»ی خط چهارم می‌نویسند. این کلید صدای سازهایی چون بربط (عود)، تارباس، قیچک باس، بخشی از صداهای قانون و همینطور صدای آوازی مردان بم‌خوان را پوشش می‌دهد.

۴- زمان در موسیقی

در بخش‌های قبلی با مفهوم نت آشنا شدیم و دانستیم که هر قطعه موسیقی در اثر حرکت زیر و بمی نت‌ها بوجود می‌آید. اما همان گونه که می‌دانیم زمانی که به یک قطعه موسیقی گوش می‌دهیم،

می‌بینیم در یک قطعه موسیقی علاوه بر اجزای نت‌ها با صداهای مختلف، زمان کشش‌های نت‌ها نیز تاثیر بسزایی در زیبایی موسیقی دارد به عبارت دیگر زمان بین اجرای نت نخست با نت‌های بعدی است که به قطعه موسیقی زیبایی می‌بخشد.

بی گمان اگر در اجرای نت‌های مختلف معیار زمان نباشد، هیچ گاه موسیقی تولید نمی شود به عنوان مثال فرض کنید نت «سل» را توسط یک ساز بادی نظیر سرنا یا ترومپت به مدت یک دقیقه و بدون وقفه بنوازیم در این مدت چه خواهیم شنید؟ بطور قطع آنچه خواهید شنید یک قطعه موسیقی نیست، بلکه بوق ممتدی نظیر آژیر خطر است. با این مثال تا حدودی به ارزش زمان در موسیقی پی می‌برید. بنابراین لازم است برای تولید یک قطعه موسیقی نت‌های مختلف با زمان های کوتاه و بلند از پی هم قرار بگیرند تا بتوان یک آهنگ قابل درک و خوشایندی ایجاد گردد.

در خط موسیقی، نشانه‌ای که به یاری آن دیرند یا ارزش زمانی صداها مشخص می‌شود، در شکل نت نمودار می‌گردد. به گفته دیگر، به کمک شکل نت است که می‌توان مقدار ارزش زمانی هر نت را نسبت به نت یا شکل دیگر بازشناخت. روشن است که هر شکل نت می‌تواند در هر جای خط حامل قرار گیرد. در واقع دیرند هر صوت موسیقی ربطی به نواک[1] آن ندارد.

با گوش کردن دقیق یک ملودی ، که سازی تنها آن را می نوازد ، یا گروهی از سازها با هم همکاری می کنند ، میتوان احساس کرد که اصوات ملودی یا لایه های آرایشی آن، گاه کوتاه تر و گاه کشیده تر، و نیز در لحظه ای زیر و در لحظه ای دیگر بم هستند.

نت های موسیقی می توانند به هفت شکل و احیانا کمی بیشتر ، ظاهر شوند در جدول پایین شکل های هفتگانه و دیرند آنها نسبت به یکدیگر نشان داده شده است:

[1] - زیر و بمي صدا

نام نت :	گرد	سفید	سیاه	چنگ	دولاچنگ	سه لاچنگ	چهارلاچنگ
شکل نت :							
نسبت زمانی هر شکل :	۱	۱ / ۲	۱ / ۴	۱ / ۸	۱ / ۱۶	۱ / ۳۲	۱ / ۶۴
ارزش زمانی :	بزرگترین کشش	نصف گرد	نصف سفید	نصف سیاه	نصف چنگ	نصف دولاچنگ	نصف سه لاچنگ

همان گونه که در جدول بالا مشاهده می‌کنیم، شکل نت‌ها و نسبت زمانی هر شکل نسبت به نت گرد آماده است. شکل‌های دولاچنگ، سه لاچنگ، و چهارلاچنگ نیز همان رابطه‌ای را از نظر نظر ارزش زمانی با نت گرد دارند که نت‌های قبلی دارند. به عنوان مثال ارزش زمانی شکل دولاچنگ علاوه به نسبت یک دوم با چنگ، یک چهارم سیاه، یک هشتم سفید، و یک شانزدهم گرد را دارا است.

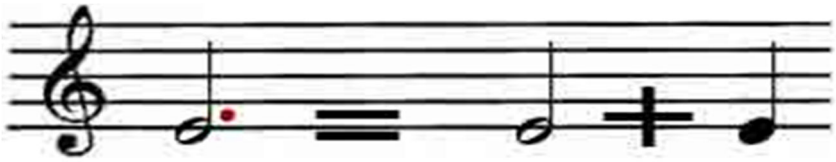
هر گاه شکل‌های چنگ، دولاچنگ، سه لاچنگ، و چهارلاچنگ (نت هایی که دارای چنگک هستند) در جایی از قطعه موسیقی بیش از یکی و پی در پی آمده باشند. می‌توان آن‌ها را به صورت گروهی نوشت، یعنی چنگکشان را حذف کرد و انتهای دمهایشان را با خطی (یا خطهایی) به هم پیوند داد .

در این حالت هر چنگک تبدیل به یک خط می‌شود.

در جدول صفحه قبل، همه نت‌ها به جز گرد دارای دم هستند و از آنجا که همه این شکل‌ها به منظور تشخیص زیر و بمی شان روی خط حامل نوشته می‌شوند، باید گفت که معمولاً هر گاه جای نت (یا جای سر نت) پایین تر از خط سوم حامل باشد، دمش رو به بالا، چنانچه سر نت در نیمه بالای حامل نوشته شود، دم آن به پایین کشیده می شود. در حالت اول، دم از سوی راست سر نت به بالا و در حالت دوم از کناره ی چپ سر به پایین و در هر دو حالت دم مماس با سر رسم می شود.

۴-۱- نقطه ونقش آن

نت گرد و هر یک از شکل‌های بعدی دیرند، تنها می توانند به ۸.۴.۲ و... قسمت تقسیم شوند، در حالی که گاه لازم است که یک ارزش زمانی به ۳ (یا مضرب هایی از ۳) قسمت تقسیم گردد. یکی از نشانه هایی که تقسیم‌های سه تایی را میسر می‌سازد، نقطه است که سمت راست هر شکل نت می‌گیرد به عبارت دیگر، هر گاه سمت راست نت، نقطه ای گذاشته شود، نصف بر دیرند آن نت می‌افزاید. به عنوان مثال نت گرد نقطه دار برابر با سه نت سفید است. با یک محاسبه ساده می‌توان فهمید که نت نقطه دار می‌تواند به ۳ و هم به ۲ قابل تقسیم باشد.



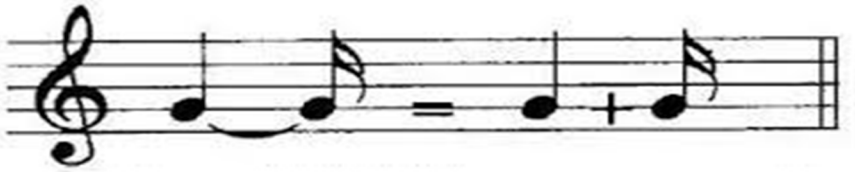
اگر سمت راست شکل نت ۲ نقطه بگذاریم، نقطه اول (نقطه نزدیکتر به نت) طبق قانون، نیم برابر به ارزش نت می‌افزاید و نقطه دوم نیم برابر ارزش نقطه اول (یا یک چهارم نت اصلی) را در مجموع افزایش می‌دهد. اگر ارزش زمانی نت اصلی را یک بگیریم، با دو نقطه سمت راست آن، مجموع ارزش برابر است با $۱\frac{۳}{۴}$ نت اصلی:

$$۱+\frac{۱}{۲}+\frac{۱}{۴}=(\frac{۴}{۴}+\frac{۲}{۴}+\frac{۱}{۴})=\frac{۷}{۴}$$

به همین ترتیب اگر به تعداد نقطه‌ها افزوده شود (که این حالت به ندرت پیش می‌آید) هر نقطه به اندازه نیمی از نقطه پیش از خود، به مجموع ارزش زمانی می‌افزاید.

۴-۲- خط اتحاد و خط اتصال

در بخش ۴-۱ دیدیم که تعداد نقطه می‌تواند از یکی بیشتر باشد، و با نگاهی دقیق دریافتیم که تعداد نقطه هر چه بیشتر شود، ارزش نت را به دو برابر نمی رساند، در حالی که به یاری خط اتحاد می‌توان ارزش نت را به هر اندازه دلخواه افزود. اگر بر روی دو نت همصدا، یا بیشتر، خطی منحنی به صورتی که در شکل زیر می‌بینیم قرار گیرد، ارزش‌های نت ها با هم جمع می‌شوند. خط اتحاد می‌تواند روی یک زوج یا بیش از یک زوج از نت‌های هم صدا قرار گیرد.



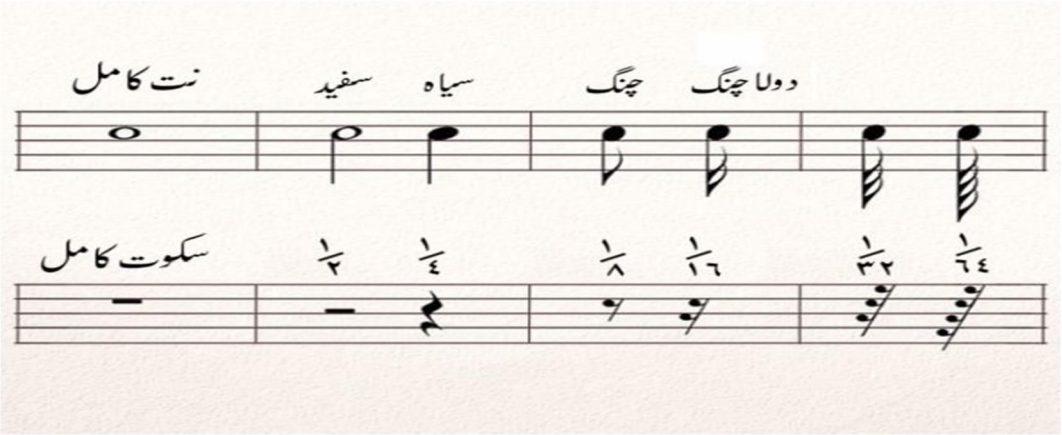
۴-۲-۲- خط اتصال

خط اتصال نشانه‌ای همشکل خط اتحاد است. با این تفاوت که نت‌های غیرهمصدا (یا همصدا و غیر هم صدا) را در بر می‌گیرد. نقش این نشانه به طور کلی این است که نت‌های در برگرفته شده، اجرایی به هم پیوسته باید داشته باشند. این نت‌ها اگر برای آواز نوشته شده باشند، سراینده، آن‌ها را با یک نفس و اگر برای ساز ویولن (یا هر ساز زهی آرشه ای) باشند، نوازنده آنها را بدون تغییر جهت آرشه وبی توقف اجرا می‌کند.

نقش خط اتصال، جز آن که در بالا گفته شد، بیشتر این است که جمله یا نیم جمله موسیقی به کمک آن مشخص می‌شود.

۴-۳- سکوت

در طی یک اجرا گاه لازم می‌شود که موسیقی لحظه‌ای ساکت شود، و به گفته دیگر، در آن لحظه (مثلاً به اندازه دیرند یک گرد، یک سیاه، یک سفید، یک چنگ و....) سکوت برقرار شود. و نیز هر گاه صداهای یک آهنگ به صورت بریده بریده و مقطع اجرا شوند، ممکن است میانشان سکوت به کار رفته باشد. نشانه‌های سکوت، عیناً مانند نت ها، دارای شکل‌ها و ارزش‌های گوناگون هستند، و به عبارتی دیگر، به ازای هر شکل نت (با دیرینه یا کشش زمانی ویژه خود) نشانه سکوتی وجود دارد.



سکوت ها نیز، مانند نت‌ها می‌توانند نقطه دار باشند، نقطه سکوت، همانند نقطه نت، در سمت راست سکوت گذاشته می‌شود. نقطه در سکوت عیناً همان نقشی را دارد که در بخش ۴-۱ به آن اشاره شد.

۴-۴- تاکید در موسیقی

موسیقی در صورتی قابل درک است که به هر حال دارای تاکید باشد. صوت‌های یک آهنگ ساده معمولاً به صورت تاکید دار و بی تاکید تقسیم می‌شود و غالباً دارای دوره‌های منظم هستند. برای درک صوت‌های تاکید دار و بی تاکید مثال تیک تاک ساعت مثالی می آوریم، تیک تاک ساعت. هر گاه با اندکی دقت به آن گوش فرا دهیم، یک دوره دو ضربی را مرتباً تکرار می‌کند

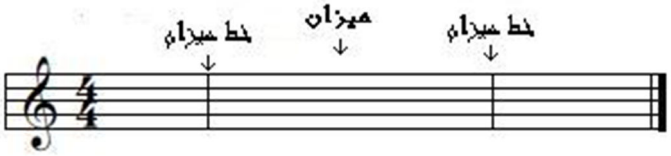
در دوره‌های دو ضربه ای، ضربه‌های اول موکد و به ضربه‌های دوم ناموکد هستند.

به همین ترتیب روال دوره های سه، چهار، پنج و شش ضربه‌ای می‌توانند در احساس انسان شکل بگیرند. در واقع مغز انسان دارای آنچنان استعدادی است که می‌تواند دوره‌های وزن را، با هر تعداد ضرب حفظ کند و ادامه دهد.

۴-۵- میزان

در موسیقی نشانه هایی برای مشخص کردن دوره‌ها و تعداد ضرب در هر دوره و تشخیص همه ی نکته های وابسته وضع شده است. هر دوره به وسیله ی خطی عمودی، که حامل را قطع می کند از دوره های پیش و پس از خود جدا می شود.

محتوای میان هردوخط عمودی را میزان و خط عمودی را خط میزان می نامند. بنابراین محتوای میزان عبارت است از تعداد ضرب‌ها در هر دوره و جای ضرب های قوی و ضعیف.



۴-۵-۱- وزن در موسیقی

وزن عبارت است از تکرار متناوب یک رشته دیرندهای گوناگون. این دیرندها می توانند از صوت (صرف نظر از نواکشان) یا سکوت تشکیل شده باشند که در حالت ساده ی خود دارای دوره ی تناوبی در فضای یک میزان و گاه بیش از یک میزان متجلی شوند.

۴-۵-۲- گونه‌های میزان

میزان بر چهار گونه است:

۱- **میزان دوتایی:** هر دوره آن متشکل است از دو ضرب (یا توان هایی از دو ضرب). در هر میزان دوتایی، (8,4,2...) شکل سفید، سیاه، چنگ و یا ... به کار می رود.

۲- میزان **سه تایی:** هر دوره آن متشکل از سه ضربه است. در هر میزان سه تایی ۳سفید،۳سیاه، ۳ چنگ ... جای می گیرد.

گونه های ۱ و ۲ را میزان ساده گویند.

۳- **میزان ترکیبی:** در واقع میزان ساده ای است که هر واحد ضربه در آن به سه قسمت تقسیم شده و این قسمت های سه گانه تا اندازه ای واحدهای ضرب پنداشته شده اند این میزان را از آن جهت ترکیبی می گویند که هم واحد اصلی هم واحد سه تایی در آن نقش دارند.

۴- **میزان پیچیده و لنگ :** در ساده ترین تعریف هر دوره ی این میزان آمیزه ای است از مضرب‌ها و مجموع های اعداد ۲ و۳.

۴-۵-۳- کسر میزان

وقتی نوع میزان‌بندی یک آهنگ مشخص شد و خواستیم آن را روی کاغذ بنویسیم، لازم است نوع میزان در همان ابتدای آهنگ نوشته شود تا نوازنده یا خواننده هنگام اجرای آن متوجه باشد که با چه نوع میزانی سر و کار دارد. این کار با استفاده از اعداد کسری صورت می‌گیرد. ولی معمولاً خط کسری بین دو عدد را

نمی‌نویسند. مثل $\frac{۲}{۴}, \frac{۳}{۴}, \frac{۴}{۴}$. به این اعداد کسر میزان می‌گویند. عدد بالایی تعداد ضرب های یک میزان را معلوم می‌کند. پس $\frac{۴}{۴}$ یعنی میزانی که دارای چهار ضرب و هر ضرب برابر با یک سیاه است.

۴-۵-۴- میزان‌های ترکیبی

اکنون صحبت از میزان‌های ساده دوضربی، سه‌ضربی و چهارضربی بود و چنان که اشاره شد. هر ضرب یک میزان ساده فقط می‌تواند به ۲، ۴ و یا ۸ قسمت مساوی تقسیم شود. همین نکته موجب تفاوت عمده میزان‌های ساده و ترکیبی است. زیرا در میزان ترکیبی هر ضرب فقط می‌تواند به ۳، ۶ و یا ۱۲ قسمت مساوی تقسیم شود. برای ترکیبی کردن میزان‌های ساده کافی است در مقابل هر ضرب یک نقطه قرار دهیم.

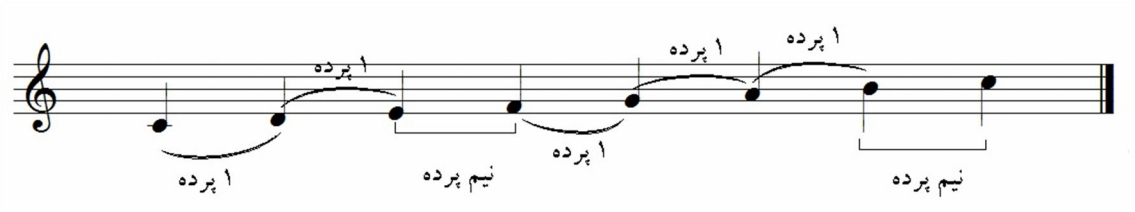
مهم ترین وجه تمایز یک میزان ساده و یک میزان ترکیبی، کشش زمانی ضرب‌هاست. یعنی در میزان ساده، هر ضرب به اندازه‌ی یک سیاه و در نوع ترکیبی، هر ضرب به اندازه‌ی یک و نیم سیاه کشش زمانی دارد. اما مهم تر از آن کیفیت شنیداری میزان‌های ترکیبی است که تفاوت بسیار محسوسی با میزان‌های ساده دارد.

۵- فاصله در موسیقی

در مباحث ابتدایی این فصل به نام نت‌های هفتگانه آشنا شدیم و دیدیم که مهمترین عامل تشخیص جدای از صدایی که گوش انسان آن را تمیز می‌دهد اختلاف بسامد بین نت‌ها است.

۵-۱- پرده و نیم پرده

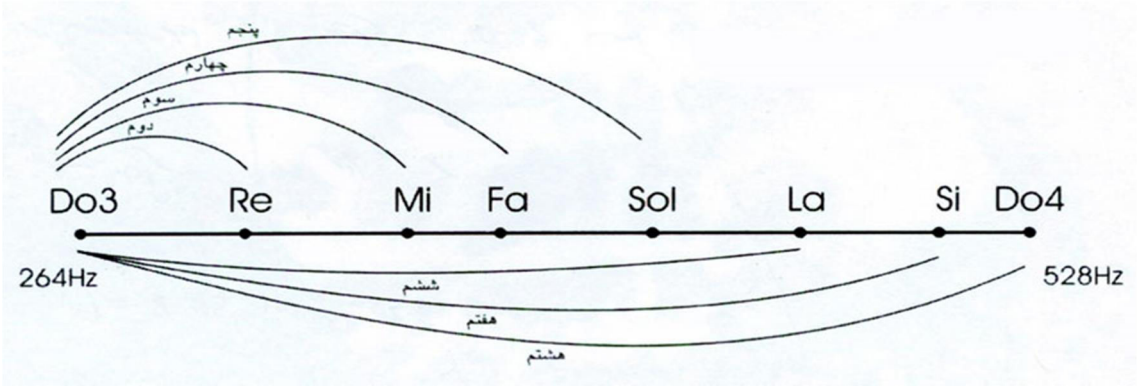
همان گونه که قبلاً نیز اشاره شد، در موسیقی بین نت‌های هفتگانه فاصله هایی از نظر بسامد وجود دارد نام فاصله‌های بلند را پرده و نام فاصله‌های کوتاه را نیم پرده گویند.



۵-۲- فاصله در موسیقی

پیش از این در بحث پرده و نیم پرده دانستیم که اختلاف بسامد در نت، سبب ایجاد فاصله می‌شود و این فاصله صرفاً بسامدی است. یک نت علاوه بر نت‌های همسایه، با نت‌های دورتر نیز فاصله دارد که لازم است همه آن‌ها را بشناسیم.

در دانش موسیقی فاصله هر نت با خودش را «یکم درست» می‌نامند. این فقط یک قرارداد است اگر فاصله‌های ایجاد شده را مطابق شکل زیر بدین صورت می‌توان نام برد:



فاصله هر نت با یک نت بعد از خودش، فاصله دوم، فاصله هر ن، دو نت بعد از خودش، فاصله سوم، و... باید توجه کرد که فاصله بین Do3 , D04 از نوع هشتم است و نباید با فاصله یکم درست اشتباه شود. زیرا نام هر دو نت یکسان است، اما بسامد آن یکی نیست. بسامد Do4= 528Hz دو برابر Do3=264Hz است یعنی نسبت ۲ به ۱.

۵-۲-۱- فاصله ساده- فاصله ترکیبی

هر گاه فاصله در نت از مرز یک اکتاو تجاوز نکند، این فاصله را ساده گویند و اگر دو نت فاصله بیرون از مرز اکتاو باشد، آن را فاصله ترکیبی می‌نامند. به عبارت روشن تر، فاصله یکم، دوم، .... هشتم، فاصله ساده هستند.

از فاصله نهم به بالا را، فاصله‌های ترکیبی می‌نامند.

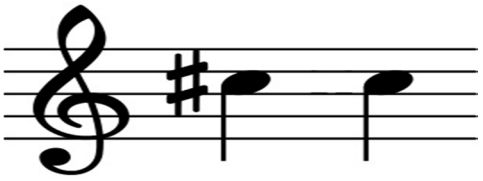
به منظور بررسی و شناخت دقیق تر، فاصله در موسیقی ابتدا لازم است ما برخی از نشانه‌های تغییر دهنده که هم در نت نویسی و نت خوانی و نیز در درک بهتر فاصله ما را یاری می‌رسانند؛ آشنا شویم.

۵-۲-۲- نشانه‌های تغییر دهنده

در موسیقی نشانه هایی وجود دارد که نقششان زیر و بم کردن نت‌ها به اندازه نیم پرده است. بدون آن که نام نت‌ها (و محل قرار گیری آن‌ها روی حامل) تغییر کند، این نشانه‌ها از زبان فرانسه گرفته شده است و عبارتند از:

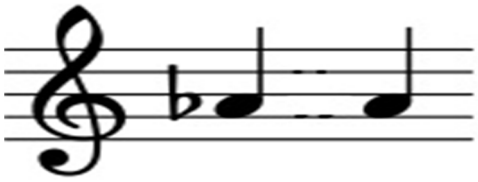
۵-۲-۲-۱- دیز (#, Diese)

هرگاه سمت چپ نتی علامت دیز قرار گیرد، آن را نیم پرده بالا می برد. این نشانه می تواند جلوی هر نت موسیقی گذاشته شود. اگر فاصله میان دو نت پی در پی یک پرده باشد و نشانه دیز سمت چپ نت اول گذاشته شود، طبیعی است که فاصله میان آن دو تبدیل به نیم پرده می شود، و هرگاه فاصله میان دو نت پیاپی نیم پرده باشد و نشانه دیز سمت چپ نت دوم قرار گیرد، فاصله آن دو تبدیل به یک پرده می شود.



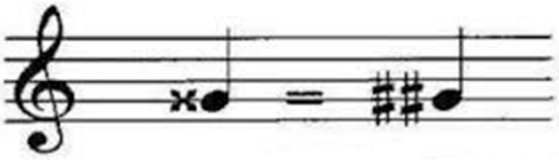
۵-۲-۲-۲- بمل (b, BEMOL)

هرگاه سمت چپ نتی قرار گیرد، آن را نیم پرده پائین می آورد، (بم می کند). این نشانه نیز، مانند دیز میتواند جلوی هر نت گذاشته شود. برای آنکه فاصله یک پرده تبدیل به نیم پرده گردد، نشانه بمل باید پیش از نت زیرتر نوشته شود و برای یک پرده ای کردن فاصله نیم پرده، این نشانه باید جلوی نت بم تر قرار گیرد. شکل زیر نکته‌های بیان شده در خصوص علامت بمل را تشریح می‌کند.



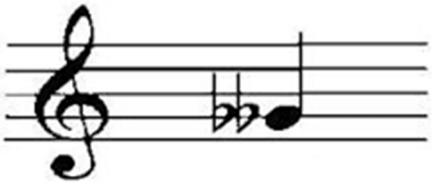
۵-۲-۲-۳- دوبل دیز (# #, X, Double Dies)

هر گاه سمت چپ نت واقع شود آن نت را دوبار و هر بار به فاصله نیم پرده بالا می‌برد به کمک این نشانه می‌توان نت «دو دیز» (یا نت دوای که یک بار دیز شده) را بار دیگر نیم پرده بالا برد و نت «دو دوبل دنیز» را به شکل پایین به دست آورد.



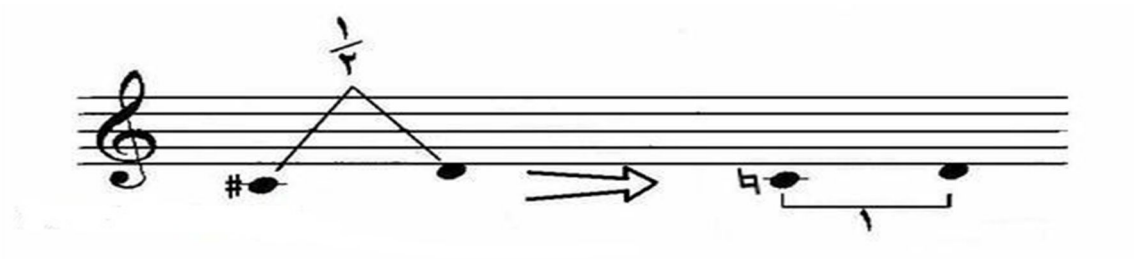
۵-۲-۲-۴- دوبل بمل (bb, Double, be mol)

هر گاه سمت چپ نتی گذاشته شود، آنت نت را دوباره به فاصله نیم پرده پایین می‌آورد. هر گاه مثلاً نت «سی بمل» (نت «سی»ای که یک بار به وسیله بمل پایین آمده) را باز بخواهیم نیم پرده پایین بیاوریم، نت سی دوبل بمل به شکل زیر به دست خواهد آمد.



۵-۲-۲-۵- بکار (Becerra)

هر گاه سمت چپ نتی که قبلاً «دیز» یا «بمل» شده گذاشته شود، اثر نشانه پیشین را خنثی می‌کند. هر گاه نت مزبور پیش تر «دوبل دیز» یا «دوبل بمل» بوده باشد، برای خنثی کردن یکی از دیزها یا یکی از بمل‌ها باید نشانه بکار را با دیز یا بمل با هم به کار گرفت. شکل زیر اثر دو حالت بیان شده را نشان می‌دهد.



۵-۲-۲-۶- سُری (Sori) و کرن (P, korn)

این دو نشانه را برای تغییرهای ویژه موسیقی ایرانی، علینقی وزیری وضع و پیشنهاد کرده است، نشانه هایی هستند که با همان روش کاربرد نشانه‌های دیگر، نت‌ها را تقریباً ربع پرده به ترتیب بالا یا پایین می‌برند. مثلاً نت دو- سری میان نت‌های دو و «دو- دیز» و نت «سی- کرن» میان نت‌های «سی» و «سی- بمل» قرار دارند.

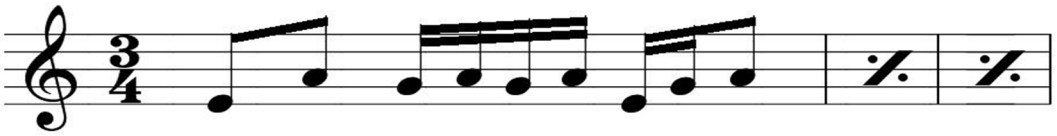
درباره این دو نشانه دو نکته را بایستی گفت:

۱- نشانه‌های «دوبل- سری» و «دوبل- بمل» وجود ندارند.

۲- برای خنثی کردن این دو نشانه، نشانه بکار همچنان موثر است.

۵-۲-۲-۷- تکرار میزان

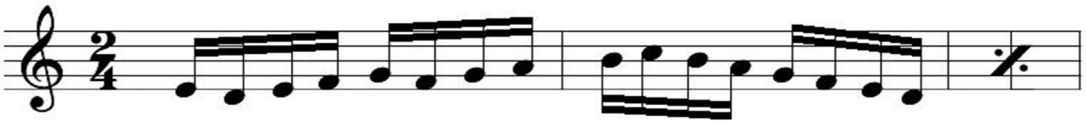
در نت نویسی برای سهولت نوشتن و خواندن، تا حد ممکن از علامت های اختصاری کمک می گیرند. برای مثال اگر در یک قطعه، محتوای دو یا چند میزان، دقیقاً یکی باشد، به جای آن که همه نت‌ها نوشته شوند، فقط نت های میزان اول نوشته می شود و در بقیه میزان های تکراری، علامت مخصوص تکرار میزان می آید.



در نتیجه کسی که قصد اجرا کردن آن قطعه را داشته باشد، با دیدن علامت تکرار میزان، برابر با تعداد علامت ها، میزان اول را اجرا می کند. شکل بالا نشانه تکرار میزان را نشان می‌دهد.

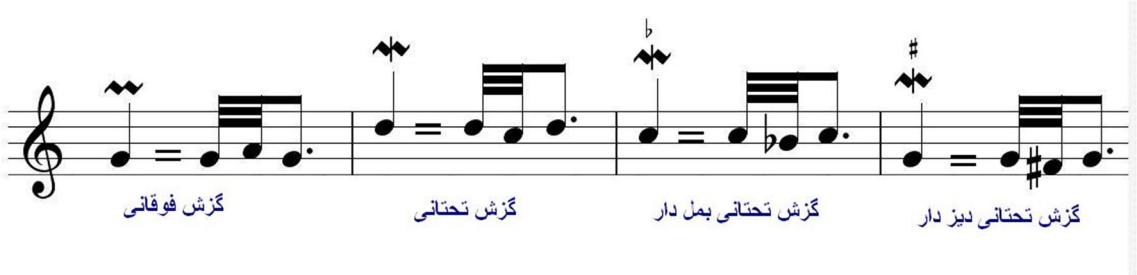
۵-۲-۲-۸- تکرار دو میزانی

اگر در قطعه ای نیاز باشد که دو میزان پشت سر هم، دوباره تکرار شوند، همان علامت مخصوص تکرار میزان، روی خط میزان قرار می گیرد. شکل زیر نشانه تکرار دو میزانی را نشان می‌دهد



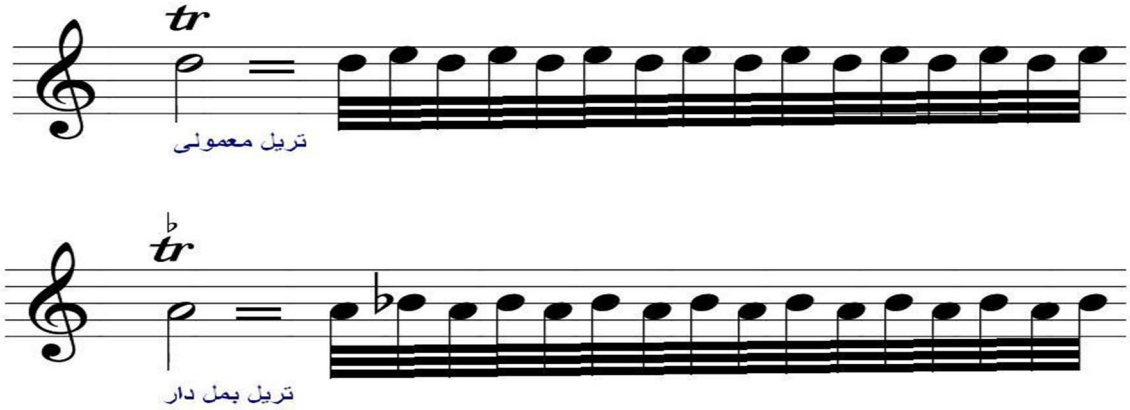
۵-۲-۲-۹- گزش

این علامت در موسیقی به دو گونه است، که وقتی روی یک نت قرار بگیرد، اشاره ای خیلی تند به نت پایین تر و یا نت بالاتر می کند و باز به نت اصلی برمی گردد. اگر اشاره به بالا باشد، گزش فوقانی و اگر به پایین باشد، گزش تحتانی نام می‌گیرد. اگر علاوه بر علامت گزش، روی آن علامت های، bیا # قرار بگیرد، به معنی آن است که نت مورد اشاره در گزش می باید بمل یا دیز باشد. شکل زیر نمونه هایی از علامت گزش را نشان می‌دهد.



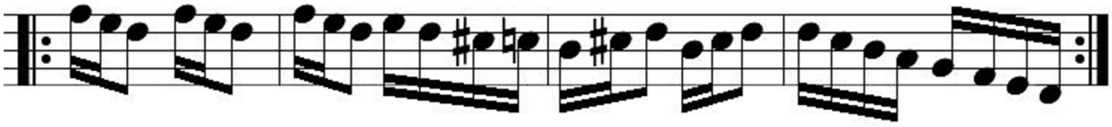
۵-۲-۲-۱۰- علامت تریل (tr)

نشانه ای است که اگر روی نتی گذاشته شود، آن نت و یک نت بالاترش پی در پی و به صورت تناوبی کم و بیش تند، به اجرا در می آیند. این عمل دقیقاً به اندازه ارزش زمانی همان نت ادامه پیدا می کند. اگر روی علامت تریل نشانه های b و # قرار بگیرد، به معنی آن است که عمل تناوب نت اصلی، می بایست با لحاظ کردن آن نشانه باشد.



۵-۲-۲-۱۱- برگشت ساده

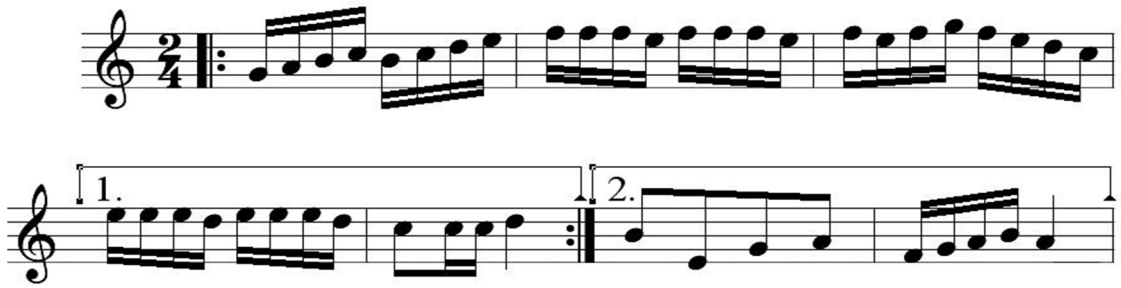
بیشتر قطعات موسیقی از تکرار جمله های گوناگون تشکیل می شوند. در نتیجه هنگام نوشتن آن ها، لازم نیست همه تکرارها نوشته شوند. بلکه یک بار جمله اصلی نوشته می شود و با استفاده ازیک علامت مخصوص، اجرا کننده متوجه می شود که این جمله باید کاملاً از نو اجرا شود. علامت برگشت ساده، یک دولاخط نقطه دار است که در ابتدا و انتهای جمله موسیقایی درج می شود شکل زیر چگونگی استفاده از این علامت را نشان می‌دهد.



۵-۲-۲-۱۲- برگشت متنوع

در این نوع برگشت، تمام جمله موسیقایی تکرار نمی شود. بلکه تنها بخشی از آن تکرار می شود و پس از تکرار، در ادامه اش یک بخش جدید می آید. در این صورت از علامت برگشت متنوع استفاده می شود که

با شماره گذاری های ۱ و ۲ به اجرا کننده می فهمانند بار اول اجرای بخش ۱ را به جمله قابل تکرار بچسباند و بار دوم از قسمت ۱ صرف نظر کند و قسمت ۲ ر ابه آن بخش تکرار شونده بچسباند.



۶-خلاصه

تا اینجا با بسیاری از اصطلاحات، اشکال و به طور کلی زبان موسیقی آشنا شدیم. در این فصل، موسیقی را با اصطلاحات خود موسیقی مطالعه نمودیم. همانگونه که برای خواندن و نوشتن هر زبان محاوره‌ای لازم است شخص حتماً با اصول الفبای آن زبان از جمله رسم الخط، اعراب و قواعد آشنا شود، لازم بود در این فصل به تفضیل با قواعد کل زبان موسیقی آشنا گردیم و تمامی اصطلاحات را با همان روش و قاعده موسیقدانان پیگیری نماییم.

Abstract

Among the knowledge communication instruments, we can find the ones that are placed in the group of “musical language”. Musical language is a language with its own specific structures and grammars, which have been changing along the past hundreds of years. Getting to know this language needs many years of practice.

This thesis aims at recognition and performing of basic musical notes. Along this path, the input image, consisting of a composition that is presented as an image of an input, is taken into account.

The preprocessing on the image, specifically correction of the input image angle and noise reduction is performed and the compositions are divided into smaller unbreakable segments. During segmentation, the semantic surges resulting from the picture resolution are modified.

After identification of each element and its coordinates, we try to identify the signs by the artificial neural network method. It should be noted that the direction of reading the musical notes is from left to right. Thus, all signs are identified in the order of placement on the staff line. Finally, by the MIDI instrument, we create the MIDI matrix and translate all signs, schedules and spaces into the MIDI language and create their matrix. Finally, we play the composition by and instrument prepared by MATLAB to play the MIDI files.

Keywords:

Musical language, Musical note, Staff lines, Radon Transform, Artificial Neural Networks, MIDI

Shahrood University of Technology

Faculty of Computer Engineering

Music note recognition and perform from music note's images

By:Kamran Shokri

Dr. Hamid Hassanpour

September 2017