

صاحب امتیاز: انجمن فناوری اطلاعات و ارتباطات ایران

مدیر مسؤول: دکتر مسعود شفیعی

سر دبیر: دکتر کریم فائز

دبیر داخلی: دکتر هومان تحیری (دانشگاه شیراز)

شمارگان: ۱۰۰۰

حروفچینی و صفحه آرایی: انسبه مرجانی، فاطمه مرجانی

هیئت تحریریه:

فرخ حجت کاشانی (استاد)، دانشگاه علم و صنعت ایران

سید احمد رضا شرافت (استاد)، دانشگاه تربیت مدرس

مسعود شفیعی (استاد)، دانشگاه صنعتی امیرکبیر

محمد رضا عارف (استاد)، دانشگاه صنعتی شریف

کریم فائز (استاد)، دانشگاه صنعتی امیرکبیر

رضا فرجی دانا (استاد)، دانشگاه تهران

کمال محامدپور (استاد)، دانشگاه خواجه نصیرالدین طوسی

فرخ مروستی (استاد)، دانشگاه صنعتی شریف

سید احمد معتمدی (استاد)، دانشگاه صنعتی امیرکبیر

روزبه معینی (استاد)، دانشگاه صنعتی امیرکبیر

کاظم اکبری (دانشیار)، دانشگاه صنعتی امیرکبیر

کامبیز بدیع (دانشیار)، مرکز تحقیقات مخابرات ایران

علی اکبر جلالی (دانشیار)، دانشگاه علم و صنعت ایران

حمیدرضا ربیعی (دانشیار)، دانشگاه صنعتی شریف

امیر البدوی (دانشیار)، دانشگاه تربیت مدرس

محمد تشنه لب (دانشیار)، دانشگاه خواجه نصیر طوسی

اعضای هیئت مدیره:

• اعضای اصلی:

دکتر مسعود شفیعی (رئیس انجمن)

دانشگاه صنعتی امیرکبیر

مهندس علی نادپور (خزانه دار)

شرکت تکفام

دکتر محمود کمره ای

دانشگاه تهران

دکتر کمال محامدپور

دانشگاه صنعتی خواجه نصیرالدین طوسی

دکتر کاظم اکبری

دانشگاه صنعتی امیرکبیر

دکتر سعادت پورمظفری

دانشگاه صنعتی امیرکبیر

دکتر رمضانعلی صادق زاده

دانشگاه صنعتی خواجه نصیرالدین طوسی

دکتر علی اصغر عمیدیان

وزارت ارتباطات و فناوری اطلاعات

دکتر قاسم رمضانپور نرگسی

سازمان پژوهش های علمی و صنعتی ایران

دکتر وحید یزدانیان

وزارت ارتباطات و فناوری اطلاعات

مهندس عبدالمجید ریاضی

شرکت ایز ایران

• اعضای علی البدل:

دکتر احمد خادم زاده (مرکز تحقیقات مخابرات ایران)

امیر شهاب شاه میری (دانشگاه آزاد اسلامی)

بازرسان:

• بازرس اصلی:

مهندس علی اکبر عسگری

• بازرس علی البدل:

مهندس اکبر کاری دولت آبادی (دانشگاه شهید ستاری)

براساس نامه شماره ۳/۴۸۱۷ مورخ ۱۳۸۶/۶/۱۱ دفتر امور پژوهشی وزارت

علوم، تحقیقات و فناوری و بیانیه رأی کمیسیون بررسی نشریات

علمی کشور در تاریخ ۱۳۸۶/۴/۲۳، درجه علمی - پژوهشی به این فصلنامه

اعطا شده است

فهرست

فصلنامه علمی - پژوهشی فناوری اطلاعات و ارتباطات ایران

سال چهارم، شماره‌های ۱۳ و ۱۴، پاییز و زمستان ۱۳۹۱

سرمقاله

- ۱ ■ اصلاح ردیاب انتقال متوسط برای ردگیری هدف با الگوی تابشی متغیر
پیمان معلم، جواد عباس‌پور، علیرضا معمارمقدم، مسعود کاوش تهرانی
- ۹ ■ تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فراطیفی بر مبنای الگوریتم ژنتیک
فرهاد صمدزادگان، حدیثه سادات حسنی
- ۲۵ ■ ارزشیابی بکارگیری روش ترکیبی SPIHT-DCT با استفاده از مقیاس‌پذیری مکانی و مکانی-زمانی
برای کد کردن تصاویر ویدئویی HDTV
وحیدرضا صیرفیان، سیامک طالبی
- ۳۷ ■ بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه
جواد راستی، سید امیرحسین منجمی، عباس وفایی
- ۵۷ ■ تقویت محور مرکزی ساختارهای لوله‌ای و کاربرد آن در استخراج محور مرکزی سیاهرگ پورتال
امیرحسین فروزان، رضا آقایی‌زاده ظروفی، یوشی‌نبو ساتو، ماساتوشی هوری
- ۶۷ ■ پردازش تصاویر ورق‌های فولادی به منظور آشکارسازی عیوب به کمک موجک گابور
مصطفی صادقی، مسعود شفیعی
- ۷۵ راهنمای نگارش
- ۷۷ معرفی انجمن فناوری اطلاعات و ارتباطات
- ۷۸ فرم اشتراک نشریه
- ۷۹ فرم‌های عضویت حقوقی و حقیقی انجمن
- ۸۱ اعضای حقوقی و حقیقی جدید انجمن
- چکیده انگلیسی مقالات
-

پس از حمد و ثنای پروردگار متعال و در روزهای پایانی سال جهاد اقتصادی، برای تمامی خدمتگزاران این سرزمین بزرگ موفقیت و سربلندی را مسألت می‌نماییم.

شماره‌های دیگر از نشریه انجمن فناوری اطلاعات و ارتباطات ایران به مرحله چاپ و نشر رسیده است که همچون گذشته در برگیرنده گزارش دستاوردهای علمی محققان و پژوهشگران کشور عزیزمان و در یک کلام عصاره و نتیجه حیات طیبه علمی آن‌هاست. این دستاوردها در زمینه‌هایی از دانش و فناوری حاصل شده‌اند که عملاً لبه‌های علم محسوب می‌شوند یعنی فناوری اطلاعات و ارتباطات از نظر زمانی نیز دستیابی به این موفقیت‌ها قابل تأمل است.

هموطنان عزیز ما عموماً و محققان ما خصوصاً روزهایی را سپری می‌کنند که تحریم‌های ظالمانه از هر سو نه تنها فعالیت‌های علمی که حتی زندگی روزمره مردم ما را تحت تأثیر قرار داده است و در کمال حیرت دانشمندان و پژوهشگران ما همچنان به تحقیق و نوآوری آن هم در لبه‌های دانش و فناوری مشغولند و روز به روز از نظر کیفی و کمی به تعداد مقالات و محصولات ناشی از فعالیت‌هایشان افزوده می‌شود. شاید تا چند سال قبل هیچ‌گاه حقانیت حدیث شریف نبوی "اگر علم در ثریا هم باشد، مردانی از پارس بدان دست خواهند یافت" را چنین باور نمی‌داشتیم و این نیز معجزه‌ای است از معجزات الهی که پس از قرن‌ها و با تمام دشواری‌های مادی و غیر مادی، مردم سرزمین پارس نه به ثروت و نه قدرت‌های بی پایه و مایه مادی بلکه به علم و فناوری در جهان شهره هستند و در این راه به خود خستگی و ضعف راه نمی‌دهند. در زمینه ارتباطات و فناوری اطلاعات نیز همچون سایر زمینه‌های دیگر، کشورمان با همت متخصصان خود در راه غلبه برهم تحریم‌ها را یافته است و در راه تعالی گام بر می‌دارد.

آنچه اندیشمندان و پژوهشگران بخش ارتباطات و فناوری اطلاعات را به استمرار صف شکنی‌های عالمانه خود دلگرم می‌سازد، ثبات در سیاست‌های کلان این بخش و حمایت‌هایی است که از فعالیت‌های تحقیقاتی آنان به عمل می‌آید که امیدواریم این حمایت‌ها روز به روز گسترش یافته و زمینه‌های مناسبی را برای توسعه و تکامل کشور فراهم سازد، به هر روی مفتخریم که بازگو کننده و ناشر بخشی اندک از تلاش‌های اساتید، دانشجویان و پژوهشگران بخش ارتباطات و فناوری اطلاعات در این نشریه هستیم و بر این باوریم که پایمردی و استقامت خردمندان و اندیشمندان این کشور عزیز و اسلامی دیر یا زود به ثمر خواهد نشست و آحاد مردم از ثمرات آن بهره‌مند خواهند شد. انشا...

دکتر مسعود شفیعی

مدیرمسئول

اصلاح ردیاب انتقال متوسط برای ردگیری هدف با الگوی تابشی متغیر

پیمان معلم^{۱*} جواد عباس پور^{**} علیرضا معمارمقدم^{***} مسعود کاوش تهرانی^{****}

^{*} دانشیار، دانشکده مهندسی برق، دانشگاه اصفهان، اصفهان

^{**} کارشناس ارشد الکترواپتیک، پژوهشکده اپتوالکترونیک، دانشگاه صنعتی مالک اشتر، تهران

^{***} دانشجوی دکتری مهندسی برق الکترونیک، دانشکده مهندسی برق، دانشگاه اصفهان، اصفهان

^{****} دانشیار، دانشکده فیزیک، دانشگاه صنعتی مالک اشتر، شاهین شهر

تاریخ دریافت: ۱۳۹۱/۰۳/۰۳

تاریخ پذیرش: ۱۳۹۱/۱۰/۰۹

چکیده

یکی از روش‌های مرسوم در زمینه‌ی ردیابی تصویری اهداف غیرصلب، استفاده از روالی تکراری به نام انتقال متوسط در تعیین موقعیت مد مرکزی هدف است. نمایش هدف در ردیاب انتقال متوسط برپایه‌ی هیستوگرام ویژگی بانقاب‌گذاری مکانی با یک کرنل مستقل از جهت انجام می‌شود. بحرانی‌ترین چالش در ردیاب انتقال متوسط، تنظیم مقیاس کرنل است. تاکنون هیچ روش کارآمد و بی‌عیب و نقصی برای تنظیم و یا وفق دهی ابعاد کرنل، زمانی که ابعاد هدف تغییر می‌کند، ارائه نشده است. مشکل دیگر ردیاب انتقال متوسط در رویارویی با هدف با الگوی تابشی متغیر پیش می‌آید. در این مقاله با رویکرد حل این مشکلات، الگوریتم ردیابی انتقال متوسط همراه با اندازه بندی وفقی قوی ارائه می‌گردد، ضمن این که مشکل الگوریتم انتقال متوسط را در مواجهه با تغییرات الگوی تابشی هدف با وفق دهی مدل هدف در هر قاب حل می‌کند. در روش پیشنهادی، ابتدا با استفاده از روش محاسبه‌ی توان ناشی از مشتقات مکان-زمانی شدت پیکسل‌های تصویر، ابعاد پنجره در قاب بعد تنظیم می‌شود. سپس نتایج حاصل از اندازه بندی پنجره در قاب بعد، در ردیاب انتقال متوسط اعمال می‌شود. نتایج نشان می‌دهند که استفاده از الگوریتم پیشنهادی ضمن اینکه به کاهش خطای موقعیت یابی هدف در مقایسه با الگوریتم انتقال متوسط استاندارد می‌انجامد، در برابر تغییرات تباین^۲ و الگوی تابشی هدف نیز کارایی قابل توجهی از خود نشان می‌دهد.

کلید واژگان: ردیاب انتقال متوسط، مدل هیستوگرام، کرنل،

پنجره ردیابی، اندازه بندی وفقی، تباین.

۱. مقدمه

طی یک دهه اخیر کاربردهای ماشین بینایی به طور وسیعی افزایش پیدا کرده است. نمونه ای از این کاربردها که از زمینه‌های مورد علاقه در ماشین بینایی است، ردیابی خودکار اهداف می باشد که به علت کاربردهای تجاری-صنعتی متنوع آن، اهمیت ویژه‌ای در سالهای اخیر پیدا کرده است. در واقع ردیابی هدف، فرآیند آشکارسازی و تعیین حرکت هدف در دنباله‌ای از تصاویر ویدئویی می‌باشد. با توجه به داده‌های تصویری شیوه‌های مختلف ردیابی به چهار دسته عمده تقسیم می‌شوند [۱] ردیابی بر اساس مدل سه بعدی [۲] ردیابی بر مبنای نواحی [۳] ردیابی بر مبنای کانتورهای فعال [۴] و ردیابی بر مبنای نقاط ویژگی [۵].

در شیوه‌ی آخر به دلیل این که نقاط ویژگی متعلق به هدف، همگی حرکت یکسانی دارند، یک گروه پیکسلی را تشکیل می‌دهند که متمایز از گروه پیکسلی مربوط به اشیاء متحرک دیگر درون صحنه است. نقاط ویژگی می‌توانند لبه‌ها، مرز نواحی هدف، نواحی با رنگ-های خاص، منحنی درون بافت و یا نقاط گوشه‌ای هدف باشند [۶]. الگوریتم‌هایی که مبنای کار را این ویژگی‌ها قرار می‌دهند، الگوریتم‌های مبتنی بر ویژگی نامیده می‌شوند. گام بعدی در این الگوریتم‌ها تعریف فضای ویژگی برای توصیف هدف است. در این مقاله نیز اساساً از یک الگوریتم کرنل- پایه کمک می‌گیریم [۷]، که نوعی الگوریتم ردیابی مبتنی بر ویژگی است. در این روش، موقعیت یابی هدف بر پایه‌ی اجرای فرایند انتقال متوسط انجام می‌شود [۸]. به همین دلیل این روش به الگوریتم ردیابی بر پایه‌ی انتقال متوسط مشهور است. در واقع ردیاب انتقال متوسط از الگوی فرکانس پایین

^۱ نویسنده عهده‌دار مکاتبات (p_moallem@eng.ui.ac.ir)

ردیاب پایه بگذارد، موفقیت آن در ردیابی هدف دور از انتظار نخواهد بود. برای این منظور لازم است همراه با وفق دهی مدل هدف، اندازه بندی سریع و پایداری روی پنجره انجام شود. مرجع [۱۰] پس از بررسی و مقایسه انواع الگوریتم های اندازه بندی پنجره، کنترل وفقی ابعاد پنجره را با استفاده از روشی مبتنی بر توان ناشی از مشتقات زمان - مکانی تصویر هدف پیشنهاد داده است که از سرعت واکنش قابل قبولی متناسب با حرکت هدف برخوردار است. مهمترین کارکرد روش پیشنهادی، تعیین نسبتاً دقیق پهنای کرنل هدف در قاب جاری و البته قبل از اجرای الگوریتم انتقال متوسط است. بهترین دستاورد این ایده را می توان کاهش مؤثر نقش زمینه در الگوی هدف دانست.

مطالب این مقاله به این صورت تنظیم شده است که ابتدا به شرح کلی از الگوریتم پیشنهادی در قسمت بعد خواهیم پرداخت. سپس به مرور الگوریتم ردیابی انتقال متوسط استاندارد می پردازیم. وفق دهی مدل هدف و کنترل وفقی ابعاد پنجره، مطالب ادامه ی مقاله را تشکیل خواهند داد. در انتها نیز عملکرد روش پیشنهادی را با بکارگیری رشته تصاویر ساختگی و واقعی مورد بررسی قرار می دهیم و نتایج اجرای آن را نسبت به دو حالت استاندارد [۷] و تشکیل فضای مقیاس (کولینز) [۱۱]، مقایسه می کنیم.

۲. توصیف کلی روش پیشنهادی

بلوک دیاگرام کلی الگوریتم پیشنهادی در شکل ۱ آورده شده است. فرض کنیم که موقعیت هدف و ابعاد اولیه پنجره ردیابی در فریم اول (F1) تعیین شده است، حال به تنظیم وفقی ابعاد پنجره با ابعاد هدف در قاب جاری ($Fk, k \geq 2$) با استفاده از روش محاسبه ی تغییرات مکان-زمانی شدت پیکسل های تصویر می پردازیم [۱۰]. با انجام چنین کاری در واقع محدوده ی واقعی تری را جهت استخراج مدل هیستوگرام هدف مشخص کرده ایم. حال ابعاد پنجره در قاب جاری را به ردیاب انتقال متوسط استاندارد انتقال می دهیم. در این مرحله موقعیت یابی هدف در قاب جاری ($Fk, k \geq 2$)، با تشکیل مدلی از هدف انجام می شود که قابل تنظیم در تمامی شرایط باشد [۹]. حلقه ی الگوریتم با بهنگام سازی مشخصات پنجره با موقعیت جدید هدف بسته می شود.

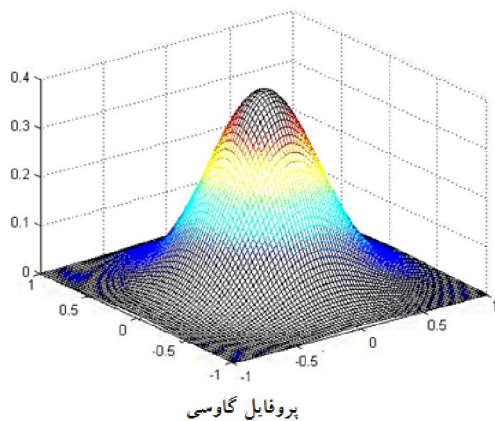
تصویر هدف کمک می گیرد و در هر قاب به دنبال مدلی از این الگوی ناحیه ای هدف در تصویر است که به مدل هیستوگرام مرجع نزدیکتر است. این مدل اصطلاحاً مدل هیستوگرام داوطلب هدف نامیده می شود. سنجه ی تعیین مدل داوطلب هدف، استفاده از ضریب تشابه بین دو مدل مرجع و داوطلب است [۷].

اما پارامتر بحرانی در ردیاب انتقال متوسط پهنای کرنل است، چراکه استفاده از یک الگوریتم اندازه بند کارآمد روی تصویر هدف در کنار این ردیاب، عملکرد آن را به شکل قابل توجهی بهبود می بخشد. تأثیر این موضوع در زمانی که ردگیری هدف با اندازه ی تصویری متغیر مدنظر است به وضوح قابل درک است. مرجع [۷] روشی را در این رابطه پیشنهاد نموده که ابعاد پنجره هدف را بر مبنای بهترین مدل هیستوگرام تنظیم می کند. مشکل اصلی این روش (بخصوص زمانی که توزیع طیفی جسم یکنواخت باشد) این است که شعاع بهینه ی بدست آمده از این روش، تمایل به کوچکتر شدن دارد. دلیل این امر این است که در مورد شعاع های ۱۰ درصد کوچکتر، همواره ضریب مشابهت بیشتری بدست می آید (حتی اگر مقیاس ثابت باشد). همچنین این روش در شرایطی که اندازه هدف نسبتاً سریع تغییر می کند، ردیابی پایداری را نتیجه نمی دهد. این موضوع از ثابت بودن میزان انبساط یا انقباض پنجره ناشی می شود. در مرجع [۱۱] کولینز موقعیت-یابی هدف و اندازه بندی پنجره ی ردیابی در ردیاب انتقال متوسط را بر مبنای تشکیل فضای مقیاس انجام می دهد. مشکلی که این روش دارد این است که اندازه بندی پنجره ی ردیابی تنها زمانی به نتیجه ی خوبی منجر می شود که هدف ردیابی شده صلب، بدون تغییر شکل و دارای توزیع یکنواختی باشد. در غیر این صورت لبه های داخلی جسم به عنوان مقیاس در نظر گرفته می شود. همچنین طول و عرض تصویر باید با نسبت یکسان بزرگ و یا کوچک شوند.

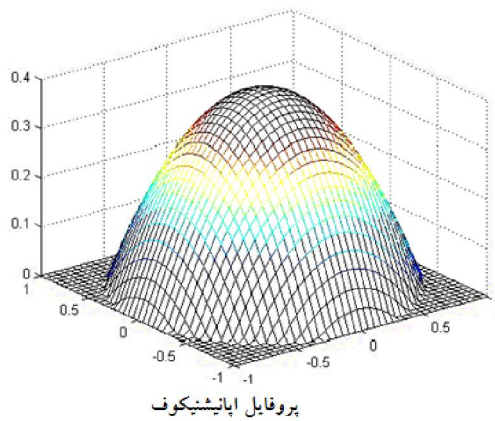
مشکل دیگر ردیاب انتقال متوسط استاندارد، ضعف در رویارویی با هدف با شدت و الگوی تابشی متغیر است، به طوری که در چنین حالتی ردیاب انتقال متوسط با مشکل مواجه می شود. منشأ اصلی پیدایش این مشکل، ثابت فرض نمودن مدل هدف در تمامی قاب ها است.

به هر حال مشکل هدف با الگوی تابشی متغیر را می توان با وفق دهی مدل هدف حل کرد [۹]، اما چنانچه بتوان الگوریتمی داشت که الگوی هدف را با کمترین تأثیر پذیری از الگوی زمینه ی آن در اختیار

۴) حال چنانچه فاصله ی موقعیت جدید هدف نسبت به حالت قبل آن بیشتر از یک پیکسل باشد مجدداً مراحل ۱ تا ۳ را تکرار کنید.



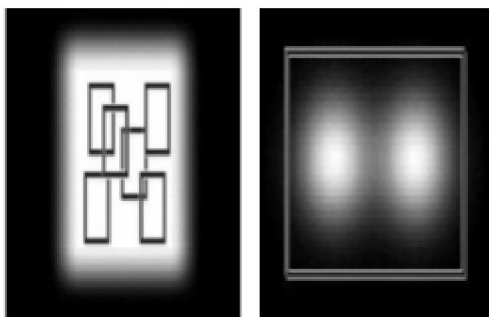
پروفایل گاوسی



پروفایل اپانیشینکوف

شکل ۲. منحنی نمایه کرنل گاوسی (چپ) و اپانیشینکوف (راست).

تدریجی انجام می‌شود. این امر باعث می‌شود اندازه‌بندی سریع و پایداری را نداشته باشیم که در نهایت با یکی از دو حالت بالا مواجه خواهیم شد.



(ب)

(الف)

شکل ۳. انتخاب مقیاس کرنل. الف- انتخاب کرنل با مقیاس بزرگتر نسبت به ابعاد هدف. در این حالت، کرنل شامل قسمت‌های ناخواسته‌ی زمینه خواهد شد. ب- انتخاب کرنل با مقیاس کوچکتر از ابعاد هدف. در این حالت کرنل در همسایگی یک مد غیر واقعی می‌چرخد، به گونه‌ای که این امر در نهایت منجر به ردیابی ضعیفی خواهد شد.

ضعف دیگر این روش در اندازه‌بندی این است که ابعاد پنجره هدف بعد از اجرای الگوریتم انتقال متوسط و برای اعمال در فریم بعدی تنظیم می‌شود که چنانچه حرکت هدف بین دو فریم از نصف اندازه این پنجره بیشتر باشد، باعث توقف ردیابی خواهد شد. بنابراین اگر بتوانیم در قابجاری و قبل از شروع الگوریتم انتقال متوسط، ابعاد پنجره ردیابی را با اندازه هدف وفق دهیم، قادر به تحقق موقعیت یابی دقیق‌تری از هدف خواهیم شد. چراکه قسمت اعظم پیکسل‌های غیرهدف، در خارج از پنجره ردیابی قرار گرفته و در محاسبات وارد نمی‌شوند. برای این منظور، ما از الگوریتم اندازه بندی تشریح شده در مرجع [۱۰] کمک گرفته‌ایم. این روش ابتدا پنجره هدف را مطابق

۳.۳. بهینه سازی الگوریتم انتقال متوسط به کمک مدل وقتی هدف

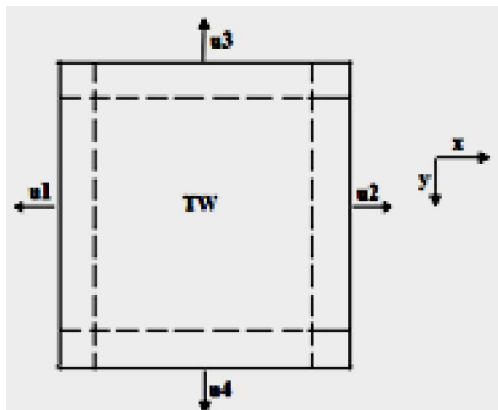
یکی از مشکلات ردیاب انتقال متوسط، کارایی ضعیف در برخورد با مواردی از قبیل تغییرات شدت و الگوی تابشی هدف می‌باشد. این مشکل را می‌توان تا حدودی با کاربرد ضریب شباهت بین دو مدل در وفق دادن مدل هدف، رفع کرد. وفق‌دهی مدل هدف با ترکیبی از دو مدل مرجع هدف در قاب آغازین \hat{q} و مدل تجمعی هدف تا قاب قبل $\hat{q}(k-1)$ انجام می‌شود [۹]. این روش تحت عنوان چشم پوشی پارامتری ذکر می‌شود. در این حالت مدل وفق‌یافته‌ی هدف در قاب جاری، $\hat{q}(k)$ ، با رابطه زیر بدست می‌آید:

$$\hat{q}(k) = \rho \cdot q + (1 - \rho) \cdot \hat{q}(k-1) \quad (9)$$

در رابطه بالا ρ ضریب تشابه بین دو مدل هدف است که قبلاً از رابطه (۶) بدست آمده‌است.

۴. بهینه‌سازی اندازه بندی وقتی در الگوریتم انتقال متوسط

هم‌چنان‌که گفتیم، پارامتر بحرانی در ردیاب انتقال متوسط، پهنای کرنل است، زیرا مطابق قسمت (الف) از شکل ۳، اگر ابعاد کرنل را بزرگتر از ابعاد هدف تنظیم کنیم، باعث حضور پیکسل‌های زمینه در تشکیل الگوی هدف خواهیم شد. این مطلب باعث می‌شود که مدل هدف آلوده شود و ادامه پیدا کردن این شرایط در نهایت منجر به خطا در ردیابی خواهد شد. برعکس اگر ابعاد کرنل را کوچکتر از ابعاد هدف تنظیم کنیم، علاوه بر این که نقش پیکسل‌های حواشی و لبه‌ای را در تشکیل مدل هدف نخواهیم داشت، در مواردی مطابق قسمت (ب) از شکل ۳، چرخش کرنل حول و حوش مد غیر واقعی را خواهیم داشت، به گونه‌ای که در نهایت ردیابی ضعیف را خواهیم داشت. مشکل روش اندازه بندی ارائه شده استاندارد [۷]، این است که در برخی کاربردها تنظیم ابعاد پنجره با اندازه هدف به‌صورت آرام و



شکل ۵. انتساب بردار یکه به هر یک از چهار ضلع ردیاب

۵. نتایج آزمایشگاهی

در این قسمت به پیاده‌سازی روش پیشنهادی روی رشته‌ای از تصاویر مختلف واقعی و ساختگی می‌پردازیم و نتایج آن را نسبت به روش استاندارد [۷] و کولینز [۱۱] مقایسه می‌کنیم. آزمایش اول اختصاص به ردگیری هدف در تصویر مادون قرمز دارد. این دنباله‌ی ویدئویی شامل ۱۰۰ قاب است که ابعاد هر قاب ۲۴۰×۳۲۰ است. مطابق شکل ۶، هدف تحت ردیابی در این فیلم، فردی است که با عبور از پله، ابعاد آن تغییر می‌کند. هدف از بررسی این فیلم، ارزیابی الگوریتم در رویارویی با تصاویر مادون قرمز و نیز هدف با ابعاد متغیر است. اولین گام در ردیابی، نمایش هدف است. در اینجا ما هدف را باناحیه‌ی مستطیل‌شکل نمایش می‌دهیم. قدم بعدی تعریف فضای ویژگی در ردیاب انتقال متوسط است. در این آزمایش ویژگی منتخب در هر سه روش استاندارد، کولینز و الگوریتم پیشنهادی، سطح خاکستری است. به منظور ملاحظات سرعت و نیز تقطیع هدف به نواحی مجزا، این فضای ویژگی را به ۳۲ قسمت مجزا تقسیم می‌کنیم. در این آزمایش کرنل مورد استفاده در ردیاب انتقال متوسط، کرنل اپانیشنیکوف است. نرخ تعداد تکرار نیز ۱۰ می‌باشد. در الگوریتم اندازه‌بندی‌مان از نرخ تصرف ۸۰٪ پنجره استفاده کرده‌ایم و آستانه‌ی پیچیدگی زمینه را ۰/۱۵ در نظر گرفته‌ایم. ثابت‌های متناظر با سرعت اندازه‌بندی پنجره (Cx, Cy) را نیز (۰/۸, ۰/۸) در نظر گرفته‌ایم.

شکل ۶ نتایج اجرا یاندازه‌بندی سه الگوریتم ردیابی استاندارد، کولینز و روش پیشنهادیماندر قاب‌های ۱، ۳۵، ۴۰ و ۶۵ را نشان می‌دهد. نتایج نشان از موفقیت روش ما در اندازه‌بندی دقیق و نرم (به‌ویژه در عبور هدف از پله) و نیز ردیابی موفق هدف در مقایسه با دو روش دیگر نشان می‌دهد.

در این قسمت به ارزیابی خطای موقعیت‌یابی و اندازه‌بندی سه روش تحت بررسی می‌پردازیم. بدین منظور کمیت‌های خطای موقعیت‌یابی و اندازه‌بندی E_p و E_s را به صورت زیر تعریف می‌کنیم:

شکل ۴ به هشت زیر پنجره لبه‌ای درونی و بیرونی تقسیم می‌کند و سپس به محاسبه نسبت توان متناظر با تغییرات مکان-زمانی^۱ هدف در هر کدام از این زیر پنجره‌ها به توان کل پنجره مطابق رابطه زیر می‌پردازد:

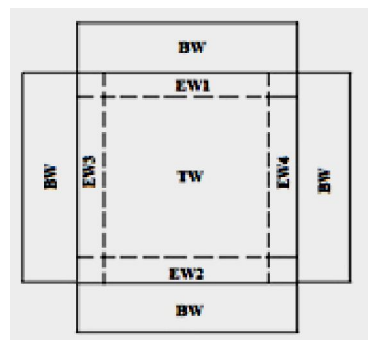
$$r_i = \frac{P_{EW_i}^T - \hat{P}_n}{P_{TW} - \hat{P}_n} \quad i=1,2,3,4 \quad (10)$$

که $P_{EW_i}^T$ توان واقعی هدف در زیر پنجره‌ی لبه‌آم، P_{TW} توان کل پنجره‌ی ردیابی و \hat{P}_n مقدار توان تخمینی ناشی از گردادیان‌های مکان-زمانی است.

بنابراین بردار جابجایی متناسب با هر ضلع پنجره با فرمول زیر

محاسبه می‌شود:

$$ds_i = \alpha_i \beta_i \hat{u}_i, \quad i=1,2,3,4 \quad (11)$$



شکل ۴. نحوه تقسیم بندی پنجره ردیابی

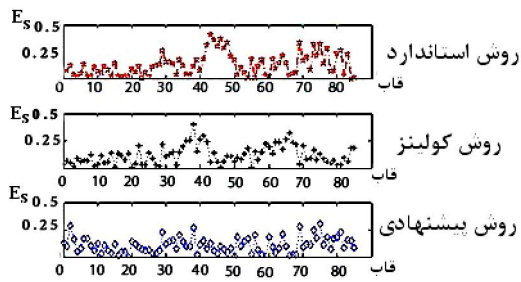
α_i و β_i به ترتیب ضرایب وزنی اندازه‌بندی و جهت‌دهی مرتبط با ضلع آم هستند، به طوری که اگر C_x و C_y را ثابت‌هایی بدانیم که سرعت اندازه‌بندی را کنترل می‌کنند، آن‌گاه:

$$\alpha_i = \begin{cases} \max\left(1, \frac{T_y}{C_y}\right) & i=1,2 \\ \max\left(1, \frac{T_x}{C_x}\right) & i=3,4 \end{cases}, \quad \beta_i = \begin{cases} -1 & r_i < T_{EL} \\ 2 & r_i \geq T_{EL} \end{cases} \quad (12)$$

T_{EL} آستانه‌ای است که بطور تجربی تعیین می‌شود. شکل ۵

اختصاص بردار جابجایی به هر ضلع را نشان می‌دهد. در واقع، هدف از طی این مراحل این است که نرخ تصرف هدف در پنجره ردیابی، بیشینه گردد. بنابراین در مواردی که پنجره‌ی ردیابی به شکلی نامناسب و بسیار بزرگتر از هدف انتخاب شده باشد الگوریتم با تنظیم سریع آن، نقاط هدف را برای به خدمت‌گیری در یک ردیاب ناحیه نظیر انتقال متوسط پالایش می‌کند.

¹ Spatiotemporal
² Occupancy Rate



شکل ۸. منحنی خطای اندازه‌بندی سه روش استاندارد، کولینز و پیشنهادی در کل قاب‌های فیلم.

در این فیلم که ابعاد هر قاب آن 576×720 است، هدف با سرعتی حدود 0.75 پیکسل بر فریم در فریم‌های آغازین شروع به حرکت کرده و سرعت آن در ادامه به تدریج افزایش می‌یابد و در انتها به حدود سه پیکسل بر فریم می‌رسد. در این آزمایش نیز ویژگی را سطح خاکستری در نظر می‌گیریم و فضای ویژگی را به 16 بخش متمایز تقسیم می‌کنیم. کرنل مورد استفاده در ردیاب انتقال متوسط، کرنل اپانیشینکوف است و تعداد تکرار ده مرتبه به ازای هر فریم می‌باشد. در روش پیشنهادی نیز نرخ تصرف را باز هم 80% در نظر گرفته‌ایم. آستانه‌ی پیچیدگی زمینه 0.15 است و ثابت‌های متناظر با سرعت اندازه‌بندی پنجره (C_x, C_y) ، $(0.8, 0.8)$ می‌باشد. شکل ۹ نتایج اجرای ردیابی سه روش استاندارد، کولینز و روش پیشنهادی را نشان می‌دهد. در هر سه روش به منظور حل معضل رویارویی با تغییرات در مدل هدف (به علت تغییرات مداوم در تباین هدف) از روش وفق‌دهی مدل استفاده می‌کنیم.

مطابق قسمت (الف) از شکل ۹ ما ابتدا شروع اندازه‌بندی را با ابعادی بزرگتر از ابعاد هدف و در جهت حرکت مدل‌های بعدی در نظر گرفته‌ایم، به گونه‌ای که ردیاب انتقال متوسط بتواند در ادامه هدف را به طور کامل ردیابی کند. در قسمت (ب) و (ج) از این شکل عملکرد ناموفق روش‌های استاندارد و کولینز را در موقعیت‌یابی همراه با اندازه‌بندی نامناسب نشان می‌دهد. روش استاندارد، به علت اندازه‌بندی غیر نرم، تدریجی و آرام، که عملاً مرتبط با ثابت در نظر گرفتن درصد افزایش یا کاهش شعاع پنجره است، نمی‌تواند موقعیت‌یابی درستی از هدف را داشته‌باشد. در روش کولینز مشاهده می‌کنیم که به علت این که نرخ رشد ابعاد هدف در راستای افقی و عمودی یکسان نیست، این روش نیز در اندازه‌بندی با مشکل مواجه می‌شود، زیرا قبلاً ذکر کرده بودیم که یکی از نقاط ضعف این روش در اندازه‌بندی هدف با توزیع غیر یکنواخت است، به ویژه زمانی که تغییر ابعاد هدف در راستای افقی و عمودی یکسان نباشد.

$$E_p = \sqrt{(x - \hat{x})^2 + (y - \hat{y})^2} \quad [\text{پیکسل}]$$

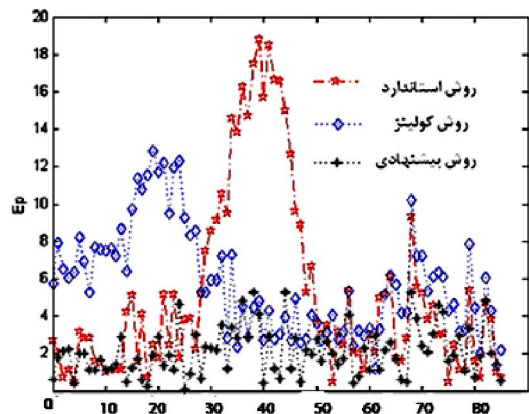
$$E_s = \frac{|A - \hat{A}|}{A} \quad (13)$$

که در این جا (x, y) مختصات واقعی هدف است که به صورت دست‌یاد در فریم تخمین زده‌ایم. (\hat{x}, \hat{y}) نیز مختصات موقعیت تخمینی به وسیله الگوریتم مورد نظر است. A مساحت واقعی پنجره ردیابی در هر فریم است و \hat{A} مساحت تخمینی در هر فریم است. شکل ۷ و ۸ منحنی خطای موقعیت‌یابی و اندازه‌بندی سه روش را در طول قاب‌های ۱ تا ۶۵ که در آن ابعاد هدف در حال تغییر است، نشان می‌دهد. این دو منحنی برتری روش ما را نسبت به دو روش دیگر نشان می‌دهد.

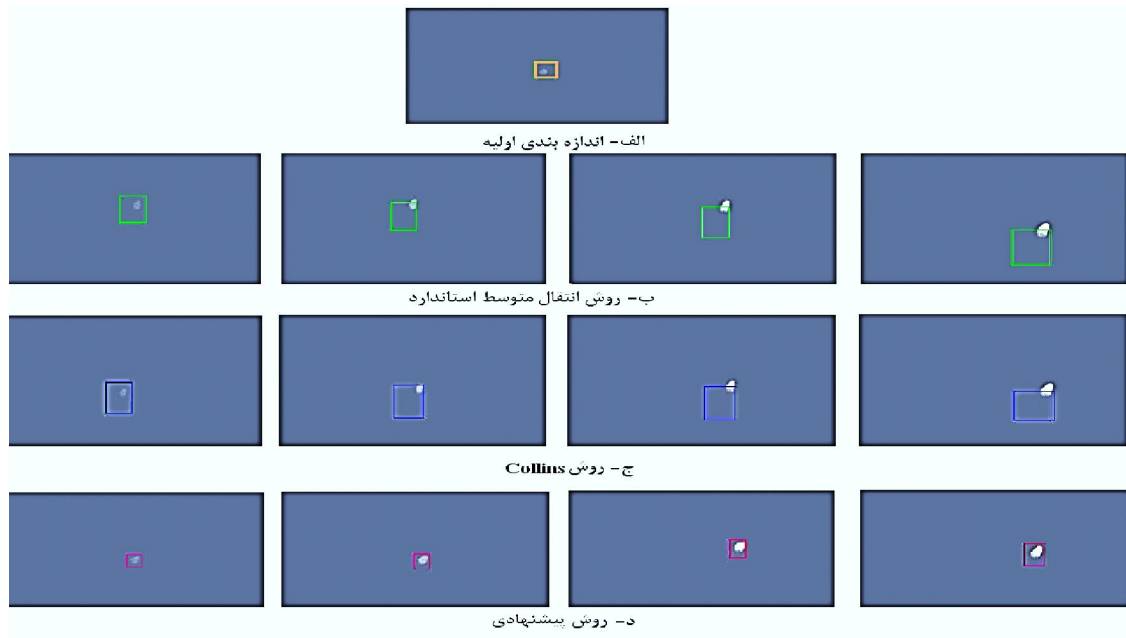
به منظور آزمایش کارایی روش پیشنهادی نسبت به الگوریتم انتقال متوسط استاندارد [۷] و روش کولینز [۱۱] در مواجهه با هدف با الگوی تابشی و ابعاد متغیر در طول زمان، یک فیلم ساختگی را به گونه‌ای شبیه‌سازی کرده‌ایم که تباین دیده شده از هدف در طول زمان افزایش و سپس کاهش می‌یابد.



شکل ۶. دنباله‌ی هدف در تصویر مادون قرمز. نتایج اندازه‌بندی سه روش استاندارد، روش کولینز و روش پیشنهادی در قاب‌های ۱، ۳۵، ۴۰ و ۶۵.

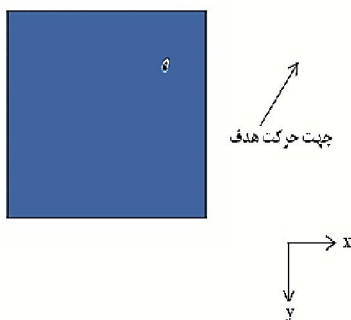


شکل ۷. منحنی خطای موقعیت‌یابی سه روش استاندارد، کولینز و پیشنهادی.



شکل ۹. نتیجه‌ی اجرای سه روش استاندارد، کولینز و روش پیشنهادی روی دنباله‌ی ویدئویی هدف با تباین و ابعاد متغیر در قاب‌های ۸۵۰، ۹۵۰، ۹۰۰ و ۱۰۰۰ همراه با اندازه‌بندی اولیه.

همانطور که در این آزمایش دیدیم، شکل هدف مرتباً در حال تغییر است، به گونه‌ای که ابعاد اولیه‌ی هدف از 18×16 پیکسل به ابعاد نهایی 31×25 پیکسل تغییر می‌کند. حال با نگاهی به شکل ۱۰ که اختلاف زمانی بین دو قاب متوالی نمایش می‌دهد، متوجه می‌شویم که جمله‌ی تغییر زمانی شدت پیکسل یعنی $T(x,y)$ برای ناحیه‌ی روشن در جلوی هدف در شکل (که به نوعی مشخص‌کننده‌ی جهت حرکت هدف است) مثبت است، در حالی که این مقدار برای ناحیه‌ی تاریک در پشت هدف منفی است. واضح است که این جمله برای ناحیه‌ی داخلی هدف تقریباً صفر است. بنابراین جمله‌ی کسر توان، یعنی رابطه‌ی (۱۰) در زیر پنجره‌های لبه‌ای متلاقی با اضلاع جلویی (پشتی) در جهت حرکت هدف افزایش (کاهش) خواهد یافت. این مطلب به این معنی است که جهت‌دهی ابعاد پنجره در جهت حرکت هدف است.



شکل ۱۰. اختلاف زمانی بین دو قاب متوالی برای تخمین جمله‌ی زمانی توان در رابطه‌ی (۱۰).

قسمت (د) از شکل ۹ موفقیت روش پیشنهادی را در مقایسه با دو روش دیگر نشان می‌دهد. دلایل این موفقیت به چند خاطر است. اول این که اندازه‌بندی این روش در دو جهت افقی و قائم برخلاف روش کولینز مستقل از یکدیگر است. دلیل دیگری که از نقاط قوت روش ما نسبت به دو روش دیگر است، اجرای اندازه‌بندی پنجره مستقل از اجرای موقعیت‌یابی ردیاب انتقال متوسط است. با استفاده از این ترفند می‌توان قبل از اجرای الگوریتم انتقال متوسط، ابعاد پنجره را در فریم بعدی تنظیم کرد و در نتیجه نقش الگوی زمینه را در تشکیل الگوی هدف به گونه‌ی مؤثری کاهش داد. اهمیت این موضوع زمانی بیشتر می‌شود که چنانچه حرکت هدف بین دو فریم از نصف اندازه پنجره بیشتر باشد، عدم استفاده از این ایده در نهایت منجر به عقب ماندن ردیاب و باز ایستادن از ردیابی می‌شود. اثبات این مطلب را می‌توان در شکل ۹ دید، جاییکه دو روش استاندارد و کولینز در نهایت در ردیابی از هدف عقب می‌افتند. بنابراین اگر بتوانیم در قاب جاری و قبل از شروع الگوریتم انتقال متوسط، ابعاد پنجره ردیابی را با اندازه هدف وفق دهیم، قادر به تحقق موقعیت‌یابی دقیق‌تری از هدف خواهیم شد. چراکه قسمت اعظم پیکسل‌های غیر هدف در خارج از پنجره ردیابی قرار گرفته و در محاسبات وارد نمی‌شوند. دلیل سوم موفقیت روش ما که به گونه‌ای با دلیل اول در ارتباط است، جهت‌دهی در اندازه‌بندی پنجره است. حال به استنباط کیفی این مطلب می‌پردازیم.

کولینز افزایش می‌دهد. البته استفاده یک پردازنده سریع نیز قید زمان و سرعت اجرا را غیر قابل لمس خواهد کرد. بنابراین هرچند افزایش دقت با در نظر گرفتن ملاحظات برخط بودن الگوریتم در روش پیشنهادی به بهای کاهش سرعت اجرای الگوریتم می‌انجامد، ولیکن این امر منجر به دستیابی به روش ردیابی ارزشمند همراه با حل مشکلات ردیاب انتقال متوسط استاندارد شده است.

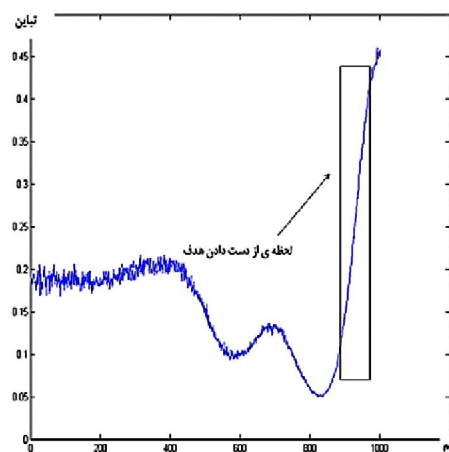
مراجع

- [1]. H. Tahvilian, P. Moallem, A. Monadjemi, Balloon Energy Based on Parametric Active Contour and Directional Walsh-Hadamard Transform and its Application in Tracking of Texture Object in Texture Background, EURASIP Journal on Advances in Signal Processing, Springer, 2012:253, Dec. 2012.
- [2]. A.I. Comport, E. Marchand, F. Chaumette, Efficient model-based tracking for robot vision, Advanced Robotics, Vol.19, No. 10, pp. 1097-1113, 2005.
- [3]. P. Salembier and F. Marques, Region-based representations of image and video: Segmentation tools for multimedia services, IEEE Transactions on Circuits, Systems and Video Technology, Vol. 9, pp. 1147-1169, 1999.
- [4]. M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, in 1st International Conference on Computer Vision, UK, pp. 259-269, 1987.
- [5]. J. Verestoy and D. Chetverikov, Comparative performance evaluation of four feature point tracking techniques, 22nd workshop of the Austrian pattern recognition group, Austria, pp. 255-263, 1998.
- [6]. ع. معمارمقدم، پ. معلم، ردیابی سریع و کارآمد هدف در رشته تصاویر ویدئویی به کمک انباره سازی خطاهای تطبیق‌دهی نقاط گوشه‌ای آن، چهارمین کنفرانس ماشین بینایی و پردازش تصویر ایران، جلد دوم، صفحات ۹۸۱ الی ۹۹۰، دانشگاه فردوسی مشهد، ۱۳۸۵.
- [7]. D. Comaniciu, V. Ramesh, P. Meer, Kernel-based object tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 5, pp. 564-577, 2003.
- [8]. D. Comaniciu, P. Meer, Mean shift: A robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 603-619, 2002.
- [9]. N.M. Artner, W. Burger, A Comparison of mean shift Tracking Methods, 12th Central European Seminar on Computer Graphics, Austrian, pp. 197-204, 2008.
- [10]. P. Moallem, A. Memarmoghaddam, M. Ashourian, Robust and Fast Tracking Algorithm in video Sequences by Adaptive Window Sizing Using a Novel Analysis on Spatiotemporal Gradient Powers, Journal of Circuits, Systems, and Computers, Vol. 16, No. 2, pp. 305-317, 2007.
- [11]. R. Collins, Mean-shift blob tracking through scale space, IEEE International Conference on Computer Vision and Pattern Recognition, USA, Vol. 2, pp. 234-240, 2003.

همانطورکه در ابتدای این آزمایش متذکر شدیم این آزمایش اختصاص به ردیابی هدفی دارد که در الگوی تابشی آن در حال تغییر است. شکل ۱۱ منحنی تغییرات تباین هدف در این فیلم ساختگی را نشان می‌دهد. مقدار تباین از فرمول زیر محاسبه شده است:

$$\text{تباین} = \frac{\mu_{L_{tgt}} - \mu_{L_{bg}}}{\mu_{L_{tgt}} + \mu_{L_{bg}}} \quad (14)$$

که در آن $\mu_{L_{tgt}}$ و $\mu_{L_{bg}}$ به ترتیب میانگین شدت هدف و زمینه در کل قاب‌ها می‌باشد. با مشاهده این منحنی متوجه می‌شویم که علی‌رغم این‌که مشکل تغییر الگوی تابشی هدف را در هر سه روش توانستیم با به‌کارگیری روش وفوق‌دهی مدل هدف حل کنیم [۹]، اما به‌مرحال لحظه‌ای از دست دادن هدف در روش استاندارد و کولینز همزمان با این تغییرات شدید در مدل هدف می‌باشد که مهمترین علت آن، اندازه‌بندی نادرست دو روش مذکور می‌باشد.



شکل ۱۱. منحنی تغییرات تباین هدف در کل قاب‌های فیلم.

۶. نتیجه‌گیری

در این مقاله با ارائه روش اندازه‌بندی ابعاد پنجره و فوق‌دهی مدل هدف در ردیاب انتقال متوسط، توانستیم روش جدیدی از ردیابی همراه با اندازه‌بندی را ارائه کنیم، به‌علاوه این‌که مشکل این روش را نیز در رویارویی با هدف با کنتراست متغیر حل کردیم. از جمله مزایای روش پیشنهادی علاوه بر اندازه‌بندی سریع، نرم و پایدار، تنظیم ابعاد پنجره ردیابی قبل از اجرای ردیاب انتقال متوسط و اعمال این اندازه‌بندی به این ردیاب می‌باشد، به‌گونه‌ای که استفاده از این ایده به‌گونه‌ی مؤثری باعث حذف آثار پیکسل‌های زمینه در تشکیل مدل هدف خواهد شد که نتیجه‌ی آن ردیابی سریع، مقاوم و قدرتمند هدف خواهد بود.

آزمایشات ما نشان می‌دهد که طراحی الگوریتم پیشنهادی با افزودن مؤلفه‌هایی همچون اندازه‌بندی به ردیاب استاندارد زمان اجرا را تا حدود ۱۱٪ نسبت به ردیاب استاندارد و ۵٪ نسبت به ردیاب

تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک

فرهاد صمدزادگان*
حدیثه سادات حسنی**

* استاد، دانشکده مهندسی نقشه‌برداری، دانشگاه تهران، تهران

** دانشجوی دکتری فتوگرامتری، دانشکده مهندسی نقشه‌برداری، دانشگاه تهران، تهران

تاریخ دریافت: ۱۳۹۱/۰۴/۰۶

تاریخ پذیرش: ۱۳۹۱/۱۱/۰۹

چکیده

امروزه تصاویر فرا طیفی به علت غنای اطلاعات طیفی یک ابزار قوی و کارآمد در سنجش از دور به حساب می‌آیند و امکان تمایز بین عوارض مشابه را فراهم می‌آورند. با توجه به پایداری ماشین‌های بردار پشتیبان در فضاهایی با ابعاد بالا، یک گزینه مناسب در طبقه‌بندی تصاویر فرا طیفی محسوب می‌شوند. با این وجود، عملکرد این طبقه‌بندی کننده‌ها تحت تأثیر پارامترها و فضای ویژگی ورودی آن‌ها می‌باشد. به منظور استفاده از ماشین‌های بردار پشتیبان با بیشترین کارایی، می‌بایست مقادیر بهینه پارامترها و همچنین زیر مجموعه بهینه از ویژگی‌های ورودی تعیین گردند. در این تحقیق از توانایی الگوریتم ژنتیک به عنوان یک تکنیک بهینه‌سازی فرا ابتکاری، در تعیین مقادیر بهینه پارامترهای ماشین‌های بردار پشتیبان و همچنین انتخاب زیرمجموعه ویژگی‌های بهینه در طبقه‌بندی تصاویر فرا طیفی استفاده شده است. نتایج عملی از به‌کارگیری روش فوق در خصوص داده‌های فرا طیفی سنجنده AVIRIS نشان می‌دهند، ویژگی‌های ورودی و پارامترها هر کدام جداگانه تأثیر بسزایی بر عملکرد ماشین‌های بردار پشتیبان دارند ولی بهترین عملکرد طبقه‌بندی کننده با حل همزمان آن دو بدست می‌آید. در حل همزمان تعیین پارامتر و انتخاب ویژگی، برای کرنل گوسین و پلی‌نومیل به ترتیب ۵٪ و ۱۵٪ افزایش دقت با حذف بیش از نیمی از باندهای تصویر حاصل شد. همچنین الگوریتم بهینه‌سازی شبیه‌سازی تبرید تدریجی به منظور مقایسه با الگوریتم ژنتیک پیاده‌سازی شد که نتایج حاکی از برتری الگوریتم ژنتیک به ویژه با بزرگ و پیچیده شدن فضای جستجو در رویکرد حل همزمان تعیین پارامتر و انتخاب ویژگی می‌باشد.

۱. مقدمه

امروزه با پیشرفت‌های اخیر در زمینه تکنولوژی ساخت سنجیده‌های فرا طیفی در سنجش از دور، امکان اخذ تصاویر با صدها باند با قدرت تفکیک طیفی بالا فراهم است [۱]. افزایش تعداد باندها و در نتیجه افزایش اطلاعات طیفی، امکان استخراج اطلاعات بیشتر در مورد عوارض موجود در تصویر و همچنین تمایز بین عوارض مشابه را فراهم می‌کند. از سوی دیگر با بالا رفتن ابعاد فضای ورودی، نیاز به پردازش‌های خاصی به منظور به‌کارگیری این تصاویر در بسیاری از کاربردهای مطرح می‌باشد [۱].

یک مرحله مهم در پردازش تصاویر فرا طیفی به منظور استخراج اطلاعات، طبقه‌بندی آنها با یکسری کلاس‌های از پیش تعیین شده می‌باشد. در تصاویر فرا طیفی، طبقه‌بندی کننده‌های پارامتریک به علت نیاز به تخمین توزیع آماری کلاس‌ها و عدم توازن بین تعداد باندها و تعداد نمونه‌های آموزشی، با پدیده هیوز^۱ مواجه می‌شوند [۲]. در این شرایط به علت بزرگ شدن فضای فرضیه و محدودیت تعداد نمونه‌های آموزشی، احتمال بیش تطابق نمودن^۲ مدل به داده‌های آموزشی وجود دارد [۳]. به منظور رفع مشکلات روش‌های پارامتریک در سال‌های اخیر طبقه‌بندی کننده‌های غیر پارامتریک مختلفی به منظور طبقه‌بندی تصاویر فرا طیفی ارائه شده‌اند، از جمله: شبکه عصبی [۴]، درخت تصمیم‌گیری [۵]، Random Forest [۶] و روش‌های کرنل مبنا [۳].

در این خصوص ماشین‌های بردار پشتیبان^۳ به عنوان یک روش کرنل مبنا، توانسته‌اند به با توجه به عملکردشان بر مبنای ویژگی‌های

کلید واژگان: ماشین‌های بردار پشتیبان، تصاویر فرا طیفی، طبقه

بندی، انتخاب مدل، انتخاب ویژگی، الگوریتم ژنتیک.

¹ Hughes phenomena

² Over-fitting

³ Support Vector Machines

بودن و اجرای همزمان جستجوی سراسری^۲ و جستجوی محلی^۳ پتانسیل بالایی در حل مسائل پیچیده بهینه‌سازی دارد [۲۱]. در این تحقیق توانایی الگوریتم ژنتیک در تعیین طبقه‌بندی کننده‌های بردار پشتیبان در سه رویکرد: تعیین مقادیر بهینه پارامترهای ماشین‌های بردار پشتیبان، انتخاب زیرمجموعه ویژگی‌های بهینه و حل همزمان هر دو مفهوم در طبقه‌بندی تصاویر فرا طیفی مورد ارزیابی قرار گرفته است.

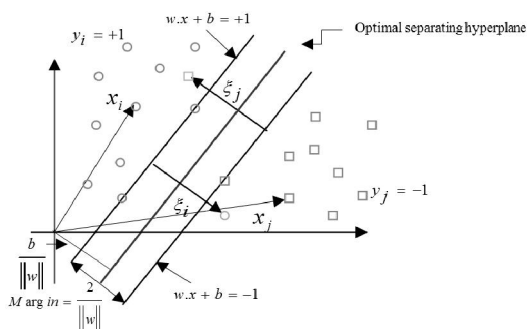
۲. ماشین‌های بردار پشتیبان

ماشین‌های بردار پشتیبان یک روش طبقه‌بندی با نظارت بر مبنای نظریه یادگیری آماری می‌باشند [۲۲]. ایده اساسی این طبقه‌بندی کننده، یافتن یک ابر صفحه بهینه به عنوان سطح تصمیم‌گیری به گونه‌ای می‌باشد که حاشیه بین دو کلاس را بیشینه کند. در صورتی که داده‌ها به صورت خطی جدا پذیر نباشد، داده‌ها با کرنلی غیرخطی به فضای با ابعاد بالاتر منتقل می‌شود و ابر صفحه بهینه در آن فضا تعیین می‌شود.

فرض کنید l داده‌های آموزشی موجود می‌باشد که هر یک با (x_i, y_i) نشان داده می‌شود، بردار ویژگی n بعدی و $y_i \in \{-1, 1\}$ برچسب آن می‌باشد. هدف یافتن ابر صفحه است که دو کلاس با برچسب ۱ و -۱ را با بیشترین حاشیه از هم جدا کند. این ابر صفحه را می‌توان با معادله (۱) بیان کرد.

$$w^T \phi(x) + b = 0 \quad (1)$$

در این رابطه، بردار وزن w ، برداری عمود بر ابر صفحه، b بردار بایاس می‌باشد که به منظور اندازه‌گیری فاصله ابر صفحه تا مبدأ استفاده می‌شود و $\phi(\cdot)$ کرنلی برای انتقال داده به فضای با ابعاد بالاتر می‌باشد (شکل ۱).



شکل ۱- طبقه‌بندی داده‌هایی که به صورت خطی جدا پذیر نیستند، توسط ماشین‌های بردار پشتیبان

هندسی و عدم نیاز به تخمین ویژگی‌های آماری، ابزاری بسیار کارا و قدرتمند برای طبقه‌بندی تصاویر فرا طیفی به حساب آیند [۷-۹]. ایده ماشین‌های بردار پشتیبان یافتن یک صفحه تصمیم‌گیری بهینه برای جداسازی دو کلاس می‌باشد به صورتی که دو کلاس بیشترین حاشیه جداسازی را در یک طبقه‌بندی باینری داشته باشند. در صورتی که نمونه‌ها به صورت خطی جدا پذیر نباشند، ابتدا با یک کرنل به فضایی با ابعاد بالاتر منتقل می‌شود و صفحه جداکننده در آن فضا تعریف می‌شود [۱۰].

هر چند در سال‌های اخیر ماشین‌های بردار پشتیبان با موفقیت در طبقه‌بندی بسیاری از تصاویر فرا طیفی به کار گرفته شده‌اند، با این حال می‌بایست به تأثیر دو عامل که بر عملکرد آن‌ها تأثیرگذار هستند، توجه داشت: پارامترهای ماشین‌های بردار پشتیبان و فضای ویژگی ورودی (باندهای ورودی طبقه‌بندی). به منظور طراحی یک سیستم بهینه طبقه‌بندی برای تصاویر فرا طیفی بر مبنای ماشین‌های بردار پشتیبان، انتخاب مقادیر بهینه پارامترهای مدل و انتخاب زیر مجموعه‌ای از باندهای بهینه چالش‌هایی است که در این زمینه وجود دارند. به علت اهمیت موضوع، مطالعات زیادی در سال‌های اخیر در این زمینه انجام شده است که می‌توان آن‌ها را به سه دسته تقسیم کرد. در دسته اول، از همه ویژگی‌های ورودی داده استفاده می‌شود و پارامترهای مدل بهینه می‌گردد تا کارایی ماشین‌های بردار پشتیبان بالا رود [۱۱-۱۳]. انتخاب ویژگی‌های بهینه با ثابت در نظر گرفتن پارامترها گروه دوم تحقیقات می‌باشد [۱۴-۱۶] که با حذف ویژگی‌های اضافی و وابسته، دقت و سرعت طبقه‌بندی را بهبود می‌بخشند. در این دسته از مطالعات، در ابتدا مقادیر پارامترهای مدل با استفاده از یک روش کلاسیک محاسبه شده و یا از مقادیر پیش فرض استفاده می‌کنند. سپس در پروسه انتخاب ویژگی آن مقادیر ثابت در نظر گرفته می‌شوند. با توجه به تأثیر فضای ورودی بر مقدار بهینه پارامترها و بالعکس، دسته سوم الگوریتم‌ها به حل همزمان تعیین پارامترها و انتخاب زیر مجموعه ویژگی‌های بهینه می‌پردازند [۱۷-۲۰].

به منظور حل این مسائل، با توجه به بزرگ بودن فضای مجهولات نیاز به یک الگوریتم بهینه‌سازی قوی و کارا می‌باشد تا بتواند جواب بهینه سراسری را بدست آورد. از آنجایی که الگوریتم‌های بهینه سازی مرسوم معمولاً در فضای جستجوی بزرگ با مشکل مواجه می‌شوند و به جواب زیر بهینه می‌رسند، در این تحقیق از الگوریتم فرا ابتکاری^۱ ژنتیک استفاده گردیده است. الگوریتم ژنتیک به علت جمعیت مینا

² Exploration
³ Exploitation

¹ Meta-heuristic

در روش «یک در مقابل یک»، برای هر زوج کلاس ممکن از یک ماشین بردار پشتیبان باینری استفاده می‌شود. به این ترتیب برای M کلاس، $\frac{M(M-1)}{2}$ طبقه‌بندی کننده بردار پشتیبان باینری نیاز داریم. در نهایت همه ماشین‌های بردار پشتیبان باینری با روش رای گیری حداکثر ادغام می‌شود.

روش «یک در مقابل مابقی»، روش مرسوم دیگر می‌باشد که در آن هر ماشین بردار پشتیبان باینری، داده‌های یک کلاس را از داده‌های کلاس‌های دیگر جدا می‌کند. در این روش، برای M کلاس، M طبقه‌بندی کننده باینری خواهیم داشت. پس از طبقه‌بندی داده جدید با M طبقه‌بندی کننده، داده به کلاسی که بیشترین نتیجه مثبت را داشته باشد، نسبت داده می‌شود [۱۰].

۳. مروری بر تحقیقات

با توجه به اطلاعات طیفی غنی موجود در تصاویر فرا طیفی، همراه با پیشرفت تکنولوژی ساخت سنجیده‌های فرا طیفی، الگوریتم‌های پردازش این دسته از تصاویر نیز رشد چشمگیری داشته‌اند که در این بین الگوریتم‌های طبقه‌بندی به منظور استخراج اطلاعات و همچنین تهیه نقشه‌های موضوعی بسیار مورد توجه قرار گرفته‌اند. از این رو مطالعات گسترده‌ای در زمینه ارزیابی عملکرد طبقه‌بندی کننده‌های گوناگون در تصاویر فرا طیفی صورت گرفته است. از اوایل سال ۲۰۰۰، نخستین تحقیقات بر مبنای استفاده از ماشین‌های بردار پشتیبان در طبقه‌بندی تصاویر فرا طیفی صورت پذیرفت. در سال‌های بعد عملکرد طبقه‌بندی کننده‌های دیگر از قبیل شبکه‌های عصبی، نزدیک‌ترین همسایه، کمترین فاصله، بیشترین شباهت، درخت تصمیم‌گیری و آنالیز تفکیک کرنل فیشر در مقایسه با ماشین‌های بردار پشتیبان بررسی شد که نتایج حاکی از برتری ماشین‌های بردار پشتیبان می‌باشد [۳، ۴، ۵، ۶]. از این رو در این تحقیق از طبقه‌بندی کننده ماشین‌های بردار پشتیبان به منظور پردازش تصاویر فرا طیفی استفاده شده است.

هرچند ماشین‌های بردار پشتیبان به عنوان ابزاری کارا در طبقه‌بندی تصاویر فرا طیفی به کار برده شده‌اند، به منظور بهبود عملکرد این طبقه‌بندی کننده به لحاظ دقت و پیچیدگی‌های محاسباتی، می‌بایست به دو عامل تأثیرگذار توجه شود: پارامترهای طبقه‌بندی کننده و باندهای ورودی آن. از این رو مطالعات گوناگونی بر مبنای بهینه‌سازی سیستم طبقه‌بندی تصاویر فرا طیفی ارائه شده‌اند. تحقیقات موجود در این زمینه را می‌توان در سه بخش انتخاب مدل، انتخاب ویژگی و حل همزمان این دو دسته‌بندی کرد.

انتخاب بهینه پارامترهای تشکیل دهنده مدل در ماشین‌های بردار پشتیبان از تأثیر مستقیمی در کارایی این روش برخوردار می‌باشند

بیشینه نمودن حاشیه بین دو کلاس معادل کمینه کردن اندازه w می‌باشد که منجر به حل مسئله کمینه‌سازی مقید (۲) می‌شود.

$$\text{Min} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^k \xi_i \right) \quad (2)$$

$$\text{Subject to: } y_i (w\phi(x_i) + b) > 1 - \xi_i \quad i=1, \dots, k$$

که پارامتر C ، پارامتر تنظیم در ماشین‌های بردار پشتیبان می‌باشد. به منظور در نظر گرفتن نویز موجود در داده و تداخل بین داده‌های آموزشی، از متغیر $\xi_i > 0$ استفاده می‌شود. وجود قید ضمانت می‌کند که داده‌ای در حاشیه قرار نمی‌گیرد. هرچند برای جلوگیری از بیش تطابق نمودن به داده‌های نویزی، این قید با متغیرهای ξ_i نرم شده است.

سطح تصمیم‌گیری بهینه با حل مسئله مقید (۲) بر مبنای روش لاگرانژ طبق معادله (۳) محاسبه می‌شود.

$$f(x) = \sum_{x_i \in SV} y_i \alpha_i \phi(x_i) \cdot \phi(x) + b \quad (3)$$

در این رابطه α_i ضرایب لاگرانژ می‌باشد که در پروسه بهینه‌سازی محاسبه می‌شود، SV بردارهای پشتیبان هستند که ضریب لاگرانژ متناظر آن‌ها بزرگ‌تر از صفر است. این داده‌های آموزشی، نزدیک‌ترین نمونه‌ها به ابر صفحه هستند. همان طور که رابطه (۳) نشان می‌دهد، تنها بردارهای پشتیبان هستند که در مرحله آموزش شرکت می‌کنند. در نتیجه ماشین‌های بردار پشتیبان نیاز به تعداد نمونه آموزشی زیاد ندارند. در رابطه (۳)، ضرب داخلی بین دو کرنل نگاشت شده، می‌تواند با کرنل آن دو نمونه محاسبه گردد. از پرکاربردترین کرنل‌ها، کرنل گوسین و پلی‌نومینال هستند که به ترتیب با روابط (۴) و (۵) تعریف می‌شوند.

$$\phi(x_i, x) = e^{-\frac{\|x_i - x\|^2}{2\sigma^2}} \quad (4)$$

$$\phi(x_i, x) = (x_i \cdot x + 1)^d \quad (5)$$

در این روابط، σ پارامتر کرنل گوسین و d متغیر کرنل پلی‌نومینال می‌باشد [۲۳].

الگوریتم پایه ماشین‌های بردار پشتیبان برای طبقه‌بندی باینری توسعه داده شده است. از آنجایی که در بیشتر کاربردها، بیش از دو کلاس وجود دارد، الگوریتم‌های مختلفی برای حل مسئله چند کلاسه به کار گرفته شده است [۱۰]. یک روش مرسوم در این زمینه تجزیه مسئله چند کلاسه به مسئله‌ای با چندین طبقه بندی کننده باینری می‌باشد. الگوریتم‌های «یک در مقابل یک» و «یک در مقابل مابقی»، دو الگوریتم پرکاربرد در این زمینه می‌باشند.

انتخاب پارامترهای بهینه مدل در ماشین‌های بردار پشتیبان استفاده شده‌اند [۲۳، ۲۶]. الگوریتم ژنتیک به دو شکل در انتخاب مدل می‌تواند استفاده گردد: کدگذاری اعداد حقیقی [۱۳، ۲۳] و کدگذاری باینری [۱۵، ۲۶].

دسته دیگر مطالعات در زمینه بهینه‌سازی سیستم‌های طبقه‌بندی، به مسئله انتخاب زیرمجموعه بهینه باندهای ورودی می‌پردازند. الگوریتم‌های انتخاب ویژگی را می‌توان به دو دسته پوششی^۱ و فیلتر^۲ تقسیم کرد. تکنیک‌های پوششی به منظور ارزیابی کیفیت زیرمجموعه ویژگی‌های انتخاب شده، از دقت طبقه‌بندی کننده استفاده می‌کنند. در مقابل الگوریتم‌های فیلتر، از معیارهای دیگری مانند جداسازی بین کلاس‌ها، قدرت نمایش و غیره استفاده می‌کنند. این دسته از تکنیک‌ها دارای سرعت بالایی هستند ولی لزوماً دارای دقت طبقه‌بندی مناسبی نیستند [۱۴]. با توجه به اهمیت دقت طبقه‌بندی در این تحقیق از روش‌های پوششی که در آن کیفیت باندهای انتخابی به وسیله دقت طبقه‌بندی کننده ماشین‌های بردار پشتیبان بدست می‌آید، استفاده خواهد شد.

با توجه به ابعاد بالای ورودی در تصاویر فرا طیفی، انتخاب یک زیر مجموعه بهینه از باندها بدون پیش فرضی در مورد تعداد باندهای بهینه یک مسئله بهینه‌سازی NP-hard می‌باشد [۲۹]. در این راستا در سال‌های اخیر الگوریتم‌های بهینه‌سازی مختلفی از قبیل جستجوی ممنوع^۳ [۳۰]، الگوریتم‌های خرد جمعی [۳۱، ۳۲] و الگوریتم ژنتیک [۱۴، ۱۶، ۳۳] برای حل این موضوع ارائه گردیده است. با توجه به ساختار الگوریتم ژنتیک در حالت باینری، مسئله انتخاب ویژگی سازگاری بسیار مناسبی با الگوریتم ژنتیک دارد. لذا تحقیقات بسیاری در رابطه با انتخاب ویژگی بر مبنای الگوریتم ژنتیک در زمینه‌های گوناگون و با طبقه‌بندی کننده‌های مختلف وجود دارد [۱۴، ۳۳].

در سال ۲۰۰۶ هوئرتا و همکاران از الگوریتم ژنتیک به منظور کاهش ابعاد داده‌های پزشکی استفاده کردند. در این تحقیق در ابتدا در مرحله پیش پردازش با منطبق فازی ابعاد فضای ورودی را کاهش داده و سپس با الگوریتم پوششی بر مبنای الگوریتم ژنتیک در فضای کاهش یافته، ویژگی‌های بهینه انتخاب گردید. همچنین پارامترهای ماشین‌های بردار پشتیبان ثابت در نظر گرفته شد [۱۴].

در سال ۲۰۰۸، بهاتیا و همکاران از الگوریتم ژنتیک که با اعداد صحیح کدگذاری شده بود به منظور انتخاب ویژگی با طبقه‌بندی کننده ماشین‌های بردار پشتیبان استفاده کردند [۳۴]. در الگوریتم ارائه شده،

[۱۳]. دو دسته پارامتر در این طبقه‌بندی کننده‌ها وجود دارد. «پارامتر تنظیم» که تعادل بین کمینه شدن خطا و کمینه شدن پیچیدگی‌های مدل را برقرار می‌کند (پارامتر C) و «پارامترهای کرنل» که با توجه به کرنل انتخابی، متغیرهای آن کرنل جزء مجهولات می‌شوند. مانند σ در کرنل گوسین و d در کرنل پلی نومیال [۱۳].

همان طور که در بخش قبل بیان شد، ماشین‌های بردار پشتیبان ذاتاً باینری هستند و برای مسائل چند کلاسه از ترکیب چندین طبقه‌بندی کننده باینری استفاده می‌شود. از این نقطه نظر می‌توان الگوریتم‌های موجود در تعیین پارامترهای ماشین‌های بردار پشتیبان که انتخاب مدل نیز نامیده می‌شوند، را به دو دسته تقسیم کرد. در دسته اول تکنیک‌های انتخاب مدل، برای همه‌ی ماشین‌های بردار پشتیبان باینری یک دسته پارامتر یکسان در نظر گرفته می‌شود [۱۳، ۲۶]. در حالی که در دسته دوم، برای هر طبقه‌بندی کننده باینری یکسری پارامتر متفاوت تعیین می‌شود [۲۳، ۲۷]. تحقیقات انجام شده نشان داده است که با اضافه شدن مجهولات در اکثر موارد نه تنها دقت طبقه‌بندی بالا نمی‌رود بلکه به دلیل بیش تطابق نمودن به داده‌های آموزشی، سیستم طبقه‌بندی دارای قدرت تعمیم کمی خواهد بود [۲۷]. در نتیجه در این تحقیق از یک دسته پارامتر برای همه ماشین‌های بردار پشتیبان باینری استفاده می‌شود.

در سال‌های اخیر روش‌های متنوعی به منظور تعیین پارامترهای بهینه در ماشین‌های بردار پشتیبان، توسط محققین مختلف ارائه گردیده است [۱۱-۱۳]. الگوریتم جستجوی شبکه‌ای، روش رایج در انتخاب پارامترهای بهینه مدل می‌باشد. در این روش، شبکه‌ای k بعدی بر روی محدوده پارامترها قرار می‌گیرد که k تعداد پارامترهای مجهول می‌باشد (در این تحقیق برای به کارگیری هر دو کرنل گوسین و پلی-نومیال $k=2$ در نظر گرفته شده است). سپس کیفیت تمام مجموعه جواب‌های ممکن در نقاط شبکه ارزیابی می‌شود و آن دسته از پارامترهایی که کمترین خطای طبقه‌بندی را دارند، به عنوان پارامترهای ماشین‌های بردار پشتیبان انتخاب می‌شوند [۲۸]. به علت پیوسته بودن مقادیر پارامترهای مورد نظر، برای رسیدن به دقت بالا می‌بایست شبکه‌ای با تراکم بالا در نظر گرفته شود که بررسی تمام این نقاط شبکه زمان محاسبات را به شدت افزایش می‌دهد.

با توجه به محدودیت‌های روش جستجوی شبکه‌ای در زمان و محاسبات، الگوریتم‌های بهینه‌سازی مختلفی برای حل این مسئله در نظر گرفته شده‌اند: الگوریتم‌های خرد جمعی [۱۲، ۲۷]، شبیه‌سازی تبرید تدریجی [۱۸] و الگوریتم ژنتیک [۱۳، ۲۳، ۲۶].

در این بین، الگوریتم ژنتیک از الگوریتم‌های فرا ابتکاری هستند که به طور موفقیت آمیز و گسترده‌ای در سال‌های اخیر در زمینه

^۱Wrapper

^۲Filter

^۳Tabu Search

استفاده از توابعی همچون سیگموئید به بازه $[0,1]$ منتقل می‌شود و برای ارزیابی ویژگی‌های انتخاب شده، این ویژگی‌ها به فضای باینری منتقل می‌شوند [۲۰، ۳۸]. در روش دوم پارامترهای ماشین‌های بردار پشتیبان، تبدیل به فرمت باینری شده و در کنار بردار ویژگی‌های انتخابی قرار می‌گیرند [۲۰].

در سال ۲۰۰۵، سوزا و کاروالو مقدار پارامتر تنظیم و زیرمجموعه بهینه ویژگی‌ها را بر مبنای الگوریتم ژنتیک باینری تعیین نمودند و مقدار پارامتر کرنل در این روش ثابت در نظر گرفته شد. همچنین در این تحقیق تنها سه مقدار ممکن برای پارامتر تنظیم در نظر گرفته شده است [۳۹].

در سال ۲۰۰۶، هوانگو وانگ پارامترهای کرنل RBF به همراه ویژگی‌های ورودی در حالت باینری بهینه گردید. نتایج بدست آمده در این روش با تکنیک جستجوی شبکه‌ای که قابلیت انتخاب ویژگی ندارد، مقایسه گردیده است که نشان می‌دهد، الگوریتم پیشنهادی با ویژگی‌های انتخاب شده به دقت بالاتری می‌رسد [۲۰]. در همین سال بگی و ملگانی الگوریتم ژنتیک را با کروموزم‌هایی به طول تعداد ویژگی‌ها به علاوه ۲ (دو پارامتر کرنل) پیاده سازی کردند. برای ارزیابی الگوریتم ارائه شده، از تصاویر فرا طیفی استفاده کردند [۷].

در سال ۲۰۰۷، مِجیا-گوئواراو اکوری-مولارساز ماشین‌های بردار پشتیبان برای رگرسیون استفاده کردند که در این حالت سه پارامتر مجهول برای ماشین‌های بردار پشتیبان وجود دارد. در این مقاله علاوه بر ویژگی و پارامترها، احتمال ترکیب و جهش در الگوریتم ژنتیک نیز به عنوان مجهول در نظر گرفته شده است. برای نمایش راه حل با کروموزم، از کدگذاری باینری استفاده شد [۳۶].

ژو و همکاران در سال ۲۰۰۸ الگوریتم ارائه شده در [۲۰] را بر روی تصاویر فرا طیفی هایپریون آزمایش کردند و مشاهده گردید نتایج بدست آمده نسبت به زمانی که از همه باندها استفاده شده باشد، بهبود یافته است [۱۵]. در این مطالعه تأثیر روش پیشنهادی تنها با استفاده از یک کرنل و یک روش چند کلاسه در ماشین‌های بردار پشتیبان بررسی شد.

در سال ۲۰۰۹، اکوی الگوریتمی ارائه داد که علاوه بر ویژگی و پارامترها، نوع کرنل نیز در یک پروسه مبتنی بر الگوریتم ژنتیک بهینه می‌گردید. در این مقاله، از انتقال موجک برای استخراج ویژگی استفاده گردیده و برای این منظور ۱۶ نوع موجک و ۸ نوع کرنل در نظر گرفته شده است. بعد از پایان الگوریتم، نوع موجک و کرنل بهینه به همراه پارامترهای آن بدست می‌آید [۳۸].

در سال ۲۰۱۱، ژائو و همکاران از الگوریتم ژنتیک در بهینه‌سازی ماشین‌های بردار پشتیبان به منظور طبقه‌بندی داده‌های UCI استفاده

طول کروموزم به تعداد ویژگی‌هایی است که انتخاب خواهد شد و درایه‌های کروموزم شماره‌های ویژگی‌های انتخاب شده می‌باشد. پارامترهای ماشین‌های بردار پشتیبان نیز قبل از انتخاب ویژگی با بررسی چندین مقدار و انتخاب بهترین آن‌ها محاسبه شده است [۳۴]. محدودیت این روش نیاز به دانستن تعداد ویژگی‌های بهینه به عنوان اطلاعات از پیش تعریف شده در مسئله می‌باشد. از آنجایی که این پارامتر از قبل مشخص نیست، به ازای مقادیر مختلف آزمایش تکرار شده و زیر مجموعه ویژگی‌ها با بهترین دقت طبقه‌بندی انتخاب شد که البته این روش در فضای با ابعاد بالا امکان پذیر نیست.

در سال ۲۰۰۹، ژانگ و همکاران الگوریتمی برای سرعت بخشیدن به انتخاب ویژگی بر مبنای الگوریتم ژنتیک ارائه دادند [۱۶]. در این مقاله برای ارزیابی هر عضو در الگوریتم ژنتیک، ابتدا پارامترهای ماشین‌های بردار پشتیبان با الگوریتم جستجوی شبکه‌ای محاسبه و سپس با ارزیابی ۱۰ قسمتی، دقت محاسبه می‌شود. به علت حجم محاسباتی بالای این الگوریتم از تکنیک محاسباتی موازی سازی^۱ در الگوریتم ژنتیک و ماشین‌های بردار پشتیبان استفاده شده است [۱۶].

در سال ۲۰۱۲، سنتوس و همکاران از الگوریتم ژنتیک به منظور انتخاب ویژگی بر اساس روش فیلتر استفاده کردند و سه روش طبقه‌بندی ماشین‌های بردار پشتیبان، نزدیک‌ترین همسایگی و شبکه‌های عصبی به منظور ارزیابی نتایج پیاده‌سازی شد [۳۵].

از آنجایی که پارامترهای ماشین‌های بردار پشتیبان دارای تأثیر متقابل بر زیر مجموعه ویژگی‌های مورد استفاده در طبقه‌بندی است، راه حل بهینه در انجام طبقه‌بندی حل همزمان پارامترها و فضای ویژگی می‌باشد [۲۰]. با این وجود با بزرگ شدن فضای جستجو و پیچیده شدن آن، نیاز به استفاده از تکنیک‌های بهینه‌سازی قدرتمندی در این وضعیت مطرح می‌گردد. در سال‌های اخیر الگوریتم‌های فرا ابتکاری متنوعی برای حل این مسئله مانند الگوریتم‌های خرد جمعی [۱۷، ۱۹]، شبیه‌سازی تبرید تدریجی [۱۸] و الگوریتم ژنتیک [۱۵، ۲۰، ۳۶، ۳۷] توسط محققین مختلف ارائه شده است.

در حل همزمان انتخاب مدل و ویژگی توجه به ماهیت متفاوت این دو مسئله ضروری می‌باشد که در آن پارامترهای ماشین‌های بردار پشتیبان دارای ماهیتی پیوسته است، برخلاف ویژگی‌های ورودی که دارای ماهیتی گسسته می‌باشد. به منظور ادغام این دو مفهوم، دو راه حل در مطالعات ارائه شده است [۲۰، ۳۸]. در روش اول بردار ویژگی‌ها نیز در کنار پارامترها به صورت پیوسته مدل‌سازی شده و با

^۱ Parallelization

گردان^۱ کروموزم‌های منتخب مشخص و وارد نسل بعد می‌شود. انتقال نخبه^۲ هر مرحله که مستقیماً به مرحله بعد منتقل می‌شود نیز به منظور حفظ نتایج مطلوب در نظر گرفته شده است.

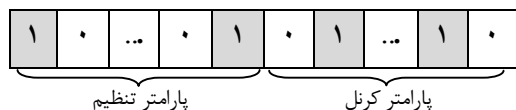
یک تکنیک رایج مرحله تلفیق^۳، روش تک نقطه‌ای^۴ می‌باشد که در این تحقیق نیز مورد استفاده قرار گرفته است. تلفیق تک نقطه‌ای توسط انتخاب یک نقطه به صورت تصادفی در کروموزم و تعویض اطلاعات دو کروموزم والد از نقطه مشخص شده، صورت می‌پذیرد.

به منظور ارزیابی مناطق جدید در فضای جستجو، از جهش^۵ استفاده می‌شود تا الگوریتم قابلیت جستجوی تصادفی را نیز داشته باشد. عملگر جهش باینری^۶ با تغییر مقدار درایه‌هایی (۰ را به ۱ و بالعکس) که احتمال آنها از احتمال جهش بیشتر شده است، عمل می‌کند. این پروسه تکرار می‌شود تا شرط توقف (حداکثر تکرار و یا عدم تغییر بهترین مقدار در طی تکرارهای مشخص) برقرار شود [۲۱].

با توجه به انعطاف پذیری الگوریتم ژنتیک نسبت به دیگر روش‌های دیگر در زمینه بهینه‌سازی فرا ابتکاری و حساسیت کمتر آن به تعریف پارامترهای مبنایی، در این تحقیق از الگوریتم ژنتیک به منظور انتخاب مدل، انتخاب ویژگی و حل همزمان این دو مسئله استفاده گردید.

۴.۱. انتخاب مدل

در روش پیشنهادی این تحقیق، با در نظر گرفتن عملگرهای معمول در کدگذاری باینری و همچنین ذات باینری قسمت‌های بعدی، از کدگذاری باینری استفاده شده است. همچنین هر کروموزم نمایشگر دو پارامتر تنظیم و کرنل می‌باشد که با یک رشته از صفر و یک تعریف می‌شود (شکل ۲). طول کروموزم متناسب با محدوده تغییرات پارامترها و همچنین دقت مورد نیاز مسئله تعیین می‌گردد.



شکل ۲- نمایش باینری پارامترهای ماشین‌های بردار پشتیبان

برای ارزیابی کیفیت هر عضو، می‌بایست هر یک از دو قسمت کروموزم به عدد حقیقی تبدیل شود. برای این منظور، از رابطه (۶) استفاده می‌کنیم.

کردند [۴۰]. در همین سال چن و همکاران در پیش‌بینی ورشکستگی بانک‌ها از الگوریتم بهینه‌سازی ژنتیک به منظور بهبود عملکرد سیستم تصمیم‌گیری بر مبنای انتخاب زیرمجموعه ویژگی‌های بهینه و تعیین پارامترهای شبکه عصبی (پیش‌بینی کننده) استفاده شد [۴۱].

۴. تعیین طبقه‌بندی کننده بهینه ماشین‌های بردار پشتیبان

بر مبنای الگوریتم ژنتیک

در سال‌های اخیر ماشین‌های بردار پشتیبان به عنوان یکی از کاراترین طبقه‌بندی کننده‌های پایدار در تصاویر فرا طیفی مطرح شده‌اند [۳]. یکی از عوامل موثر بر عملکرد ماشین‌های بردار پشتیبان، پارامترهای آن می‌باشند [۲۳]. انتخاب زیر مجموعه بهینه به عنوان ورودی طبقه‌بندی یک گام مهم دیگر در بهینه‌سازی ماشین‌های بردار پشتیبان به منظور طبقه‌بندی تصاویر فرا طیفی است. هر چند ماشین‌های بردار پشتیبان در فضای با ابعاد بالا پایدار می‌باشند ولی انتخاب زیرمجموعه بهینه از ویژگی‌ها می‌تواند با حذف ویژگی‌های اضافی و زائد عملکرد طبقه‌بندی کننده را از لحاظ دقت، سرعت و هزینه بهبود ببخشد [۱۶].

با توجه به عوامل تأثیرگذار بر عملکرد ماشین‌های بردار پشتیبان، به منظور تعیین یک سیستم طبقه‌بندی بهینه در فضای با ابعاد بالای تصاویر فرا طیفی، می‌بایست از یک تکنیک بهینه‌سازی قدرتمند بهره برد [۲۰]. الگوریتم‌های فرا ابتکاری، تکنیک‌های محاسباتی هستند که در یک پروسه تکراری با توجه به تابع هدفی که کیفیت راه‌حل را بیان می‌کند، راه‌حل‌های کاندید را بهبود می‌بخشند (بدون پیش فرضی راجع به مسئله) و معمولاً در فضای جستجوی بزرگ کارا عمل می‌کنند [۲۴]. در نتیجه این دسته الگوریتم‌ها گزینه‌ای مناسب در بهینه‌سازی سیستم طبقه‌بندی تصاویر فرا طیفی بر مبنای ماشین‌های بردار پشتیبان باشند.

الگوریتم ژنتیک یک نمونه از الگوریتم‌های فرا ابتکاری می‌باشد که با موفقیت در بسیاری از مسائل بهینه‌سازی در کاربردهای مختلف به کار گرفته شده است [۱۴، ۲۰، ۲۵]. این الگوریتم، روشی جمعیت مبنایی می‌باشد که در یک پروسه تکراری تکاملی کیفیت جمعیت را بهبود می‌بخشد. در این الگوریتم ابتدا می‌بایست راه‌حل را به صورت یک رشته باینری که کروموزم نامیده می‌شود، نمایش دهیم. در مرحله بعد جمعیت اولیه به صورت تصادفی ساخته و کیفیت آن به وسیله تابع هدف ارزیابی می‌شود.

سه عملگر اصلی در الگوریتم ژنتیک عبارتند از: انتخاب، تلفیق و جهش. پس از اندازه‌گیری کیفیت اعضاء، احتمال انتخاب هر عضو برای شرکت در مرحله تلفیق مشخص می‌شود و با استفاده از چرخ

¹ Roulette wheel

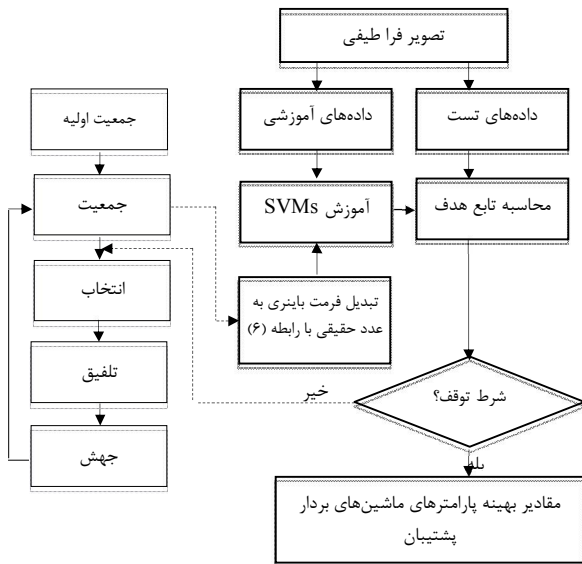
² Elite

³ Crossover

⁴ Single Point

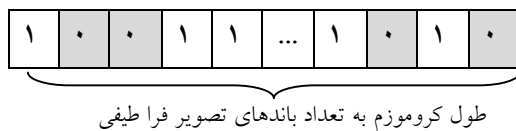
⁵ Mutation

⁶ Bit-Flip



شکل ۳- فلوجارت تعیین مقادیر بهینه پارامترهای ماشین‌های بردار پشتیبان بر مبنای الگوریتم ژنتیک

استفاده شد که در آن هر عضو جمعیت به وسیله یک رشته از ۰ و ۱ به طول تعداد باندهای تصویر فرا طیفی نمایش داده می‌شود (شکل ۴). در این کروموزم، مقدار بیت ۰ به معنای حذف باند متناظر و ۱ به معنای انتخاب آن می‌باشد.



شکل ۴- نمایش کروموزم به منظور انتخاب ویژگی بر مبنای الگوریتم ژنتیک

به منظور تعریف معیاری برای ارزیابی کیفیت زیر مجموعه ویژگی‌های انتخاب شده، می‌بایست دو پارامتر دقت طبقه‌بندی و تعداد ویژگی‌های انتخاب شده را در نظر گرفت. به عبارت دیگر، طبقه‌بندی بر مبنای زیرمجموعه ویژگی‌های مطلوب دارای دقت طبقه‌بندی بالاتر و تعداد ویژگی‌های انتخاب شده کمتر می‌باشد. لذا تابع هدف با قرار دادن این دو معیار در یک تابع تعریف می‌شود و مسئله به بیشینه سازی معادله (۸) تبدیل می‌شود.

$$f = w \times accuracy + (1 - w) \times \frac{1}{N_f} \quad (8)$$

در این رابطه، f مقدار تابع هدف، w یک کمیت ثابت در بازه [۰،۱] می‌باشد که وزن بین دقت طبقه‌بندی و تعداد ویژگی‌ها را مشخص می‌کند. همچنین $accuracy$ دقت طبقه‌بندی می‌باشد که با رابطه (۷) محاسبه می‌شود و N_f تعداد باندهای انتخاب شده می‌باشد. پس از ارزیابی اعضای جمعیت، سه عملگر انتخاب، تلفیق و جهش با توجه به کیفیت هر عضو عمل می‌کنند و مجدداً جمعیت

$$p = \min_p + \frac{\max_p - \min_p}{2^l - 1} \times d \quad (6)$$

که در این رابطه، p مقدار حقیقی پارامتر، \max_p و \min_p به ترتیب کمینه و بیشینه مقدار پارامتر، l تعداد بیت‌های نمایشگر پارامتر و d مقدار عددی رشته باینری در پایه ۱۰ می‌باشد.

در مرحله بعد ماشین‌های بردار پشتیبان به وسیله داده‌های آموزشی و پارامترهای محاسبه شده، آموزش دیده و ابر صفحه‌های مورد نظر ساخته می‌شوند. سپس به منظور محاسبه تابع هدف، داده‌های تست به وسیله ماشین‌های بردار پشتیبان آموزش دیده، طبقه‌بندی می‌شود و سپس ماتریس خطا تشکیل می‌شود. از ضرب کاپا به علت استفاده از تمام اطلاعات ماتریس خطا، به عنوان دقت طبقه‌بندی و تابع هدف در این بخش استفاده گردید که با رابطه (۷) تعریف می‌شود.

$$Kappa = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})} \quad (7)$$

در این رابطه، N تعداد کل نمونه‌ها، r تعداد کلاس‌ها، x_{ii} عناصر روی قطر اصلی ماتریس خطا، x_{i+} جمع حاشیه‌ای سطرها و x_{+i} جمع حاشیه‌ای ستون‌ها می‌باشد. پس از ارزیابی اعضا، سه مرحله انتخاب، تلفیق و جهش بر روی فرمت باینری پارامترها انجام می‌شود و جمعیت جدید ساخته می‌شود و این مراحل تکرار می‌شود تا شرط توقف برقرار شود (شکل ۳).

۴.۲. انتخاب ویژگی

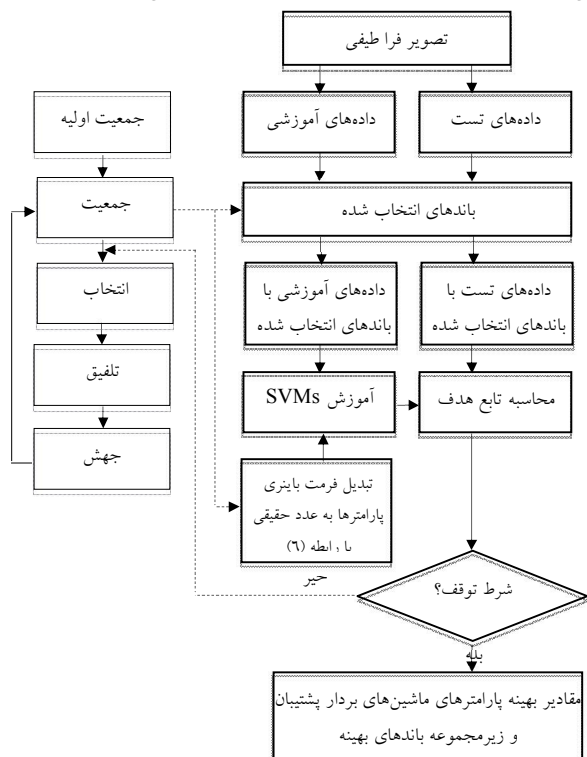
انتخاب ویژگی یکی از مراحل تأثیرگذار در طبقه‌بندی تصاویر فرا طیفی بر مبنای ماشین‌های بردار پشتیبان می‌باشد که در آن با حذف باندهای نامربوط و نویزی، عملکرد طبقه‌بندی کننده را از لحاظ دقت و سرعت بهبود می‌بخشد.

در روش پیشنهادی این تحقیق، انتخاب ویژگی بر مبنای الگوریتم ژنتیک با استفاده از تکنیک پوششی پیاده‌سازی شده است. در مرحله پیش پردازش روش پیشنهادی ابتدا با حضور همه باندها، مقادیر در پروسه اصلی انتخاب ویژگی‌های بهینه، از کدگذاری باینری پارامترهای ماشین‌های بردار پشتیبان با روش جستجوی شبکه‌ای تعیین و این مقادیر در طول پروسه انتخاب ویژگی ثابت در نظر گرفته شدند.

تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک

ساخته شده توسط تابع هدف ارزیابی خواهد شد و این پروسه تکرار می‌شود تا شرط توقف برقرار گردد (شکل ۵).

خطا تشکیل و دقت طبقه‌بندی با رابطه (۷) محاسبه می‌شود. سپس با توجه به قسمت اول هر عضو که بیانگر تعداد باندهای انتخابی است و دقت بدست آمده، مقدار تابع هدف با رابطه (۸) محاسبه می‌شود. در ادامه مشابه مسائل قبل، مراحل انتخاب، تلفیق و جهش انجام و این مراحل تکرار می‌شود تا شرط توقف برقرار شود (شکل ۷).



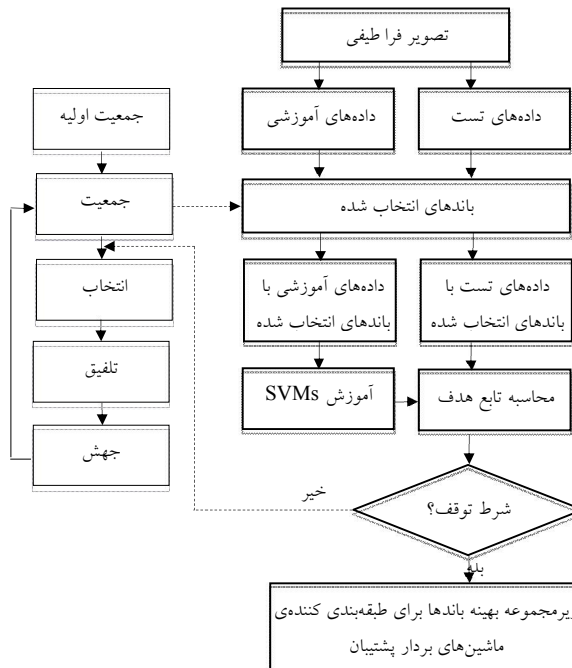
شکل ۷- فلوجارت تعیین مقادیر بهینه پارامترهای ماشین‌های بردار پشتیبان و زیرمجموعه باندهای بهینه بر مبنای الگوریتم ژنتیک

۵. نتایج عملی

به منظور ارزیابی توانایی روش‌های ارائه شده در این تحقیق، نسبت به پیاده‌سازی و به کارگیری آن‌ها در طبقه‌بندی تصویر فرا طیفی سنجنده AVIRIS اقدام گردید. نتایج حاصل در قالب سه گروه: انتخاب پارامترهای مدل، انتخاب ویژگی و حل همزمان هر دو شرح داده شده است. در ماشین‌های بردار پشتیبان از دو کرنل گوسین و پلی‌نومیال و دو روش چند کلاسه «یک در مقابل یک» و «یک در مقابل مابقی» استفاده شده است. به منظور مقایسه نتایج بدست آمده، الگوریتم شبیه‌سازی تبرید تدریجی نیز پیاده‌سازی شد.

۵.۱. مشخصات داده

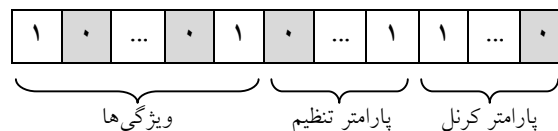
برای پیاده‌سازی الگوریتم‌های ارائه شده از تصویر فرا طیفی مربوط به منطقه ای کشاورزی-جنگلی که توسط سنجنده AVIRIS در سال ۱۹۹۲ از شمال شرقی ایالت ایندیانا اخذ شده، استفاده گردید. به دلیل شباهت بین کلاس‌ها، این تصویر دارای پیچیدگی‌هایی برای طبقه بندی است. این تصویر دارای ۲۲۰ باند می‌باشد که ۱۵ باند



شکل ۵- فلوجارت انتخاب زیر مجموعه بهینه باندها بر مبنای الگوریتم ژنتیک

۴.۳. حل همزمان پارامتر و ویژگی

با توجه به هدف این تحقیق که حل همزمان تعیین پارامتر و انتخاب ویژگی در این بخش می‌باشد، در روش پیشنهادی ویژگی‌ها و پارامترهای کرنل در کروموزم به صورت باینری کدگذاری می‌شود. هر کروموزم از سه قسمت تشکیل شده است: ویژگی‌ها، پارامتر تنظیم و پارامتر کرنل (شکل ۶). طول قسمت اول به تعداد باندهای تصویر و طول دو قسمت آخر به دقت مورد نیاز برای پارامترها بستگی دارد.



شکل ۶- نمایش کروموزم به منظور حل همزمان انتخاب ویژگی و پارامترهای کرنل

در این مرحله بعد از ساخت جمعیت اولیه به صورت تصادفی، به منظور ارزیابی اعضای جمعیت، با توجه به قسمت اول کروموزم (ویژگی‌های انتخاب شده) در داده‌های آموزشی با باندهای انتخاب شده استخراج می‌شود. سپس دو قسمت پارامترها با استفاده از رابطه (۶) به مقدار حقیقی تبدیل می‌شود و ماشین‌های بردار پشتیبان با استفاده از داده‌های آموزشی با باندهای منتخب و پارامترهای بدست آمده، آموزش داده می‌شود. در مرحله بعد داده‌های تست توسط ماشین‌های بردار پشتیبان آموزش دیده طبقه‌بندی می‌شوند و ماتریس

باشند. همچنین پارامتر دمای اولیه و نرخ کاهش دما، به ترتیب ۱۰۰۰ و ۰/۰۱ در نظر گرفته شد.

۵.۳. ارزیابی نتایج

یکی از پارامترهای تأثیرگذار در طبقه بندی، تعیین معیاری برای ارزیابی نتایج می‌باشد. برای این منظور ابتدا ۵۰٪ داده‌ها به صورت تصادفی به داده آموزشی و ۵۰٪ باقیمانده به داده تست تخصیص داده شد. ماشین‌های بردار پشتیبان بر روی داده‌های آموزشی، آموزش دیده و ماتریس خطا با استفاده از داده تست تشکیل داده شد. دو معیار کلی ارزیابی دقت که از ماتریس خطا محاسبه می‌شوند، عبارتند از: ضریب کاپا که با رابطه (۷) محاسبه و دقت کلی که به رابطه (۹) محاسبه می‌گردد.

$$Overall Accuracy = \frac{\sum_{i=1}^r x_{ii}}{N} \times 100 \quad (9)$$

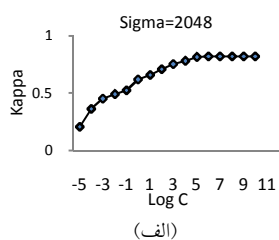
در این رابطه، N تعداد کل نمونه‌ها، r تعداد کلاس‌ها و x_{ii} عناصر روی قطر اصلی ماتریس خطا می‌باشد معیار دیگری که برای بررسی نتایج استفاده گردید، دقت برای هر کلاس می‌باشد که با رابطه (۱۰) محاسبه می‌شود.

$$k_i = \frac{N \times x_{ii} - x_{+i} \times x_{i+}}{N \times x_{i+} - x_{+i} \times x_{i+}} \quad (10)$$

از آنجایی که ضریب کاپا و دقت کلی معیارهایی کلی برای طبقه بندی می‌باشد، استفاده از دقت هر کلاس می‌تواند به صورت مناسب نمایانگر نحوه پخش خطا در بین کلاس‌ها باشد. از این رو در این تحقیق در کنار معیارهای کلی، از دقت هر کلاس نیز استفاده گردید. در ادامه نتایج بدست آمده در ۳ رویکرد بیان شده، ارائه گردیده است.

۵.۳.۱. نتایج انتخاب پارامترهای بهینه مدل

به منظور ارزیابی اثر هر یک از پارامترهای ماشین‌های بردار پشتیبان (C, d) بر روی دقت طبقه‌بندی تصاویر فرا طیفی، روند ذیل انجام پذیرفت. ابتدا در کرنل گوسین، پارامتر کرنل ثابت در نظر گرفته شد و دقت طبقه‌بندی بر اساس تغییرات پارامتر C محاسبه شد (شکل ۸-الف). سپس با تغییر پارامتر σ و ثابت نگه داشتن پارامتر C ، اثر پارامتر کرنل گوسین اندازه‌گیری شد (شکل ۸-ب). در نهایت در کرنل پلی‌نومیال، اثر درجه آن بر دقت طبقه‌بندی بدست آمد (شکل ۸-ج).



نویزی و ۲۰ باند جذبی آب حذف گردید و در نهایت از ۱۸۵ باند باقیمانده استفاده شد. تصویر دارای 145×145 پیکسل، دقت رادیومتریکی ۸ بیت و دارای ۱۶ کلاس می‌باشد که برخی از کلاس‌ها به دلیل کوچک بودن تعداد نمونه‌های آموزشی برای ارزیابی مناسب نمی‌باشد. به همین علت از ۹ کلاس که بیشترین داده را در مقایسه با دیگر کلاس‌ها داشتند، در این تحقیق استفاده شده است.

۵.۲. تنظیم پارامترها

به منظور تعیین پارامترهای ماشین‌های بردار پشتیبان و فضای ویژگی، پارامترهای الگوریتم ژنتیک می‌بایست تنظیم شوند. پارامترهای مورد استفاده در الگوریتم ژنتیک به صورت تجربی و با آزمون و خطا طبق جدول ۱ بدست آمدند.

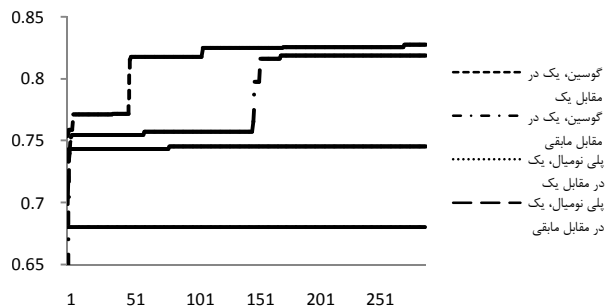
تعداد بیت‌های مورد نیاز برای نمایش پارامتر تنظیم، پارامتر کرنل گوسین و درجه پلی‌نومیال به ترتیب ۲۰، ۲۲ و ۴ تعیین شدند. در پروسه انتخاب ویژگی و حل همزمان تعیین پارامترها و انتخاب ویژگی با در نظر گرفتن ابعاد بزرگ فضای جستجو و به منظور تعیین جمعیت اولیه مناسب ابتدا یک جمعیت با ابعاد بزرگ به صورت اولیه ساخته می‌شود و سپس اعضای برتر به عنوان جمعیت اولیه انتخاب می‌شود. همچنین در تابع هدف، مطابق معادله (۸) با توجه به اهمیت بیشتر دقت نسبت به تعداد ویژگی، $w=0.8$ در نظر گرفته شد.

جدول ۱- پارامترهای الگوریتم ژنتیک

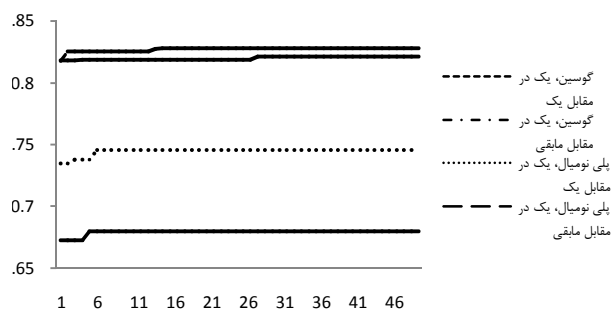
| مقدار | پارامتر | روش |
|-------|--------------------------|------------------|
| ۵ | سایز جمعیت | انتخاب مدل |
| ۴۲ | طول کروموزم (گوسین) | انتخاب ویژگی |
| ۲۴ | طول کروموزم (پلی‌نومیال) | انتخاب ویژگی |
| ۳۰۰ | سایز جمعیت اولیه | انتخاب ویژگی |
| ۳۰ | سایز جمعیت | انتخاب ویژگی |
| ۱۸۵ | طول کروموزم | انتخاب ویژگی |
| ۵۰۰ | سایز جمعیت اولیه | انتخاب ویژگی |
| ۵۰ | سایز جمعیت | انتخاب ویژگی |
| ۲۲۷ | طول کروموزم (گوسین) | انتخاب ویژگی |
| ۲۰۹ | طول کروموزم (پلی‌نومیال) | انتخاب ویژگی |
| ۰/۵ | نرخ تلفیق | پارامترهای عمومی |
| ۰/۰۵ | نرخ جهش | پارامترهای عمومی |
| ۳۰۰ | حداکثر تکرار | پارامترهای عمومی |

پارامترهای الگوریتم شبیه‌سازی تبرید تدریجی نیز به گونه‌ای تنظیم شده‌اند که دارای حجم محاسباتی یکسان با الگوریتم ژنتیک

تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک

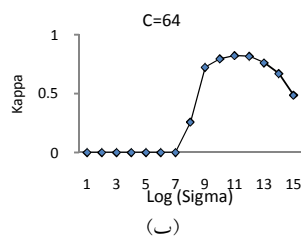


شکل ۹- نمودار همگرایی تعیین پارامترهای ماشین‌های بردار پشتیبان بر مبنای الگوریتم شبیه‌سازی تبرید تدریجی برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی

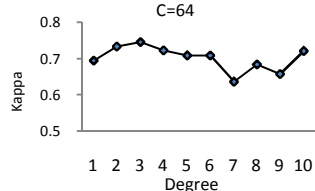


شکل ۱۰- نمودار همگرایی تعیین پارامترهای ماشین‌های بردار پشتیبان بر مبنای الگوریتم ژنتیک برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی

مقادیر بهینه پارامترها و دقت متناظر با آن، برای دو کرنل گوسین و پلی‌نومیال در دو حالت چند کلاس، یک در مقابل یک و یک در مقابل مابقی در جدول ۲ ارائه شده است. همچنین دقت هر کلاس با رابطه (۱۰) محاسبه و در جدول ۲ ارائه شده است. همان طور که در جدول ۲ نشان داده شده است، دقت کلی طبقه‌بندی در روش‌های الگوریتم ژنتیک، شبیه‌سازی تبرید تدریجی و جستجوی شبکه‌ای نزدیک به هم می‌باشد. با توجه به حجم محاسباتی کمتر الگوریتم ژنتیک، سرعت همگرایی بالا و عدم جستجوی جامع، می‌تواند در زمان محاسباتی کمتر، به دقتی نزدیک و یا بالاتر از الگوریتم جستجوی شبکه‌ای و شبیه‌سازی تبرید تدریجی برسد. ولی نحوه پخش خطا در سه روش متفاوت است. از این رو در برخی از کلاس‌ها، دقت تعیین پارامتر بر مبنای الگوریتم ژنتیک، در برخی الگوریتم شبیه‌سازی تبرید تدریجی و در برخی کلاس‌ها الگوریتم جستجوی شبکه‌ای به دقت بالاتری رسیده است.



(ب)



(ج)

شکل ۸- تأثیر پارامترهای ماشین‌های بردار بر دقت طبقه‌بندی، الف) تأثیر پارامتر تنظیم (ب) تأثیر پارامتر کرنل گوسین (ج) تأثیر درجه پلی‌نومیال

همان طور که در شکل ۸ دیده می‌شود، هر سه پارامتر تأثیر بسزایی بر عملکرد طبقه‌بندی کننده دارند و در نتیجه تعیین مقدار بهینه این پارامترها از تأثیر بسزایی در دقت طبقه‌بندی برخوردار می‌باشد. از این رو در این مرحله از جستجوی شبکه‌ای، الگوریتم شبیه‌سازی تبرید تدریجی و الگوریتم ژنتیک برای تعیین مقادیر بهینه پارامترهای دو کرنل گوسین و پلی‌نومیال استفاده گردید. به منظور تعیین پارامترهای ماشین‌های بردار پشتیبان مبتنی بر این سه روش، پارامترهای C, σ و d به ترتیب در بازه‌های $[2^{-5}, \dots, 2^{12}]$ ، $[2^{-5}, \dots, 2^{10}]$ و $[1, \dots, 10]$ در نظر گرفته شدند. نمودار همگرایی انتخاب مدل بر مبنای الگوریتم شبیه‌سازی تبرید تدریجی و الگوریتم ژنتیک به ترتیب در شکل‌های ۹ و ۱۰ نمایش داده شده است. نمودار همگرایی الگوریتم ژنتیک حاکی از همگرایی سریع آن (حداکثر ۳۰ تکرار) به مقدار بهینه می‌باشد؛ در حالی که الگوریتم شبیه‌سازی تبرید تدریجی نیاز به تکرارهای بیشتری برای همگرایی دارد.

به منظور مقایسه نتایج بدست آمده، روش کلاسیک جستجوی شبکه‌ای پیاده‌سازی گردید. تغییر نمایی پارامترهای تنظیم و کرنل گوسین در این روش به صورت تجربی بدست آمده است. ابتدا دقت طبقه‌بندی با تمام زوج پارامترها در هر کرنل محاسبه و در نهایت زوج پارامتر با بیشترین دقت طبقه‌بندی به عنوان مقادیر نهایی ماشین‌های بردار پشتیبان انتخاب شدند. از آنجایی که الگوریتم جستجوی شبکه‌ای به منظور تعیین مقادیر بهینه پارامترها جستجوی جامع انجام می‌دهد، می‌توان انتظار داشت که دقت بدست آمده از این روش، دقتی نزدیک به بهینه باشد و در نتیجه معیار مناسبی به منظور مقایسه با الگوریتم ژنتیک و شبیه‌سازی تبرید تدریجی باشد.

جدول ۲- نتایج تعیین پارامترهای بهینه بر مبنای الگوریتم ژنتیک، شبیه‌سازی تبرید تدریجی و جستجوی شبکه‌ای

| تکنیک | پارامتر تنظیم | پارامتر کرنل | کلاس | | | | | | | | | | | |
|---------|---------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|----------|-------|
| | | | کلاس ۱ | کلاس ۲ | کلاس ۳ | کلاس ۴ | کلاس ۵ | کلاس ۶ | کلاس ۷ | کلاس ۸ | کلاس ۹ | کاپا | ضریب کلی | |
| شبکه‌ای | SA | ۲۰۴۸ | ۱ | ۱ | ۱ | ۰/۶۵۳ | ۰/۹۸۲ | ۰/۵۶۹ | ۰/۸۶۸ | ۰/۸۷۲ | ۰/۵۹۴ | ۰/۸۲ | ۸۴/۸۳ | |
| | | | ۱ | ۱ | ۱ | ۰/۷۰۲ | ۰/۹۸۲ | ۰/۶۴۲ | ۰/۸۶۹ | ۰/۸۸۲ | ۰/۵۶۲ | ۰/۸۲۸ | ۸۵/۴۷ | |
| | | | ۱ | ۱ | ۱ | ۰/۷ | ۰/۹۸۲ | ۰/۶۴۲ | ۰/۸۶۸ | ۰/۸۷۲ | ۰/۵۹۵ | ۰/۸۲۸ | ۸۵/۴۷ | |
| شبکه‌ای | SA | ۴۰۹۶ | ۱ | ۱ | ۱ | ۰/۷۹ | ۱ | ۱ | ۰/۷۹ | ۰/۹۱۲ | ۰/۷۲۵ | ۰/۸۱۹ | ۸۴/۶۲ | |
| | | | ۱ | ۱ | ۱ | ۰/۷۶۷ | ۱ | ۱ | ۰/۵۴ | ۰/۹۱۲ | ۰/۸۰۸ | ۰/۷۲۵ | ۰/۸۱۹ | ۸۴/۶۲ |
| | | | ۱ | ۱ | ۱ | ۰/۷۶۷ | ۱ | ۱ | ۰/۵۴۱ | ۰/۹۱۲ | ۰/۸۱۹ | ۰/۷۲۵ | ۰/۸۲۱ | ۸۴/۳۳ |
| شبکه‌ای | SA | ۳ | ۱ | ۱ | ۱ | ۰/۹۷۱ | ۰/۹۶۴ | ۰/۴۵۳ | ۰/۹۱۲ | ۰/۶۷۲ | ۰/۶۵۸ | ۰/۷۴۶ | ۷۸/۴۲ | |
| | | | ۱ | ۱ | ۱ | ۰/۹۷۱ | ۰/۹۶۴ | ۰/۴۵۳ | ۰/۹۱۲ | ۰/۶۷۲ | ۰/۶۵۸ | ۰/۷۴۶ | ۷۸/۴۲ | |
| | | | ۱ | ۱ | ۱ | ۰/۹۷۱ | ۰/۹۶۴ | ۰/۴۵۳ | ۰/۹۱۲ | ۰/۶۷۲ | ۰/۶۵۸ | ۰/۷۴۶ | ۷۸/۴۲ | |
| شبکه‌ای | SA | ۸ | ۱ | ۱ | ۱ | ۰/۹۴۱ | ۱ | ۱ | ۰/۵۱۷ | ۰/۷۳۶ | ۰/۴۸۹ | ۰/۶۷۷ | ۷۲/۴۴ | |
| | | | ۱ | ۱ | ۱ | ۰/۹۴۱ | ۱ | ۱ | ۰/۵۱۷ | ۰/۷۳۶ | ۰/۴۸۹ | ۰/۶۵۶ | ۰/۶۸ | ۷۲/۶۵ |
| | | | ۱ | ۱ | ۱ | ۰/۹۴۱ | ۱ | ۱ | ۰/۵۱۷ | ۰/۷۳۶ | ۰/۴۸۹ | ۰/۶۵۶ | ۰/۶۸ | ۷۲/۶۵ |

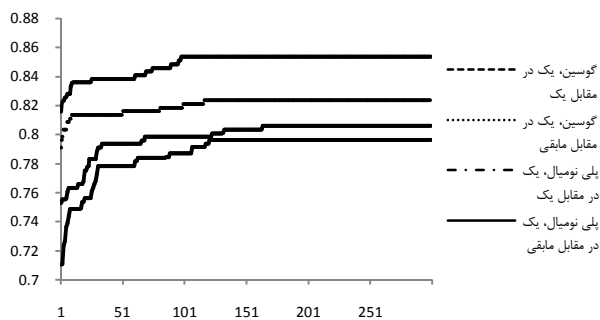
۵.۳.۲. نتایج انتخاب ویژگی

الگوریتم شبیه‌سازی تبرید تدریجی و ژنتیک، معیارهای ضریب کاپا، دقت کلی و دقت هر کلاس، به کار گرفته شد (جدول ۳). نتایج نشان می‌دهند با حذف ویژگی‌های اضافی نه تنها سرعت طبقه‌بندی بالا می‌رود بلکه دقت آن نیز افزایش می‌یابد.

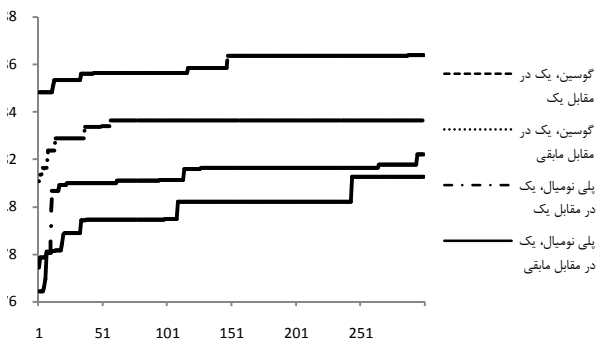
در این قسمت هدف یافتن زیرمجموعه‌ای از ویژگی‌های بهینه تصویر فرا طیفی به منظور طبقه‌بندی با بیشترین کارایی می‌باشد. در این قسمت پارامترهای ماشین‌های بردار پشتیبان، در حضور تمام باندها و با استفاده از الگوریتم جستجوی شبکه‌ای بدست آمده و مقدار آن‌ها در طول پروسه انتخاب ویژگی ثابت در نظر گرفته می‌شود.

شکل ۱۱ و ۱۲ به ترتیب نمودار همگرایی الگوریتم شبیه‌سازی تبرید تدریجی و الگوریتم ژنتیک برای ۴ مورد در نظر گرفته شده را نشان می‌دهد. همان طور که در این دو شکل دیده می‌شود، انتخاب ویژگی موجب بهبود قابل توجهی بر دقت کرنل‌های پلی‌نومیال و گوسین می‌شود. با توجه به شکل‌های ۹ و ۱۰، کرنل گوسین دقت بسیار بالاتری نسبت به کرنل پلی‌نومیال دارد ولی پروسه انتخاب ویژگی با تأثیر بالایی که بر کرنل پلی‌نومیال داشته، باعث نزدیکی دو کرنل در دو حالت یک در مقابل یک و یک در مقابل مابقی گردیده است که حاکی از تأثیر بیشتر انتخاب ویژگی بر کرنل پلی‌نومیال نسبت به کرنل گوسین می‌باشد. نکته دیگر تفاوت در سرعت همگرایی دو کرنل می‌باشد. با توجه به شکل ۱۲، کرنل گوسین در تکرارهای کمتری به دقت بهینه همگرا می‌شوند، از طرف مقابل دقت کرنل پلی‌نومیال تا تکرارهای پایانی افزایش پیدا می‌کند و به تکرارهای بیشتری برای رسیدن به ثبات احتیاج دارد.

به منظور مقایسه دقیق‌تر نتایج در سه حالتی که از همه باندها استفاده شده باشد و نتایج بعد از مرحله انتخاب ویژگی توسط



شکل ۱۱-نمودار همگرایی انتخاب ویژگی بر مبنای الگوریتم شبیه‌سازی تبرید تدریجی برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی



شکل ۱۲-نمودار همگرایی انتخاب ویژگی بر مبنای الگوریتم ژنتیک برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی

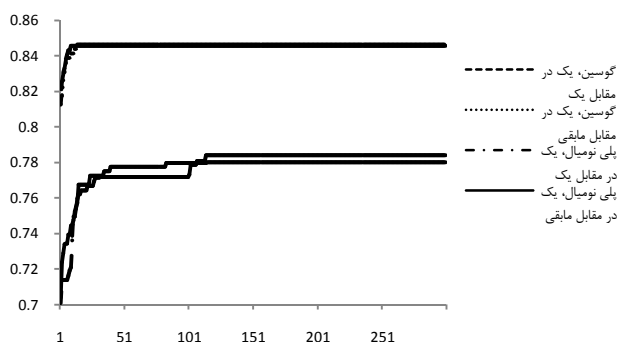
تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک

جدول ۳- نتایج حاصل از انتخاب ویژگی بر مبنای الگوریتم ژنتیک و شبیه‌سازی تبرد تدریجی در مقایسه با حضور همه باندها

| تکنیک | تعداد باندها | کلاس ۱ | کلاس ۲ | کلاس ۳ | کلاس ۴ | کلاس ۵ | کلاس ۶ | کلاس ۷ | کلاس ۸ | کلاس ۹ | ضریب کاپا | دقت کلی |
|------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-----------|---------|
| همه باندها | ۱۸۵ | ۰/۸۱۶ | ۱ | ۱ | ۰/۶۵۳ | ۰/۹۸۲ | ۰/۵۶۹ | ۰/۸۶۸ | ۰/۸۷۲ | ۰/۵۹۴ | ۰/۸۲ | ۸۴/۸۳ |
| | SA | ۰/۸۲ | ۱ | ۱ | ۰/۷۴۵ | ۱ | ۰/۶۴۱ | ۰/۸۶۸ | ۰/۹۰۸ | ۰/۶۲۸ | ۰/۸۵۳ | ۸۷/۶۱ |
| | GA | ۰/۸۵۱ | ۱ | ۱ | ۰/۷۲۳ | ۱ | ۰/۷۱۷ | ۰/۸۶۸ | ۰/۸۸۷ | ۰/۶۹۴ | ۰/۸۶۴ | ۸۷/۴۶ |
| همه باندها | ۱۸۵ | ۰/۷۲۲ | ۱ | ۱ | ۰/۷۸۹ | ۱ | ۰/۴۹۲ | ۰/۹۱۲ | ۰/۸۱۹ | ۰/۷۲۵ | ۰/۸۱۹ | ۸۴/۶۲ |
| | SA | ۰/۷۸۳ | ۱ | ۱ | ۰/۷۴۷ | ۱ | ۰/۴۶۹ | ۰/۹۱۲ | ۰/۸۳ | ۰/۷۲۵ | ۰/۸۲۳ | ۸۵/۰۴ |
| | GA | ۰/۸۱۶ | ۱ | ۱ | ۰/۷۲۵ | ۱ | ۰/۵۱۴ | ۰/۹۱۲ | ۰/۸۴۴ | ۰/۷۵۹ | ۰/۸۳۶ | ۸۶/۱۱ |
| همه باندها | ۱۸۵ | ۰/۷۳۴ | ۰/۹۷۱ | ۱ | ۰/۵۴۶ | ۰/۹۶۴ | ۰/۴۵۳ | ۰/۹۱۲ | ۰/۶۷۲ | ۰/۶۵۸ | ۰/۷۴۶ | ۷۸/۴۲ |
| | SA | ۰/۷۴۹ | ۰/۹۷۱ | ۱ | ۰/۶۵۹ | ۰/۹۸۲ | ۰/۴۸۵ | ۰/۹۱۲ | ۰/۷۸۵ | ۰/۷۹۵ | ۰/۸۰۵ | ۸۳/۱۲ |
| | GA | ۰/۷۹۷ | ۱ | ۱ | ۰/۶۳۳ | ۱ | ۰/۶۳۶ | ۰/۹۱۲ | ۰/۷۷۳ | ۰/۸۲۷ | ۰/۸۲۵ | ۸۵/۰۴ |
| همه باندها | ۱۸۵ | ۰/۵۹۷ | ۰/۹۴۱ | ۱ | ۰/۵۱۷ | ۱ | ۰/۵۱۸ | ۰/۷۳۶ | ۰/۴۸۹ | ۰/۶۵۷ | ۰/۶۷۷ | ۷۲/۴۴ |
| | SA | ۰/۷۸۳ | ۰/۹۴۱ | ۱ | ۰/۶۱۴ | ۰/۹۸۲ | ۰/۵۵۹ | ۰/۸۲۳ | ۰/۷۴۴ | ۰/۶۹۱ | ۰/۷۹۶ | ۸۱/۶۲ |
| | GA | ۰/۸۰۱ | ۱ | ۱ | ۰/۶۷۹ | ۱ | ۰/۶۶۱ | ۰/۸۲۵ | ۰/۷۳۱ | ۰/۷۹ | ۰/۸۱۲ | ۸۳/۹۷ |

ژنتیک را برای ۴ وضعیت در نظر گرفته شده نشان می‌دهند. با توجه به این دو شکل با در نظر گرفتن همزمان پارامترها و ویژگی‌ها، دقت طبقه‌بندی بدست آمده بیشتر تحت تأثیر نوع کرنل می‌باشد و تأثیر نوع روش چند کلاسه کاهش یافته است. از طرف دیگر با توجه به بزرگ بودن ابعاد فضای جستجو در این بخش، زمان یافتن جواب بهینه نیز در هر ۴ وضعیت افزایش یافت.

به منظور ارزیابی دقیق‌تر نتایج، مقادیر پارامترهای ماشین‌های بردار پشتیبان، تعداد باندهای انتخابی، ضریب کاپا و دقت کلی به عنوان معیارهای دقت کلی و دقت کلاس‌ها به صورت مجزا در جدول ۴ ارائه شده است.



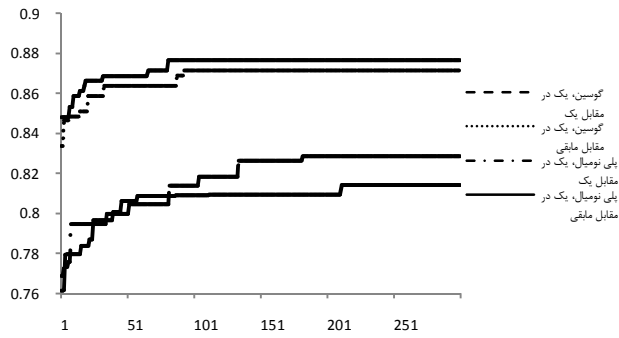
شکل ۱۳- نمودار همگرایی تعیین همزمان پارامترهای ماشین‌های بردار پشتیبان و زیرمجموعه بهینه ویژگی بر مبنای الگوریتم شبیه‌سازی تبرد تدریجی برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی

در تمام موارد الگوریتم ژنتیک به دقتی بالاتر از الگوریتم شبیه‌سازی تبرد تدریجی دست یافت. همچنین در نتایج بدست آمده از الگوریتم ژنتیک با کرنل گوسین در حالت یک در مقابل یک، مرحله انتخاب ویژگی با حذف ۸۳ باند اضافی، توانست دقت طبقه‌بندی را نسبت به حالتی که تمام باندها حضور دارند، ۴٪ بهبود ببخشد. در مورد کرنل گوسین و در حالت یک در مقابل همه، ۸۶ باند اضافی حذف و دقتی ۲٪ بالاتر از زمان حضور همه باندها حاصل گردید.

کرنل پلی‌نومیال با حذف باندهای بیشتر، افزایش دقت بیشتری نسبت به کرنل گوسین حاصل نمود. در حالت چند کلاسه‌ی یک در مقابل یک، دقت طبقه‌بندی با حذف ۹۱ باند اضافی، ۷٪ افزایش یافت. در بین کرنل‌های در نظر گرفته شده، کرنل پلی‌نومیال در حالت یک در مقابل مابقی، با حذف بیشترین باندهای اضافی، بالاترین افزایش دقت را داشت. در این مورد، بیش از نیمی از باندها (۱۰۴ باند) حذف گردید و دقت ۱۱٪ افزایش یافت. همچنین دقت هر کلاس هم پس از انتخاب ویژگی در اکثر موارد افزایش بهبود قابل توجهی نسبت به حضور همه باندها داشته است.

۵.۳.۳. نتایج حل همزمان انتخاب مدل و ویژگی

با توجه به تأثیر مقدار پارامترها در نحوه انتخاب زیرمجموعه بهینه ویژگی‌ها و بالعکس، در این مرحله نسبت به تعیین همزمان مقادیر پارامترها و ویژگی‌ها اقدام گردید. شکل‌های ۱۳ و ۱۴ به ترتیب نمودار همگرایی الگوریتم شبیه‌سازی تبرد تدریجی و الگوریتم



شکل ۱۴- نمودار همگرایی تعیین همزمان پارامترهای ماشین‌های بردار پشتیبان و زیر مجموعه بهینه ویژگی بر مبنای الگوریتم ژنتیک برای کرنل گوسین و پلی‌نومیال در دو حالت یک در مقابل یک و یک در مقابل مابقی

جدول ۴- نتایج حاصل از حل همزمان تعیین پارامترها و انتخاب ویژگی بر مبنای الگوریتم ژنتیک و شبیه‌سازی تدریجی

| وضعیت | روش | تعداد باندها | پارامتر تنظیم | پارامتر کرنل | کلاس ۱ | کلاس ۲ | کلاس ۳ | کلاس ۴ | کلاس ۵ | کلاس ۶ | کلاس ۷ | کلاس ۸ | کلاس ۹ | ضریب کاپا | دقت کلی |
|-------------------------------|-----|--------------|---------------|--------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-----------|---------|
| گوسین- یک در مقابل یک | SA | ۸۱ | ۶۲۷/۳۳۹ | ۱۵۵۰/۵۱ | ۰/۷۸۸ | ۱ | ۱ | ۰/۷۴۷ | ۰/۹۸۲ | ۰/۶۹۱ | ۰/۸۶۸ | ۰/۸۸۵ | ۰/۶۲۶ | ۰/۸۴۶ | ۸۶/۹۶ |
| گوسین- یک در مقابل مابقی | GA | ۹۰ | ۱۸۱/۱۶ | ۲۰۵۷/۵ | ۰/۸۵۱ | ۱ | ۱ | ۰/۷۶۹ | ۱ | ۰/۶۴۴ | ۰/۹۱۲ | ۰/۹۱ | ۰/۷۹۴ | ۰/۸۷۶ | ۸۹/۵۳ |
| گوسین- یک در مقابل مابقی | SA | ۷۸ | ۲۸۹/۷۵۳ | ۲۸۲۹/۲۶ | ۰/۸۵ | ۱ | ۱ | ۰/۷۶۹ | ۱ | ۰/۵۶۸ | ۰/۹۱۲ | ۰/۸۰۹ | ۰/۷۹۳ | ۰/۸۴۶ | ۸۶/۹۶ |
| پلی‌نومیال- یک در مقابل یک | GA | ۹۸ | ۵۶۵/۶۷ | ۲۰۹۳/۲۳ | ۰/۸۵۲ | ۱ | ۱ | ۰/۷۶۸ | ۱ | ۰/۶۹ | ۰/۹۱۳ | ۰/۸۸۷ | ۰/۷۲۵ | ۰/۸۷۱ | ۸۹/۱ |
| پلی‌نومیال- یک در مقابل یک | SA | ۹۳ | ۶۲۷/۳۳۹ | ۵/۱۹۶ | ۰/۷۹۸ | ۰/۹۷ | ۱ | ۰/۵۲۲ | ۰/۹۶۴ | ۰/۶۵۷ | ۰/۹۱۲ | ۰/۷۰۵ | ۰/۷۲۷ | ۰/۷۸۴ | ۸۱/۶۲ |
| پلی‌نومیال- یک در مقابل مابقی | GA | ۸۹ | ۴۰۴/۱ | ۴/۶ | ۰/۷۳۷ | ۰/۹۷۱ | ۱ | ۰/۶۱۱ | ۱ | ۰/۶۸۹ | ۰/۹۱۲ | ۰/۸۳۱ | ۰/۸۲۸ | ۰/۸۲۹ | ۸۵/۴۷ |
| پلی‌نومیال- یک در مقابل مابقی | SA | ۸۸ | ۹۵۵/۱۸۴ | ۴/۹۶۸ | ۰/۷۶۵ | ۰/۹۴۱ | ۱ | ۰/۵۴۸ | ۱ | ۰/۶۳۴ | ۰/۷۸ | ۰/۶۸۹ | ۰/۸۶۹ | ۰/۷۸ | ۸۱/۲ |
| پلی‌نومیال- یک در مقابل مابقی | GA | ۸۹ | ۷۳۸/۱۵ | ۱۰ | ۰/۸۱۸ | ۰/۹۴ | ۱ | ۰/۵۹۶ | ۱ | ۰/۷۱۲ | ۰/۸۲۴ | ۰/۷۶۹ | ۰/۷۹۲ | ۰/۸۱۴ | ۸۴/۱۹ |

انتخاب ویژگی بالاترین دقت را در ۴ وضعیت دارد. همچنین می‌توان از شکل ۱۲ نتیجه گرفت با توجه به افزایش دقت بیشتر در حل همزمان برای کرنل گوسین، تأثیر پارامترها بر فضای ورودی در این کرنل بیشتر از کرنل پلی‌نومیال می‌باشد.

۶. نتیجه گیری و پیشنهادات

در این تحقیق سه رویکرد در ایجاد یک طبقه‌بندی کننده بهینه مبتنی بر ماشین‌های بردار پشتیبان بر مبنای الگوریتم ژنتیک ارائه گردید. نتایج حاصل در مقایسه با الگوریتم شبیه‌سازی تدریجی، بیانگر برتری الگوریتم ژنتیک به خصوص با افزایش ابعاد فضای جستجو می‌باشد. همچنین نتایج بهینه با به‌کارگیری الگوریتم ژنتیک در حل همزمان تعیین پارامتر و انتخاب ویژگی حاصل می‌شود که در آن با انتخاب باندهای کمتر، دقت بالاتر حاصل شد. از این رو با استفاده از طبقه‌بندی کننده‌ی کارا و قدرتمند ماشین‌های بردار پشتیبان در کنار الگوریتم بهینه سازی و فرا ابتکاری ژنتیک، می‌تواند یک سیستم طبقه‌بندی ترکیبی بهینه برای تصاویر فرا طیفی طراحی کرد.

مقایسه نتایج این بخش با مراحل قبل، حاکی از رسیدن به دقت-های بالاتری در هر چهار وضعیت می‌باشد که نشان از بهینه بودن نتایج در حل همزمان می‌باشد. همچنین مقایسه نتایج حاصل از دو روش حاکی از برتری بیشتر الگوریتم شبیه‌سازی تدریجی نسبت به الگوریتم ژنتیک می‌باشد که بیانگر توانایی الگوریتم ژنتیک در فضای جستجوی پیچیده می‌باشد.

۵.۳.۴. مقایسه نتایج

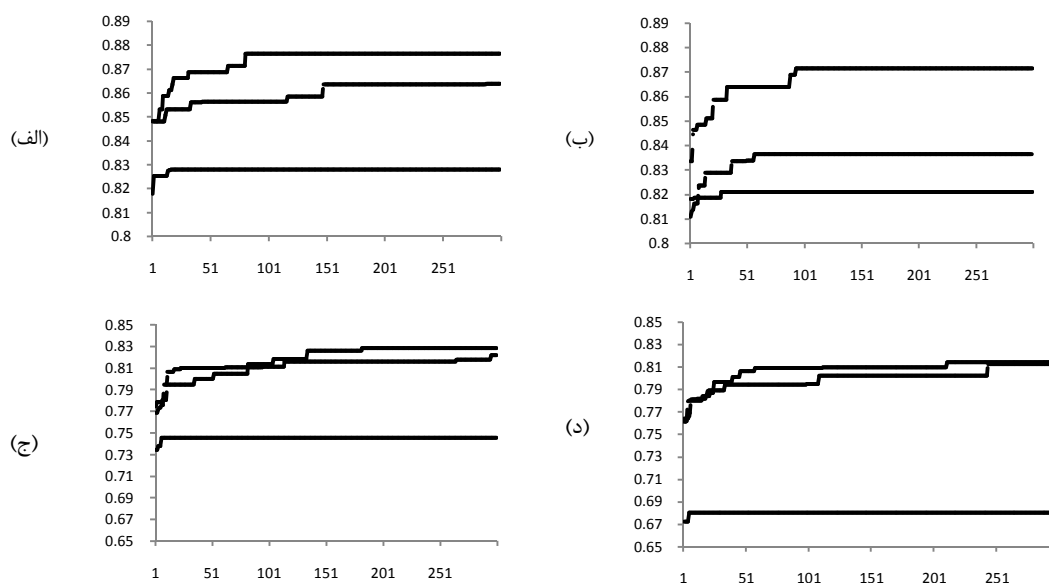
به منظور مقایسه بهتر بین نتایج بدست آمده، منحنی همگرایی الگوریتم ژنتیک در سه رویکرد ارائه شده (تعیین پارامترها، انتخاب ویژگی و حل همزمان هر دو) را بر روی ۴ وضعیت در نظر گرفته شده، در شکل ۱۵ نمایش داده شد. همان طور که در این شکل دیده می‌شود، تعیین پارامترهای بهینه ماشین‌های بردار پشتیبان در تکرارهای اولیه به همگرایی رسیده و نسبت به دو رویکرد دیگر بهبود کمتری در دقت ایجاد کرده است. در حالی که، انتخاب ویژگی با توجه به استفاده از پارامترهای بدست آمده از جستجوی شبکه‌ای دقت را نسبت به تعیین پارامترها افزایش داده است. در نهایت حل همزمان پارامترها و

تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک

پشتیبان در تصاویر فرا طیفی استفاده گردد. همچنین بهینه‌سازی اتوماتیک پارامترهای الگوریتم ژنتیک که در این نوشته با آزمون و خطا بدست آمد، از موضوعات دیگر پیش رو می‌باشد.

پیشنهاد می‌شود در تحقیقات آینده با توجه به بزرگ بودن فضای جستجوی مسئله ارائه شده از الگوریتم‌های فرا ابتکاری دیگر از قبیل روش‌های مبنی بر خرد جمعی به منظور بهینه‌سازی ماشین‌های بردار

حل همزمان تعیین پارامترها و انتخاب ویژگی - - - انتخاب ویژگی - - - - - تعیین پارامترها - - -



شکل ۱۵- مقایسه نتایج حاصل از الگوریتم ژنتیک در سه رویکرد ارائه شده بر روی الف) گوسین، یک در مقابل یک ب) گوسین، یک در مقابل مابقی ج) پلی‌نومیل، یک در مقابل یک د) پلی‌نومیل، یک در مقابل مابقی

مراجع

[1]. C. Chang, *Hyperspectral data exploitation: theory and applications*: Wiley-Blackwell, 2007.

[2]. G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, pp. 55-63, 2002.

[3]. G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 1351-1362, 2005.

[4]. P. Du, K. Tan, W. Zhang, and Z. Yan, "ANN Classification of OMIS Hyperspectral Remotely Sensed Imagery: Experiments and Analysis," *Congress on Image and Signal Processing*, pp. 692-696, 2008.

[5]. T. Waheed, R. Bonnell, S. Prasher, and E. Paulet, "Measuring performance in precision agriculture: CART--A decision tree approach," *Agricultural water management*, vol. 84, pp. 173-185, 2006.

[6]. J. Ham, Y. Chen, M. Crawford, and J. Ghosh, "Investigation of the random forest framework for classification of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 492-501, 2005.

[7]. F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 1778-1790, 2004.

[8]. C. Dai, X. Huang, and G. Dong, "Support Vector Machine for Classification of Hyperspectral Remote Sensing Imagery," *Fourth International Conference on Fuzzy System and Knowledge Discovery*, pp. 77-80, 2007.

[9]. B. Guo, S. Gunn, R. Damper, and J. Nelson, "Customizing kernel functions for SVM-based hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 17, pp. 622-629, 2008.

[10]. P. Watanachaturaporn, M. Arora, and P. Varshney, "Hyperspectral image classification using support vector machines: A comparison with decision tree and neural network classifiers," 2006.

[11]. S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Surveys*, vol. 4, pp. 40-79, 2010.

[12]. X. Zhang, X. Chen, and Z. He, "An ACO-based algorithm for parameter optimization of support vector machines," *Expert systems with applications*, 2010.

[13]. C. Wu, G. Tzeng, Y. Goo, and W. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," *Expert systems with applications*, vol. 32, pp. 397-408, 2007.

[14]. E. Huerta, B. Duval, and J. Hao, "A hybrid GA/SVM approach for gene selection and classification of microarray data," *Applications of Evolutionary Computing*, pp. 34-44, 2006.

[15]. L. Zhuo, J. Zheng, F. Wang, X. Li, B. Ai, and J. Qian, "A genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector

- machine," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 397-402, 2008.
- [16]. T. Zhang, X. Fu, R. Goh, C. Kwok, and G. Lee, "A GA-SVM feature selection model based on high performance computing techniques," *IEEE International Conference on System, Man and Cybernetics*, pp. 2653-2658, 2009.
- [17]. C. Huang, "ACO-based hybrid classification system with feature subset selection and model parameters optimization," *Neurocomputing*, vol. 73, pp. 438-448, 2009.
- [18]. S. Lin, Z. Lee, S. Chen, and T. Tseng, "Parameter determination of support vector machine and feature selection using simulated annealing approach," *Applied soft computing*, vol. 8, pp. 1505-1512, 2008.
- [19]. S. Lin, K. Ying, S. Chen, and Z. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines," *Expert systems with applications*, vol. 35, pp. 1817-1824, 2008.
- [20]. C. Huang and C. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert systems with applications*, vol. 31, pp. 231-240, 2006.
- [21]. R. Haupt, S. Haupt, and J. Wiley, *Practical genetic algorithms*: Wiley Online Library, 1998.
- [22]. V. Vapnik, *The nature of statistical learning theory*: Springer Verlag, 2000.
- [23]. A. Lorena and A. de Carvalho, "Evolutionary tuning of SVM parameter values in multiclass problems," *Neurocomputing*, vol. 71, pp. 3326-3334, 2008.
- [24]. T. Weise, "Global Optimization Algorithms—Theory and Application," *Abrufdatum*, vol. 1, 2008.
- [25]. U. Maulik, "Medical image segmentation using genetic algorithms," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 166-173, 2009.
- [26]. E. Pourbasheer, S. Riahi, M. Ganjali, and P. Norouzi, "Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity," *European journal of medicinal chemistry*, vol. 44, pp. 5023-5028, 2009.
- [27]. B. de Souza, A. de Carvalho, R. Calvo, and R. Ishii, "Multiclass SVM model selection using particle swarm optimization," 2006, p. 31.
- [28]. C. Hsu, C. Chang, and C. Lin, "A practical guide to support vector classification," Citeseer, 2003.
- [29]. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.
- [30]. M. Tahir, A. Bouridane, F. Kurugollu, and A. Amira, "Feature selection using tabu search for improving the classification rate of prostate needle biopsies," *Pattern Recognition*, vol. 2, pp. 335-33, 2004.
- [31]. C. Tu, L. Chuang, J. Chang, and C. Yang, "Feature selection using PSO-SVM," *IAENG International journal of computer science*, vol. 33, pp. 111-116, 2007.
- [32]. D. Niu, Y. Wang, and D. Wu, "Power load forecasting using support vector machine and ant colony optimization," *Expert systems with applications*, vol. 37, pp. 2531-2539, 2010.
- [33]. H. Frohlich, O. Chapelle, and B. Scholkopf, "Feature selection for support vector machines by means of genetic algorithm," 2003, pp. 142-148.
- [34]. S. Bhatia, P. Prakash, and G. Pillai, "SVM Based Decision Support System for Heart Disease Classification with Integer-Coded Genetic Algorithm to Select Critical Features," 2008.
- [35]. A.B. Santos, C.S.F. de S. Celes, A. de A. Araújo, D. Menotti, "Feature selection for classification of remote sensed hyperspectral images: A filter approach using genetic algorithm and cluster validity," *The 2012 International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICCV'12)*, 2012.
- [36]. Y. Bazi and F. Melgani, "Toward an optimal SVM classification system for hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 3374-3385, 2006.
- [37]. I. Mejía-Guevara and Ákuri-Morales, "Evolutionary feature and parameter selection in support vector regression," *MICAI 2007: Advances in Artificial Intelligence*, pp. 399-408, 2007.
- [38]. E. Avci, "Selecting of the optimal feature subset and kernel parameters in digital modulation classification by using hybrid genetic algorithm-support vector machines: HGASVM," *Expert systems with applications*, vol. 36, pp. 1391-1402, 2009.
- [39]. B. F. de Souza and A. P. d. L. F. de Carvalho, "Gene selection based on multi-class support vector machines and genetic algorithms," *Genetics and Molecular Research*, pp. 599-607, 2005.
- [40]. M. Zhao, C. Fu, L. Ji, K. Tang, and M. Zhou, "Feature selection and parameter determination for support vector machines: A new approach based on genetic algorithm with feature chromosomes," *Expert System with Applications*, Vol. 38, No. 5, pp. 5197-5204, 2011.
- [41]. N. Chen, B. Ribeiro, A. S. Vieira, J. Duarte, J. C. Neves, "A genetic algorithm-based approach to cost-sensitive bankruptcy prediction," *Expert System with Application*, Vol. 38, No. 10, pp. 12939-12945, 2011.

بکارگیری روش ترکیبی SPIHT-DCT با استفاده از مقیاس پذیری مکانی و

مکانی - زمانی برای کد کردن تصاویر ویدئویی HDTV

وحید رضا صیرفیان^{۱*} سیامک طالبی^{**}

* کارشناس ارشد، دانشکده مهندسی برق، دانشگاه آزاد اسلامی واحد میمه، اصفهان

** کارشناس ارشد، دانشکده مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان

تاریخ دریافت: ۱۳۹۱/۰۴/۰۱

تاریخ پذیرش: ۱۳۹۱/۱۰/۱۱

چکیده

در این مقاله، کد کننده ترکیبی با استفاده از دو ویژگی مقیاس پذیری مکانی و مقیاس پذیری مکانی-زمانی برای کد کردن ویدئو با درجه تفکیک بالا ارائه شده است. در روش ترکیبی، قاب های ویدئویی Intra و Inter به دو روش متفاوت کد می شوند. قاب های Intra با استفاده از الگوریتم SPIHT^۱ که مبتنی بر تبدیل موجک است کد می شوند. قاب های Inter به روش معمول استاندارد MPEG-2 و بر اساس تبدیل DCT کد می شوند. با کد کردن ویدئو با درجه تفکیک بالا به دو روش مقیاس پذیری مکانی و مقیاس پذیری مکانی-زمانی، ویدئو از طریق دو یا سه لایه ارسال می شود. داده های ارسالی از لایه ها، ویدئویی با وضوح و کیفیت متفاوت به کاربر عرضه می کنند. به این ترتیب کاربر می تواند بر اساس نیاز خود سرویس مناسب را انتخاب کند.

در مقیاس پذیری مکانی لایه پایه و لایه ارتقا ساختار کدینگ یکسانی دارند. ولی در مقیاس پذیری مکانی-زمانی لایه ارتقا دوم، به دلیل اینکه فقط شامل قاب های Inter است، فقط مبتنی بر روش استاندارد کد می شود.

نتایج شبیه سازی های انجام شده روی ویدئوهای مختلف با درجه تفکیک بالا، بهبود کیفیت تصویر نهایی در روش ترکیبی پیشنهادی با مقیاس پذیری را در لایه های مختلف، نسبت به روش مبتنی بر استاندارد MPEG-2 نشان می دهد.

کلید واژگان: استاندارد MPEG-2، الگوریتم SPIHT، تبدیل

DCT، مقیاس پذیری، ویدئو با درجه تفکیک بالا.

۱. مقدمه

کد کننده های ویدئویی مقیاس پذیر^۲، کاربردهای زیادی در شبکه های مختلف ارتباطی دارند. هر شبکه نیز بر اساس سطوح مختلف سرویسی که به کاربر عرضه می کند ارزش گذاری^۳ و با ارسال ویدئو در کیفیت های مختلف و یا دقت مکانی یا زمانی متفاوت، کاربر می تواند بر اساس نیاز و محدودیت های گیرنده خود، سرویس مورد نظر را انتخاب می کند [۱]. این مورد برای ارسال ویدئو با درجه تفکیک بالا^۴ اهمیت بیشتری دارد.

بررسی کد کردن ویدئو با درجه تفکیک بالا با مقیاس پذیری از این جنبه دارای اهمیت است که عرضه کنندگان محصولات ویدئویی و تصویری، همچنین سرویس دهندگان رسانه های تصویری، به این نوع ویدئو توجه ویژه ای پیدا کرده اند و رویکرد همه جانبه ای به آن داشته اند [۲]. بنابراین به دلیل حجم بالای اطلاعات ارسالی و لزوم انطباق فنی گیرنده ها برای پخش تصاویر با درجه تفکیک بالا، قابلیت ارسال و دریافت تصاویر ویدئویی با اندازه و نرخ متفاوت، می تواند منجر به ارتقا کیفیت سرویس دهی سیستم شود.

از زمان مطرح شدن ارسال ویدئو با درجه کیفیت بالا، ارسال اطلاعات این نوع ویدئو از طریق دو یا چند لایه مطرح شد [۳]. لایه ای که بتواند اطلاعات لازم برای تصویر ویدئویی استاندارد^۵ را انتقال دهد و لایه ای که با افزوده شدن اطلاعات آن به تصویر ویدئویی استاندارد، ویدئو با درجه تفکیک بالا ارائه کند [۳ و ۴]. بنابراین کد کننده ویدئویی مقیاس پذیر به عنوان یک نیاز مطرح بوده است.

¹ Set Partitioning Hierarchical Trees

² Scalable

³ QoS

⁴ HDTV

⁵ SDTV

هم چنین بکاربردن روش ترکیبی SPIHT-DCT برای ویدئوهای با درجه تفکیک بالا، باعث ارتقا کیفیت نسبت به کدک استاندارد MPEG-2 شده است [۱۷].

در این مقاله، روش کدکردن ترکیبی SPIHT-DCT برای ویدئوهای با درجه تفکیک بالا، HDTV، ابتدا بر اساس مقیاس پذیری مکانی^۱ و سپس مقیاس پذیری مکانی-زمانی^۲ پیشنهاد شده است. اطلاعات ویدئو در حالت اول از طریق دو لایه و در حالت دوم از طریق سه لایه مجزا ارسال می شود. در کدک مقیاس پذیر مکانی، کد کردن هر دو لایه مشابه هم انجام می شود. کد کردن قاب های Intra مبتنی بر تبدیل موجک و کدکردن قاب های Inter بر اساس تبدیل DCT انجام شده است. کدکننده مقیاس پذیر مکانی-زمانی به غیر از دو لایه پایه و ارتقا اول، لایه سوم می دارد که کدکننده آن فقط تحت استاندارد MPEG-2 عمل می کند. دلیل آن وجود قاب های Inter در این لایه است.

هم چنین برای ارزیابی تاثیر کدکننده ترکیبی، ویدئو با درجه تفکیک بالا، با شرایط یکسان، تحت استاندارد MPEG-2 نیز کد شده است.

ادامه این مقاله به این صورت سازماندهی شده است که در بخش دوم الگوریتم کدکردن ترکیبی پیشنهادی تشریح شده است. در بخش سوم کدکننده با مقیاس پذیری مکانی و در بخش چهارم الگوریتم کد کردن ترکیبی با مقیاس پذیری مکانی-زمانی برای ویدئو با درجه تفکیک بالا توضیح داده شده است. بخش پنجم به نتایج شبیه سازی-های انجام شده به هر دو روش مقیاس پذیری اختصاص دارد. در نهایت مقاله در بخش ششم جمع بندی شده است.

۲. کدکردن تصاویر با درجه تفکیک بالا به روش

ترکیبی

برای کد کردن قاب های تصویری با درجه تفکیک بالا، ترکیب الگوریتم SPIHT مبتنی بر تبدیل موجک و تبدیل DCT به کار می رود. در روش ترکیبی SPIHT-DCT، کدکننده به دو بخش مجزا تقسیم می شود. قسمت اول کدکننده قاب های Intra است. این قسمت از کدکننده به طور کاملاً مجزا از روش مبتنی بر استاندارد MPEG-2 عمل می کند. این قاب ها با الگوریتم فشرده سازی SPIHT که مبتنی بر تبدیل موجک است، کد می شوند.

قاب های Inter که شامل قاب های P و B هستند، در بخش دیگر کدکننده بر اساس تبدیل DCT کد می شوند. قاب های Inter با

تصویر ویدئو با درجه تفکیک بالا از عرض و طول دارای وضوح دو برابر نسبت به تصویر ویدئو استاندارد است. اندازه قاب های ویدئوی HDTV ۱۲۸۰×۷۲۰ و ۱۹۲۰×۱۰۸۰ است. نسبت تصویر در ویدئو با درجه تفکیک بالا ۱۶:۹ است. این نسبت در تصویر ویدئو استاندارد برابر ۴:۳ است [۵].

بسیاری از الگوریتم های مورد استفاده در کدکردن ویدئو بر اساس تبدیل DCT پایه گذاری شده است [۶]. در استاندارد MPEG-2 نیز برای حذف افزونگی های مکانی موجود در تصویر از تبدیل DCT استفاده می شود [۷]. از سوی دیگر نتایج آزمایش های سال های اخیر نشان داده است که تبدیل موجک در فشرده سازی تصویر می تواند بر مشکلاتی که الگوریتم های مبتنی بر تبدیل بلوکی ایجاد می کنند، غلبه کند [۶]. هم چنین بکار بردن تبدیل موجک در کدکردن ویدئو می تواند نتایج کیفی بهتری نسبت به کدک های متداول مبتنی بر تبدیل DCT داشته باشد [۸ و ۹].

تبدیل موجک برای کدکردن تصویر با درجه تفکیک بالا نیز مورد استفاده قرار گرفته است [۶]. در روش های مختلف فشرده سازی، با افزایش تعداد پیکسل های یک بلوک تحت پردازش، از همبستگی بیشتر بین پیکسل ها استفاده می شود و فشرده سازی موثرتری انجام می شود [۱۰]. ولی در روش های مبتنی بر تبدیل DCT، بزرگ کردن بلوک تحت تبدیل، موجب پیچیدگی بیشتر می شود [۱۱]. تبدیل موجک از این خاصیت استفاده می کند و با اعمال تبدیل به کل تصویر یا قسمت های بزرگتر از تصویر، نتایج بهتری ارائه می کند. به همین دلیل در استاندارد JPEG2000 نیز از تبدیل موجک استفاده شده است [۱۱].

استفاده از مقیاس پذیری های کدک استاندارد، در روش های کدکردن مبتنی بر تبدیل موجک نیز بکار گرفته شده است [۱۲ و ۱۳]. استفاده از تبدیل موجک در کدکردن ویدئویی به اندازه وضوح CIF با مقیاس پذیری نسبت به استفاده از تبدیل DCT، منجر به ارتقا کیفیت در حدود ۴ dB شد [۱۲].

الگوریتم SPIHT روشی برای فشرده سازی تصویر بر اساس تبدیل موجک است [۱۴ و ۱۵]. با استفاده از این الگوریتم، فقط داده های مهم تصویر و ضرایب حساس تبدیل موجک انتقال پیدا می کنند [۱۱]. مزایای استفاده از تبدیل موجک می تواند برای کد کردن قاب های Intra و تبدیل DCT برای قاب های Inter مورد استفاده قرار گیرد. این روش برای استاندارد ویدئویی H.263 نتایج قابل قبولی به همراه داشت [۱۶]. در این مرجع شبیه سازی ها در نرخ بیت های پایین حدود ۳۰۰ تا ۱۵۰۰ کیلو بیت در ثانیه و روی ویدئوهایی با اندازه CIF و QCIF، انجام شد.

¹ Spatial Scalability

² Spatial-Temporal Scalability

می دهد. با افزوده شدن داده های لایه ارتقا به داده های لایه پایه، ویدئوی با درجه تفکیک بالا در گیرنده دریافت می شود [۱۶].

کد کننده ترکیبی مقیاس پذیر مکانی شامل دو قسمت مجزا از هم است. ساختار کدکننده هر قسمت مشابه هم است.

در شکل (۱) ساختار بلوک دیاگرام کد کننده ترکیبی پیشنهادی با استفاده از ویژگی مقیاس پذیری مکانی برای کد کردن ویدئو با درجه تفکیک بالا نشان داده شده است.

در کدکننده لایه پایه، ابتدا وضوح قاب های تصویر کاهش پیدا می کند. قاب های Intra توسط قسمت کدکننده مبتنی بر الگوریتم SPIHT کد می شوند. این قسمت مجزا از کد کننده ی قاب های Inter است و مبتنی بر تبدیل DCT است. قاب های بازسازی شده I برای پیش بینی و جبران حرکت قاب های P و B بکار می روند.

قاب های کد شده توسط لایه پایه، پس از دکد شدن از طریق مسیر بازگشت در فرستنده به وضوح اصلی تصویر با درجه تفکیک بالا برگردانده می شوند. این قاب ها به همین صورت در گیرنده قابل بازسازی هستند. داده های لایه ارتقا، حاصل تفاضل قاب اصلی تصویر و داده های قابل بازسازی در گیرنده است. در واقع در این مرحله پس از حذف داده های ارسال شده، اطلاعات باقی مانده هر قاب به روش ترکیبی کد و ارسال می شوند.

هر کد کننده سری بیت های مجزایی ایجاد می کند. برای تصویر با وضوح پایین تر، سری بیت های ارسال لایه پایه مورد استفاده قرار می گیرد. سری ویدئویی حاصل از دکد کردن بیت های این لایه، ویدئویی با وضوح SDTV (۷۲۰×۴۸۰) به کاربر ارائه می دهد. سری دیگر که از طریق لایه ارتقا ارسال می شود، شامل اطلاعاتی است که با افزوده شدن به اطلاعات لایه پایه، ویدئوی HDTV (۱۲۸۰×۷۲۰) یا (۱۹۲۰×۱۰۸۰) با وضوح کامل را عرضه می کند. با استفاده از این ویژگی در گیرنده، یک ویدئو با دو وضوح متفاوت وجود دارد که کاربر با توجه به نیاز خود، سرویس دلخواه را انتخاب می کند.

تخمین و جبران سازی حرکت در هر لایه به طور مستقل از لایه دیگر محاسبه و ارسال می شود. به دلیل اختلاف وضوح تصویر در دو قسمت کد کننده، نمی توان از نتایج حاصل از پیاده سازی الگوریتم های پیش بینی و جبران حرکت و الگوریتم های تخمین بردار حرکت (Motion Vector) یک لایه برای لایه دیگر استفاده کرد.

بکاربردن الگوریتم های پیش بینی و جبران حرکت روی قاب مرجع ایجاد می شوند. پیاده سازی الگوریتم های پیش بینی و جبران حرکت روی قاب های تصویری منجر به ایجاد فرکانس های بالای مصنوعی می شود. اگر این فرکانس های بالا تحت تبدیل الگوریتم های مبتنی بر تبدیل موجک قرار گیرد، می تواند منجر به ایجاد اعوجاج شود. اعوجاج ایجاد شده در این قسمت مزایای استفاده از تبدیل موجک در کد کردن ویدئو را تحت تاثیر قرار می دهد [۱۶].

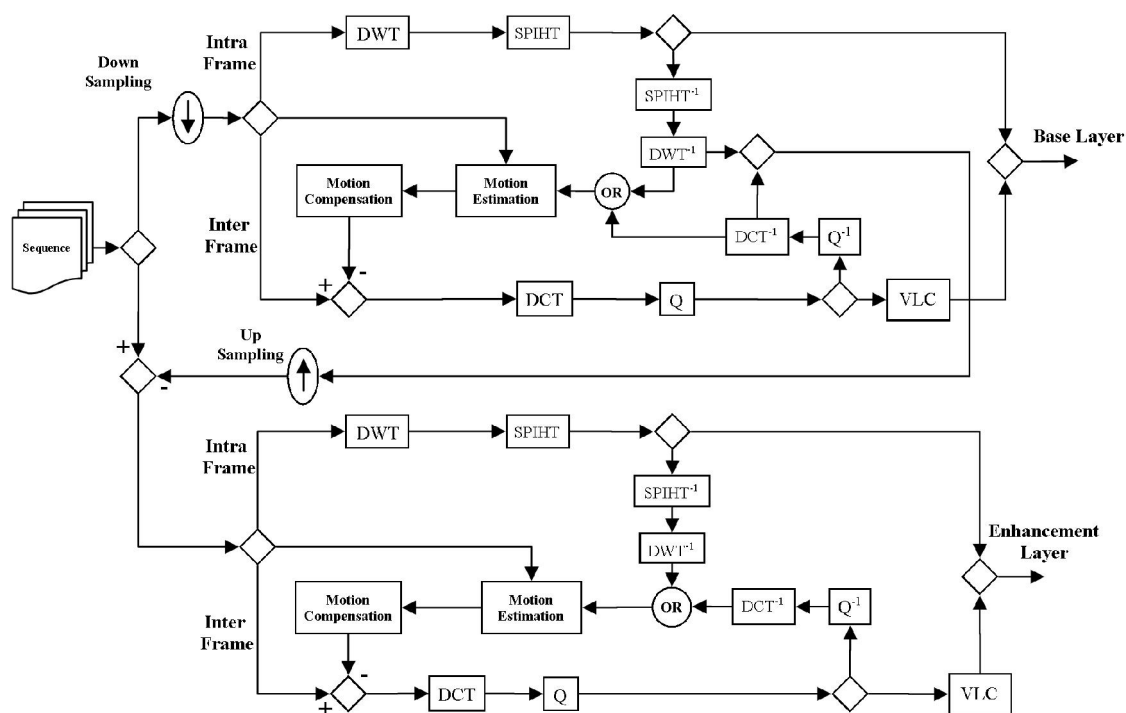
مرجع قاب های P و B، قاب I کد شده به روش مبتنی بر تبدیل موجک است. در کد کردن قاب های I با افزایش لایه های تجزیه در الگوریتم استفاده شده از همبستگی بین همه نمونه های تصویر استفاده می شود. ارتقا کیفیت قاب مرجع، منجر به بهبود کیفیت قاب های Inter می شود و می تواند در کیفیت کل ویدئوی ارسالی موثر واقع شود.

الگوریتم پیش بینی حرکت، یک مدل از قاب جاری را بر اساس داده های موجود در قاب های قبلی و بعدی کد شده ایجاد می کند. هدف از بکارگیری الگوریتم پیش بینی حرکت، مدل کردن قاب جاری با حداکثر دقت ممکن است تا اطلاعات ارسالی کاهش یابد. با کم کردن قاب مدل شده از قاب جاری، عمل جبران حرکت روی قاب انجام می شود [۱۱]. پساز جبران حرکت، قابی بدست می آید که حاوی اطلاعاتی است که دکد کننده برای بازسازی قاب جاری از روی قاب مرجع نیاز دارد. بهترین عملکرد فشرده سازی زمانی هنگامی بدست می آید که داده های قاب باقی مانده از الگوریتم جبران حرکت به کمترین مقدار برسد، در حالی که در دکد کننده بهترین کیفیت ممکن ایجاد شود. الگوی قرارگیری و کد شدن قاب ها را کد کننده مشخص می کند، در عین حال الگوی متداول ترتیب قاب ها به این صورت است: I B B P B B P B B [۵].

۳. روش کد کردن ترکیبی SPIHT-DCT با مقیاس

پذیری مکانی برای ویدئو با درجه تفکیک بالا

برای کد کردن ترکیبی ویدئوهای با درجه تفکیک بالا، مقیاس پذیری مکانی به کدکننده افزوده شده است. با استفاده از این ویژگی، دو لایه برای انتقال داده ها ایجاد می شود. لایه اول یا لایه پایه، داده های ویدئو با وضوح پایین تر را انتقال می دهد. لایه دوم داده های ارسال نشده برای ارتقا ویدئو به وضوح اصلی تصویر لازم استرا انتقال



شکل ۱: بلوک دیاگرام کد کننده ترکیبی با استفاده از مقیاس پذیری مکانی.

وضوح کامل را شامل می شود که از طریق لایه پایه ارسال نشده اند. در نتیجه مجموع اطلاعات سه لایه، ویدئو با درجه تفکیک بالا با وضوح و نرخ قاب در ثانیه ویدئوی اصلی است.

بلوک دیاگرام کد کننده ترکیبی پیشنهادی ویدئو با درجه تفکیک بالا با ویژگی مقیاس پذیری مکانی- زمانی در شکل (۲) نشان داده شده است.

در این کد کننده سه واحد جداگانه برای اجرای الگوریتم های پیش بینی و جبران حرکت وجود دارد. این الگوریتم ها برای قاب های با وضوح پایین تر به طور جداگانه محاسبه می شوند. بردارهای حرکتی این قاب ها هم مجزا از قسمت های دیگر محاسبه و ارسال می شوند. در لایه ارتقا اول نیز به دلیل افزایش وضوح تصویر باید الگوریتم پیش بینی و جبران حرکت جداگانه محاسبه شود. در لایه ارتقا دوم، قاب هایی ارسال می شوند که در لایه های قبلی کد و ارسال نشده اند. بنابراین این الگوریتم ها برای این قاب ها به طور مجزا انجام می شود.

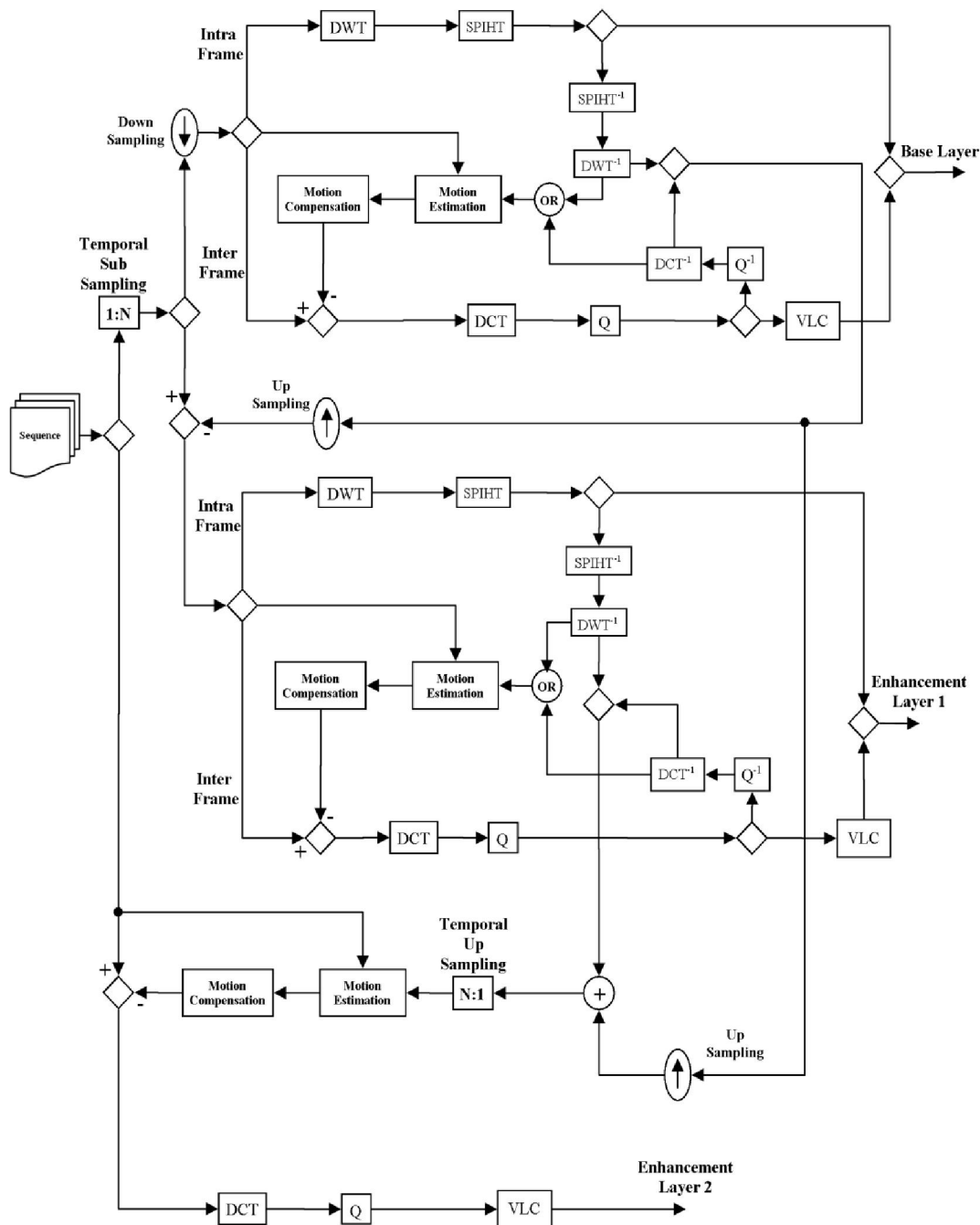
این کد کننده ترکیبی، ضمن ارتقا کیفیت قاب های مرجع که می تواند منجر به افزایش کیفیت سایر قاب های ویدئو شود، از مقیاس پذیری مکانی- زمانی نیز استفاده می کند. بنابراین کاربر می تواند رنج گسترده تری از سرویس ویدئویی در اختیار داشته باشد و بر اساس نیاز، سرویس مورد نظر خود را انتخاب نماید.

۴. روش کد کردن ترکیبی SPIHT-DCT با مقیاس پذیری مکانی- زمانی برای ویدئو با درجه تفکیک بالا

در این قسمت به کد کننده ترکیبی پیشنهادی، ویژگی مقیاس پذیری مکانی- زمانی اضافه شده است. در مقیاس پذیری مکانی، قاب های تصویر ویدئو با اندازه وضوح متفاوت ارسال می شوند. حال آنکه در مقیاس پذیری زمانی نرخ قاب در ثانیه لایه های ارسالی نیز تغییر می کند. در واقع یک قاب ویدئویی با درجه تفکیک بالا در این کد کننده از سه لایه مختلف ارسال می شود.

لایه اول یا لایه پایه، اطلاعات پایین ترین سطح وضوح و نرخ قاب در ثانیه ویدئو را انتقال می دهد. داده های انتقالی توسط این لایه، ویدئو SDTV با اندازه قاب 720×480 و نرخ ۲۵ قاب در ثانیه عرضه می کنند. در این لایه قاب های I، P، و برخی قاب های B قرار دارند. اطلاعات لایه دوم یا لایه ارتقا اول، با افزوده شدن به داده های لایه پایه، وضوح قاب های ارسالی لایه پایه را به اندازه واقعی می رسانند. در واقع ویدئو عرضه شده با مجموع اطلاعات لایه پایه و لایه ارتقا اول، نسبت به ویدئو با درجه تفکیک بالا، فقط نرخ قاب در ثانیه کمتری دارد.

اطلاعات انتقال داده شده توسط لایه سوم یا لایه ارتقا دوم، به اطلاعات دو لایه پایین ترافزوده می شوند. این لایه قاب هایی با



شکل ۲: بلوک دیاگرام کدکننده ترکیبی با استفاده از ویژگی مقیاس پذیری مکانی-زمانی.

۵. شبیه سازی ها و نتایج

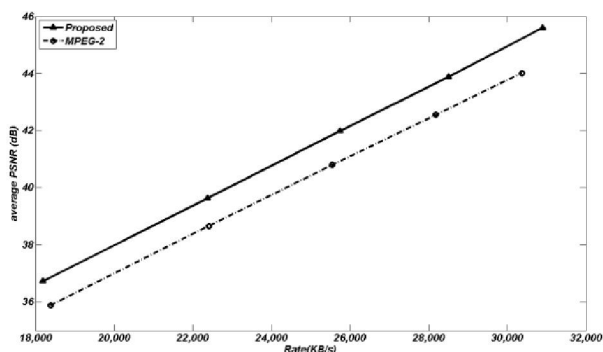
۵.۱. کدکننده ترکیبی با مقیاس پذیری مکانی

برای کدکردن ویدئو با درجه تفکیک بالا به روش ترکیبی پیشنهادی و با ویژگی مقیاس پذیری مکانی، از ویدئوهای با وضوح متفاوت استفاده شده است. اندازه وضوح ویدئوهای تحت شبیه سازی ۱۲۸۰×۷۲۰ و ۱۹۲۰×۱۰۸۰ است. شبیه سازی‌ها روی مولفه روشنایی

قاب‌های ویدئویی انجام شده است. هم چنین نرخ قاب های Intra در دو حالت متفاوت ۱ به ۵ و ۱ به ۹ در نظر گرفته شده است. در نرخ Intra ۱ به ۹ از همان الگوی قاب بیان شده استفاده شده است و در نرخ Intra ۱ به ۵ همان الگو با قطع پنجمین قاب به صورت I B B P B بکار رفته است. شکل‌های (۴ و ۳) نشان دهنده میانگین کیفیت ویدئو Shields به وضوح ۱۲۸۰×۷۲۰ در نرخ بیت های مختلف است. اختلاف بین میانگین کیفیت ویدئو کد شده با افزایش نرخ بیت تقریباً ثابت است و

مختلف را نشان می دهد. هم چنین نرخ بیت انتقال داده شده توسط لایه پایه نیز مشخص شده است. اندازه قاب های انتقال داده شده توسط لایه پایه 720×480 است. لایه پایه حدود ۳۰ درصد مجموع بیت ارسالی را به خود اختصاص داده است.

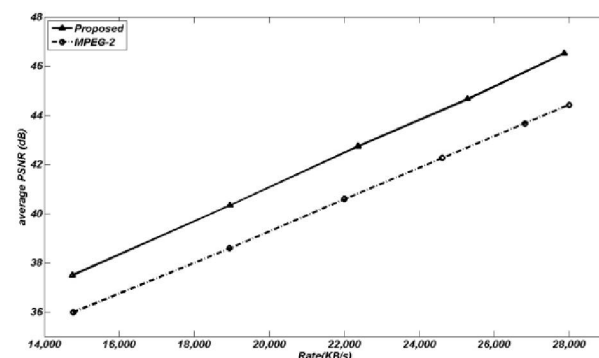
در شکل های (۵ و ۶) نمودار میانگین کیفیت ویدئو Mobcal به اندازه 1280×720 بر حسب نرخ بیت های مختلف نشان داده شده است. در شکل ۵ با نرخ ایترا ۱ به ۵، افزایش نرخ بیت، اختلاف بین میانگین کیفیت روش پیشنهادی و روش استاندارد را افزایش داده است. با افزایش نرخ بیت، روند صعودی میانگین کیفیت ویدئو کد شده به روش استاندارد کاهش می یابد. برای ویدئو کد شده به روش ترکیبی پیشنهادی، روند افزایش کیفیت تقریباً ثابت است. شکل (۶) نشان می دهد که با نرخ ایترا ۱ به ۹، اختلاف ویدئو کد شده به روش پیشنهادی و استاندارد با افزایش نرخ بیت کمی افزایش می یابد و در نرخ بیت 29900 KB/s حدود $1/4$ dB است.



شکل ۴: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه 1280×720 در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

روند افزایش تقریباً مشابهی را دارد. در نرخ ایترا ۱ به ۵ بهبود میانگین کیفیت روش پیشنهادی نسبت به روش استاندارد بیش از مورد مشابه در نرخ ایترا ۱ به ۹ است. علت کاهش میانگین کیفیت ویدئو با کاهش نرخ قاب های ایترا، کاهش اثر قاب مرجع کد شده به روش مبتنی بر تبدیل موجک در سایر قاب ها است. هم چنین با توجه به نتایج حاصل از شبیه سازی روش پیشنهادی بدون در نظر گرفتن مقیاس پذیری، می توان این نکته را بیان کرد که روش ترکیبی پیشنهادی به همراه ویژگی مقیاس پذیری مکانی بهتر از روش استاندارد به همراه این ویژگی عمل می کند و با دو لایه شدن اطلاعات ارسالی عملکرد بهتری نسبت به روش استاندارد دارد [۱۷]. در نرخ ایترا ۱ به ۹ نتایج حاصل از روش مبتنی بر تبدیل موجک در نرخ بیت های بالا، نزدیک به نتایج روش ترکیبی است. در عین حال با نزدیک شدن به نرخ بیت 30000 KB/s میانگین کیفیت روش مبتنی بر تبدیل موجک اندکی کاهش می یابد.

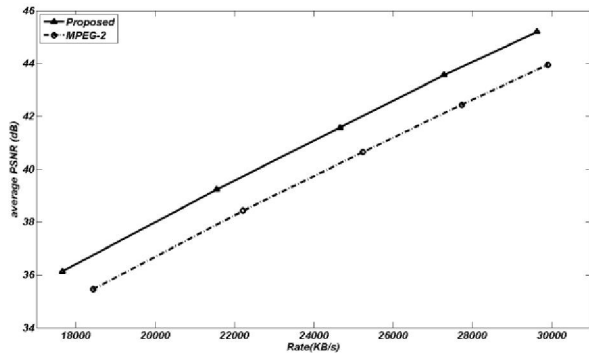
جدول (۱) نتایج شبیه سازی انجام شده روی ویدئو Shields به اندازه 1280×720 را در نرخ بیت های تقریباً یکسان برای دو لایه



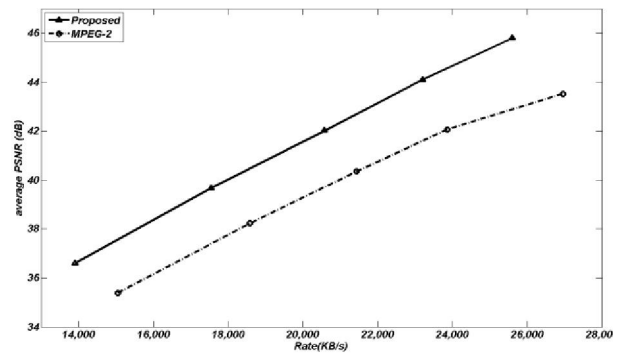
شکل ۳: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه 1280×720 در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

جدول ۱: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه 1280×720 برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

| Sequence | Size | frame/sec | Intra Coding Rate | Compression Method | Bit Rate (KB/s) | | Avg. PSNR (dB) |
|-------------|----------|-----------|-------------------|--------------------|-----------------|--------|----------------|
| | | | | | SDTV | HDTV | HDTV |
| Shields_ter | 1280*720 | 50 | 1/5 | MPEG-2 | 5,848 | 18,937 | 38.59 |
| | | | | Proposed | 5,469 | 18,947 | 40.34 |
| | | | 1/9 | MPEG-2 | 8,627 | 25,541 | 40.8 |
| | | | | Proposed | 8,425 | 25,751 | 41.99 |



شکل ۶: مقایسه میانگین کیفیت لایه های ویدئو Mobcal به اندازه ۱۲۸۰×۷۲۰ در نرخ اینترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.



شکل ۵: مقایسه میانگین کیفیت لایه های ویدئو Mobcal به اندازه ۱۲۸۰×۷۲۰ در نرخ اینترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

جدول ۲: مقایسه نتایج شبیه سازی ویدئو Mobcal به اندازه ۱۲۸۰×۷۲۰

برای نرخ ارسال بیستگرمیاً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

| Sequence | Size | frame/sec | Intra Coding Rate | Compression Method | Bit Rate (KB/s) | | Avg. PSNR (dB) |
|------------|----------|-----------|-------------------|--------------------|-----------------|--------|----------------|
| | | | | | SDTV | HDTV | HDTV |
| Mobcal_ter | 1280*720 | 59.94 | 1/5 | MPEG-2 | 7,466 | 23,862 | 42.07 |
| | | | | Proposed | 6,883 | 23,207 | 44.1 |
| | | | 1/9 | MPEG-2 | 10,419 | 29,905 | 43.93 |
| | | | | Proposed | 10,060 | 29,625 | 45.19 |

با افزایش نرخ بیت، اختلاف میانگین کیفیت ابتدا کمی کاهش و سپس اندکی افزایش می یابد.

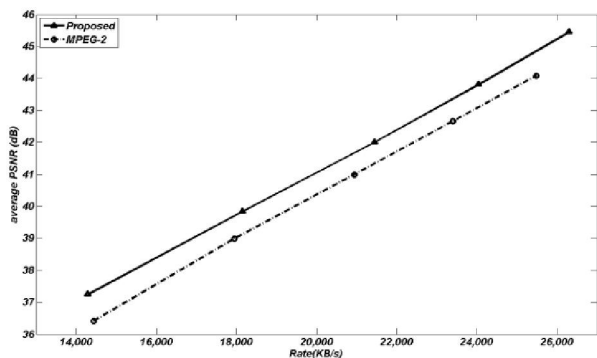
شکل (۸) در نرخ اینترا ۱ به ۹، نشان می دهد که اختلاف میانگین کیفیت دو روش کدکردن نسبت به نرخ اینترا بالاتر، کاهش یافته که به دلیل میرا بودن اثر بهبود کیفیت قاب کد شده به روش تبدیل موجک در قاب های بعدی است. در نرخ بیت برابر، میانگین کیفیت ویدئو کد شده در نرخ اینترا ۱ به ۵ حدود ۱/۶ dB بیشتر از ویدئو کد شده در نرخ ۱ به ۹ است.

جدول (۳) نتایج میانگین کیفیت ویدئو Shields به اندازه ۱۹۲۰×۱۰۸۰ در نرخ بیت تقریباً یکسان نشان می دهد. حجم اطلاعات ارسالی توسط لایه پایه هم در هر نرخ بیت مشخص شده است. در نرخ اینترا ۱ به ۵ و در نرخ بیت حدود ۲۰۰۰۰ KB/s اختلاف میانگین کیفیت دو روش کدکردن به حدود ۱/۵ dB می رسد. در نرخ اینترا ۱ به ۹ و در نرخ بیت حدود ۲۱۰۰۰ KB/s اختلاف میانگین کیفیت دو روش به حدود ۱ dB می رسد. لایه پایه حدود ۲۹ تا ۳۱ درصد حجم اطلاعات ارسالی را تشکیل می دهد.

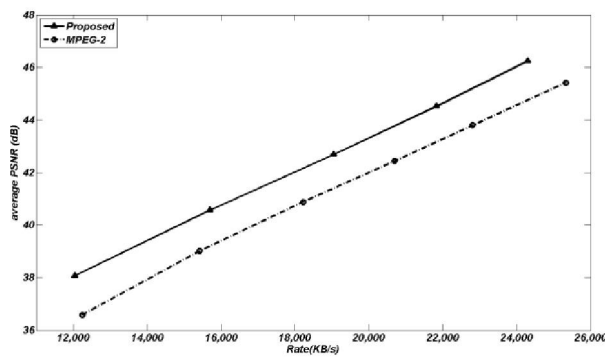
در جدول (۲) مقایسه میانگین کیفیت ویدئو Mobcal به اندازه ۱۲۸۰×۷۲۰ کد شده به سه روش پیشنهادی و تبدیل موجک و استاندارد در نرخ بیت تقریباً یکسان نشان داده شده است. حجم اطلاعات لایه پایه به اندازه وضوح ۷۲۰×۴۸۰ در مقایسه با حجم بیت ارسالی برای ویدئو با درجه تفکیک بالا نیز نشان داده شده است. اطلاعات ارسالی لایه پایه حدود ۳۰ درصد از حجم ارسالی برای انتقال ویدئو کامل را تشکیل می دهد.

شکل های (۷ و ۸) نمودارهای مقایسه میانگین کیفیت برای ویدئو Shields به اندازه ۱۹۲۰×۱۰۸۰ را نشان می دهد. این ویدئو به دو روش ترکیبی و استاندارد و با ویژگی مقیاس پذیری مکانی در دو نرخ بیت متفاوت کد شده است.

شکل (۷) در نرخ اینترا ۱ به ۵ است و شبیه سازی از ۱۲۰۰۰ KB/s تا حدود ۲۶۰۰۰ KB/s انجام شده است. اختلاف میانگین کیفیت بین دو روش کدکردن ترکیبی پیشنهادی و روش استاندارد به ۱/۵ dB هم می رسد. نتایج حاصل شده نشان می دهد که



شکل ۸: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه 1920×1080 در نرخ اینترا ۱ به ۹ کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.



شکل ۷: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه 1920×1080 در نرخ اینترا ۱ به ۵ کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.

جدول ۳: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه 1920×1080 برای ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.

| Size | frame/sec | Intra Coding Rate | Compression Method | Bit Rate (KB/s) | | Avg. PSNR (dB) |
|-----------|-----------|-------------------|--------------------|-----------------|--------|----------------|
| | | | | SDTV | HDTV | HDTV |
| 1920*1080 | 25 | 1/5 | MPEG-2 | 5,949 | 20,696 | 42.44 |
| | | | Proposed | 5,333 | 19,051 | 42.69 |
| | | 1/9 | MPEG-2 | 7,012 | 20,941 | 41 |
| | | | Proposed | 7,052 | 21,443 | 42 |

جدول (۴)، میانگین کیفیت ویدئو را در یک نرخ بیت در ثانیه تقریباً یکسان و با نرخ Intra متفاوت نشان می دهد.

شکل های ۱۱ و ۱۲ مقایسه میانگین کیفیت ویدئوی Mobcal به اندازه وضوح 1280×720 با نرخ ۵۰ قاب در ثانیه است. مقایسه در دو نرخ Intra متفاوت ۱ به ۵ و ۱ به ۹ انجام شده است. شبیه سازی کد کردن این ویدئو نیز روی ۵۰ قاب انجام شده است.

با مشاهده نتایج این ویدئو، با افزایش نرخ بیت در ثانیه میانگین کیفیت لایه های ویدئو کد شده به روش ترکیب نسبت به ویدئو کد شده به روش استاندارد افزایش پیدا می کند. با کاهش نرخ قاب Intra به نرخ ۱ به ۹، اختلاف کیفیت دو روش کاهش می یابد. این نتایج نشان می دهد با افزایش قاب های Inter وابسته، اثر ارتقا قاب مرجع در آنها میرا می شود. جدول (۵)، نتایج شبیه سازی ویدئو به هر دو روش ترکیبی و استاندارد را در نرخ بیت ارسالی تقریباً یکسان نشان می دهد. نتایج در دو نرخ متفاوت Intra نشان داده شده است. لایه پایه در نرخ Intra ۱ به ۵ حدود ۲۰ درصد حجم اطلاعات ارسالی را تشکیل می دهد. در نرخ ۱ به ۹ این مقدار به ۱۹ درصد می رسد.

اختلاف میانگین کیفیت لایه های ویدئو Shields با اندازه وضوح 1920×1080 با نرخ ۲۵ قاب در ثانیه در دو نرخ Intra متفاوت در شکل های ۱۳ و ۱۴ نشان داده شده است. اختلاف بین ویدئو کامل و

۵.۲. کد کننده ترکیبی با مقیاس پذیری مکانی-زمانی

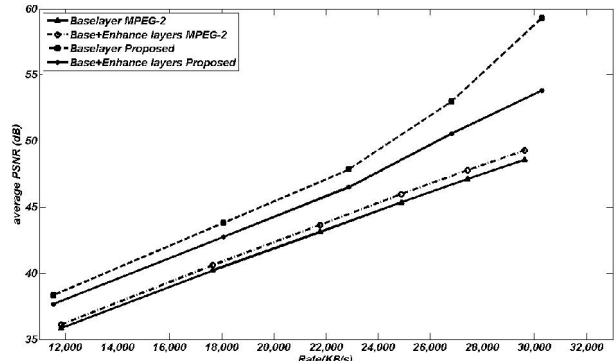
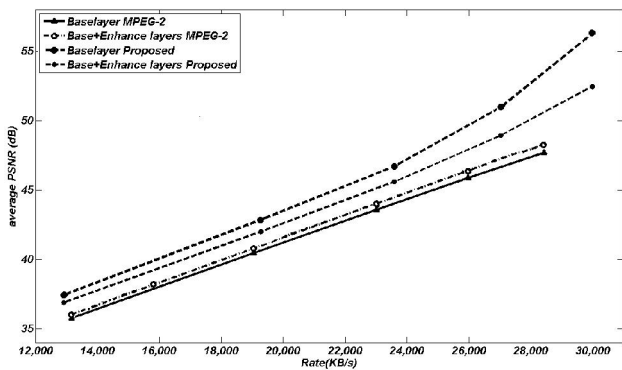
عملکرد کد کننده ترکیبی با مقیاس پذیری مکانی-زمانی با کد کننده استاندارد MPEG-2 مقایسه شده است. ویدئوهای با درجه تفکیک بالا با وضوح متفاوتی برای شبیه سازی ها در نظر گرفته شده است.

شکل های ۹ و ۱۰، میانگین کیفیت ویدئوی Shields با اندازه وضوح 1280×720 را نسبت به نرخ های بیت بر ثانیه مختلف در دو نرخ Intra ۱ به ۵ و ۱ به ۹ نشان می دهند. در این شکل ها میانگین کیفیت ویدئو کد شده به دو روش متفاوت با هم مقایسه شده است. شبیه سازی روی این ویدئو برای ۵۰ قاب انجام شده است.

ترکیب لایه پایه و لایه ارتقا اول کد شده به روش ترکیبی پیشنهادی بیشترین کیفیت را دارد. ویدئو این دو لایه با نرخ ۲۵ قاب در ثانیه است. در واقع نصف نرخ قاب در ثانیه ویدئو اصلی است. با افزایش نرخ بیت، روند افزایش کیفیت سریع تر می شود. نرخ بیت ارسالی برای لایه پایه و لایه ارتقا اول همان نرخ بیت ارسالی کل ویدئو است که ترکیب این دو لایه بخشی از آن را شامل می شود. در مجموع اختلاف میانگین کیفیت ویدئو کامل کد شده به روش ترکیبی نسبت به روش استاندارد، با افزایش نرخ بیت در ثانیه زیاد می شود.

و در نرخ **Intra** متفاوت نشان داده شده است. از این جدول می توان درصد اشغال شده پهنای باند توسط لایه های پایه و ارتقا اول را بدست آورد. ویدئو لایه پایه در هر دو نرخ **Intra** حدود ۹ درصد پهنای باند را اشغال می کند. میانگین کیفیت ویدئو با درجه تفکیک بالا و ویدئو حاصل از لایه های پایه و ارتقا اول، در نرخ بیت ارسال تقریباً یکسان نیز، مشخص است. لایه پایه برای این ویدئو همان اندازه وضوح **SDTV** یا 720×480 را ارسال می کند. هم چنین ویدئو حاصل از لایه پایه و ترکیب لایه پایه و لایه ارتقا اول نرخ ۱۳ قاب در ثانیه را دارد.

ویدئو حاصل از مجموع لایه پایه و لایه ارتقا در روش ترکیبی بیشتر از اختلاف لایه های مشابه در روش استاندارد است. با افزایش نرخ بیت ارسال، این اختلاف در روش ترکیبی افزایش می یابد، در صورتی که در روش استاندارد، این اختلاف تقریباً ثابت است. نتایج نشان می دهد که با افزایش اندازه وضوح، در نرخ بیت ارسال یکسان، میانگین کیفیت ویدئو نیز افزایش پیدا می کند. در نرخ **Intra** ۱ به ۹ اختلاف میانگین کیفیت ویدئو کد شده به روش استاندارد و ترکیبی نسبت به نرخ **Intra** ۱ به ۵ کاهش یافته است. جدول (۶) نتایج شبیه سازی ویدئو را در نرخ بیت های تقریباً یکسان



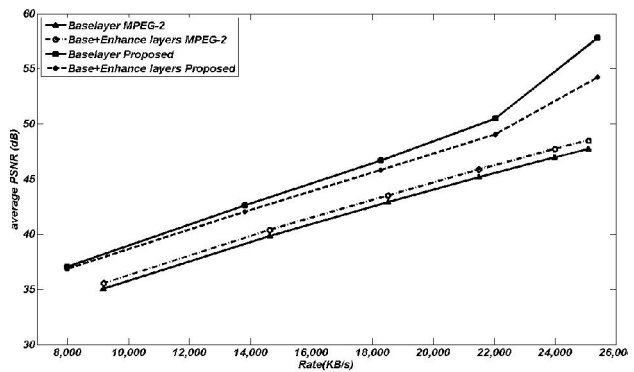
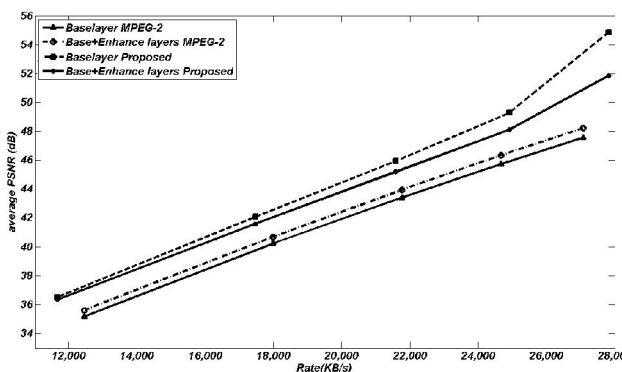
شکل ۱۰: میانگین کیفیت لایه های ویدئو **Shields** به اندازه 1280×720 در نرخ اینترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی..

شکل ۹: میانگین کیفیت لایه های ویدئو **Shields** به اندازه 1280×720 در نرخ اینترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

جدول ۴: مقایسه نتایج شبیه سازی ویدئو **Shields** به اندازه 1280×720 برای ارسال بیت

تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی

| Sequence | fps | Intra Rate | compression Method | Bit Rate (KB/s) | | | Avg. PSNR (dB) | |
|----------|-----|------------|--------------------|-----------------|----------------|--------|----------------|-------|
| | | | | Base | Base+ Enhance1 | HDTV | Base+ Enhance1 | HDTV |
| Shields | 50 | 1/5 | MPEG-2 | 4,378 | 14,275 | 21,774 | 43.13 | 43.67 |
| | | | Proposed | 4,207 | 15,405 | 22,884 | 47.86 | 46.52 |
| | | 1/9 | MPEG-2 | 4,173 | 13,607 | 23,005 | 43.56 | 44.02 |
| | | | Proposed | 4,066 | 14,216 | 23,591 | 46.72 | 45.6 |

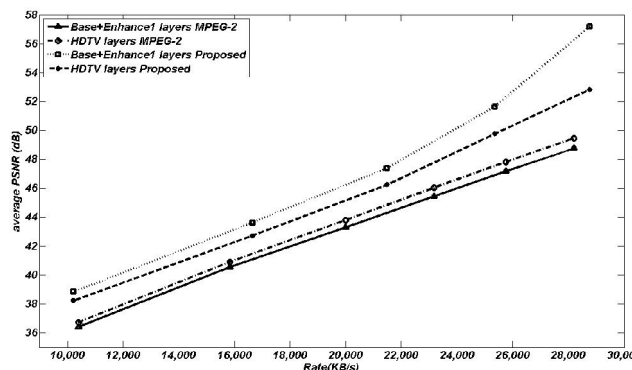
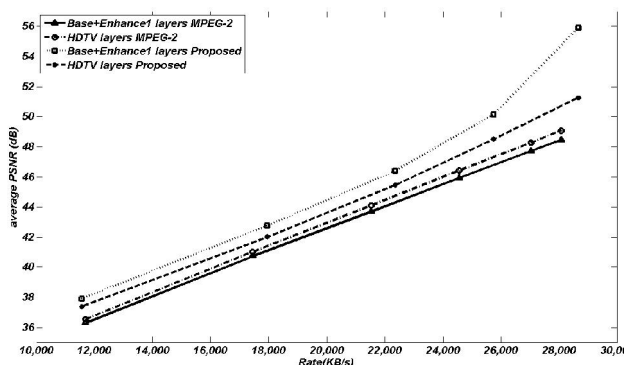


شکل ۱۲: میانگین کیفیت لایه های ویدئو **Mobcal** به اندازه 1280×720 در نرخ اینترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

شکل ۱۱: میانگین کیفیت لایه های ویدئو **Mobcal** به اندازه 1280×720 در نرخ اینترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

جدول ۵: مقایسه نتایج شبیه سازی ویدئو Mobcal به اندازه ۷۲۰×۱۲۸۰ برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

| Sequence | fps | Intra Rate | compression Method | Bit Rate (KB/s) | | | Avg. PSNR (dB) | |
|----------|-----|------------|--------------------|-----------------|----------------|--------|----------------|-------|
| | | | | Base | Base+ Enhance1 | HDTV | Base+ Enhance1 | HDTV |
| Mobcal | 50 | 1/5 | MPEG-2 | 4,408 | 14,270 | 20,515 | 42.88 | 43.53 |
| | | | Proposed | 3,953 | 14,151 | 20,286 | 46.69 | 45.81 |
| | | 1/9 | MPEG-2 | 4,813 | 15,474 | 24,668 | 45.73 | 46.33 |
| | | | Proposed | 4,629 | 15,776 | 24,917 | 49.29 | 48.11 |



شکل ۱۴: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه ۱۰۸۰×۱۹۲۰ در نرخ اینترا ۱ به ۹ کد شده به دو روش پیشنهادی و استاندارد.

شکل ۱۳: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه ۱۰۸۰×۱۹۲۰ در نرخ اینترا ۱ به ۵ کد شده به دو روش پیشنهادی و استاندارد.

جدول ۶: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه ۱۰۸۰×۱۹۲۰ برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش پیشنهادی و استاندارد.

| Sequence | fps | Intra Rate | compression Method | Bit Rate (KB/s) | | | Avg. PSNR (dB) | |
|----------|-----|------------|--------------------|-----------------|----------------|--------|----------------|-------|
| | | | | Base | Base+ Enhance1 | HDTV | Base+ Enhance1 | HDTV |
| Shields | 25 | 1/5 | MPEG-2 | 2,203 | 16,510 | 25,758 | 47.16 | 47.84 |
| | | | Proposed | 2,236 | 16,965 | 25,357 | 51.65 | 49.77 |
| | | 1/9 | MPEG-2 | 1,497 | 10,302 | 17,448 | 40.73 | 41.02 |
| | | | Proposed | 1,520 | 10,809 | 17,949 | 42.77 | 42.01 |

هایی می شود که از آن به عنوان قاب مرجع استفاده می کنند. استفاده از ویژگی مقیاس پذیری، ضمن انعطاف پذیر کردن کدکننده، باعث ارتقا کیفیت سرویس نیز می شود. مقایسه نتایج کدکننده پیشنهادی برای ویدئو با درجه تفکیک بالا با سایر کدک های استاندارد مانند MPEG-4 می تواند زمینه ساز تحقیقات در آینده باشد.

مراجع

[1].M. Domenski, A. Luczuk, S. Mackwiak, "On Improving MPEG Spatial Scalability", in proceedings of 2000 international conference of image processing ICIP Vancouver, Canada, vol. 2, pp. 848-851, 2000.

۶. نتیجه گیری

در این مقاله، کدکننده ترکیبی برای ویدئو با درجه تفکیک بالا پیشنهاد شده است. این کدکننده در دو حالت مختلف طراحی و شبیه سازی شده است. ویدئو با درجه تفکیک بالا ابتدا با استفاده از مقیاس پذیری مکانی و سپس با بکار بردن مقیاس پذیری مکانی-زمانی کد و ارسال شده است. بکارگیری روش پیشنهاد شده در این کدکننده، منجر به بهبود کیفیت ویدئو ارسالی نسبت به کدکننده مبتنی بر استاندارد MPEG-2 می شود. قاب های Intra با استفاده از الگوریتم SPIHT کد می شوند و کیفیت آن نسبت به قاب های Intra کد شده به روش استاندارد افزایش پیدا می کند. این بهبود منجر به افزایش کیفیت قاب

- [2]. L. Lima, Manerba F, Adami N, Signoroni A, Leonardi R, "Wavelet-Based Encoding for HD Applications", IEEE, ICME, pp. 1351-1354, 2007.
- [3]. T. Chiang, D. Anastassiou, "Hierarchical HDTV/SDTV Compatible Coding Using Kalman Statistical Filtering", IEEE Trans. On circuits for video Tech. Vol. 9, No. 3, 1999.
- [4]. M. Domenski, A. Luczak, S. Mackowaik, R. Swierczynski, "Hybrid coding of video with spatio-temporal Scalability using subband decomposition", in Proceedings 1999 of SPIE, vol. 3653, pp. 1018-1025, 1999.
- [5]. C. Poynton, "Digital Video and HDVT Algorithms and Interfaces", Los Altos, USA, pp. 111-126, 2003.
- [6]. S. Uzun, A. Amira "Real-time 2-D Wavelet Transform Implementation for HDTV Compression", in Elsevier Science, Real-Time Imaging 11, pp. 151-165, 2005.
- [7]. ISO/IEC International Standard 13813-2, information Technology Generic Coding of Moving Pictures and Associated Audio Information, 1995.
- [8]. J. Fowler, M. Tagliasacchi, B. Pesquet, "Wavelet-Based Distributed Source Coding of Video", In Proc. of the European signal Processing Conference, EUSIPCO, Antalia, Turkey, 2005.
- [9]. M. Weeks, M. A. Bayoumi, "Three-Dimensional Discrete Wavelet Transform Architectures", IEEE transactions on signal processing, 2002, vol. 50, No. 8.
- [10]. G. Kwon, Hyo-Kak Kim, Yoon Kim, Sung-Jea Ko, "An Efficient POCS-based Post-processing Technique Using Wavelet Transform in HDTV", IEEE Transactions Consumer Electronics, Vol. 51, No. 4, Pages: 1283-1288, 2005.
- [11]. M. Ghanbari, "Standard Codecs: Image Compression to Advanced Video Coding", IEE Telecommunication 49, London, U.K., 2003.
- [12]. B. Wu, X. Ji, D. Zhoa, W. Gao, "Wavelet Based distributed video coding with spatial scalability", circuits and systems, ISCAS, 2008 IEEE international symposium, pp. 3458-3461, 2008.
- [13]. N. Adami, M. Brescianini, Riccardo Leonardi, A. Signoroni. "A Fully Scalable Video Coding Scheme with Homologous Inter-scale Prediction.", ST Journal of Research, Vol. 3, No. 2, Pages 19-35, 2006.
- [14]. A. Said, W.a. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", in IEEE trans. circuits system video technology, vol.6, no. 3, pp. 243-250, 1996.
- [15]. N. Bozinovic, J. Konrad, "Modeling Motion for Spatial Svalability", Acoustics, Speech and Signal Processing ICASSP, Proceedings, 2006 IEEE Intenational conference, Vol. 2, PP. 29-32.
- [16]. R. Dianat, M Ghanbari, F. Marvasti "A Low Bit Rate Hybrid Wavelet-DCT Video Codec" IEEE Tran. On circuits and systems for Video Technology, VOL. 15, NO. 7, July 2005
- [17]. وحیدرضا صیرفیان، سیامک طالبی، "کدکردن تصاویر ویدئویی HDTV با استفاده از روش ترکیبی SPIHT-DCT"، هفدهمین کنفرانس مهندسی برق ایران، ICEE2009، تهران، دانشگاه علم و صنعت، ش. ۲۲۵۴، بهار ۱۳۸۸.

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه

جواد راستی^{۱*}

سید امیرحسین منجمی^{**}

عباس وفایی^{***}

* استادیار، دانشکده مهندسی پزشکی، دانشگاه اصفهان، اصفهان

** دانشیار، دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان

*** استادیار، دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان

تاریخ دریافت: ۱۳۹۱/۰۴/۲۰

تاریخ پذیرش: ۱۳۹۱/۱۱/۳۰

چکیده

یکی از مسایل مهم در پردازش خودکار تصاویر بیرونی، نحوه بخش‌بندی این تصاویر به هدف تشخیص شیء در آنها می‌باشد. مشخصات خاص این تصاویر از جمله تنوع رنگ، اثرات نوری متفاوت، وجود سایه‌های رنگی، جزئیات بافتی زیاد و وجود اشیاء کوچک و ناهمگن باعث می‌شود مسأله بخش‌بندی تصاویر بیرونی به ویژه بخش‌بندی رنگی با چالش‌های جدی مواجه شود. در تحقیقات قبلی برای خوشه‌بندی رنگی تصاویر بیرونی به هدف بخش‌بندی ابتدایی، روشی مبتنی بر الگوریتم خوشه‌بندی k-means در بستری با دقت چندگانه پیشنهاد شده بود. این روش با استفاده از محو عمودی جزئیات بافتی تصویر و حذف کلاسهای محرز در تصاویر محو شده و سپس اضافه کردن کلاسها در تصاویر با دقت بالاتر، کارایی مناسبی برای بخش‌بندی ابتدایی این تصاویر در مقایسه با روش k-means عادی نشان می‌داد. در این مقاله، یک روش تطبیق‌پذیر با تصویر با استفاده از هیستوگرام حلقوی ته‌رنگ برای تشخیص کلاس‌های محرز در تصاویر محوشده در بستری با دقت دوگانه پیشنهاد گردیده است. کارایی این الگوریتم به کمک یک روش ارزیابی‌نظارت‌شده روی دو پایگاه داده از تصاویر بیرونی بررسی شده که حدود ۲۰٪ کاهش خطای پیکسلی در بخش‌بندی و نیز دقت و حدود ۳۰٪ سرعت بیشتر در همگرایی الگوریتم خوشه‌بندی، نشانگر کیفیت بالاتر روش پیشنهادی نسبت به روش عادی است.

کلید واژگان: تصاویر بیرونی، خوشه‌بندی، بخش‌بندی رنگی، دقت تصویر.

۱. مقدمه

پردازش خودکار تصاویر بیرونی یکی از زمینه‌های مهم بینایی ماشین است که از جمله کاربردهای آن می‌توان به ساخت روبات‌های هوشمند برای ایفای نقش در محیط‌های خارجی [۱] و طراحی کامپیوترهای پوشیدنی [۲] اشاره کرد. بخش‌بندی یکی از مهمترین گامهای ابتدایی پردازش خودکار تصاویر است که موفقیت در تحلیل تصویر تا حد زیادی به آن وابسته است [۳]. بخش‌بندی، تصویر را به اجزای سازنده آن تقسیم می‌کند تا به کمک روندهای تشخیص شیء بتوان این اجزاء را شناسایی نمود. این تشخیص در زمینه‌های زیادی از بینایی ماشین مانند تحلیل صحنه و ردیابی اشیاء کاربرد خواهد داشت.

مشخصات خاص تصاویر بیرونی از جمله تغییرات نوری، وجود جزئیات بافتی زیاد و وجود اشیاء کوچک و زیاد و ناهمگن، باعث می‌شود مسأله بخش‌بندی تصاویر بیرونی با چالش‌های جدی مواجه شود [۴، ۵]. به همین دلیل استفاده از رویه‌های پیش‌پردازش که بتواند ابعاد فضای اطلاعاتی تصویر را کاهش دهد، برای بخش‌بندی این تصاویر معمولاً ضروری است.

رنگ یکی از مهمترین ویژگی‌های هر شکل است که با تکیه بر آن می‌توان تا حد زیادی به موفقیت بخش‌بندی امیدوار بود [۶، ۷]. اما تنوع رنگی تصویر که در سیستم‌های دیجیتال امروزی به صورت معمول چند میلیون رنگ را دربرمی‌گیرد، مانعی جدی در این راه به شمار می‌رود. یکی از ابزارهای معمول پیش‌پردازش برای عملیات بخش‌بندی استفاده از روش‌های کاهش و دسته‌بندی رنگهاست. روندهای کاهش رنگ،

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هستوگرام با دقت دوگانه

پیشنهاد شده در [۲۵-۲۸] اشاره کرد. هرچند با استفاده از این الگوریتم مانند آنچه پیش‌تر در [۲۱] پیشنهاد کردیم، می‌توان با حفظ کیفیت بصری تصویر، پهنای باند لازم برای انتقال و نیز فضای مورد نیاز برای ذخیره‌سازی تصویر را به نحو مطلوبی کاهش داد، اما به لحاظ تنوع رنگ و جزئیات بافتی تصاویر بیرونی، روش k -means استاندارد (مانند دیگر روش‌های خوشه‌بندی رنگی) معمولاً کارایی مناسبی در کاهش رنگ این تصاویر به هدف بخش‌بندی ابتدایی (که در آن شکل کلی اشیاء اهمیت بیشتری دارد) از خود نشان نمی‌دهد و عموماً منجر به بخش‌بندی نادرست می‌شود [۲۹، ۳۰]. به منظور بهبود روش k -means برای بخش‌بندی رنگی تصاویر بیرونی، در [۳۱] روشی مبتنی بر خوشه‌بندی سلسله‌مراتبی با حذف تدریجی خوشه‌ها با استفاده از هرم گوسی با دقت چندگانه پیشنهاد کردیم که در مقایسه با روش k -means عادی، کارایی بهتری در بخش‌بندی تصاویر بیرونی و ایجاد شکل کلی اشیاء از خود نشان می‌داد. در این مقاله، روش فوق به صورت تطبیق‌پذیر با تصویر گسترش داده می‌شود تا به کمک تجمع آماری رنگ‌های تصویر که از روی هستوگرام توزیع رنگ به دست می‌آیند، بتواند نسبت به شرایط نوری و بافتی متنوع تصاویر بیرونی مقاوم باشد. کارایی این الگوریتم نسبت به روش k -means استاندارد و نسخه بهبود یافته آن، به کمک یک روش ارزیابی نظارت‌شده روی دو مجموعه تصویر استاندارد در بستری با دقت دوگانه بررسی خواهد شد.

در بخش ۲ به معرفی روش k -means و عملکرد آن در کاهش رنگ تصاویر بیرونی خواهیم پرداخت. بخش ۳ این تحقیق به تشریح الگوریتم k -means بهبود یافته و نسخه گسترش یافته آن اختصاص دارد. روش و نتایج ارزیابی الگوریتم در بخش ۴، تحلیل نتایج در بخش ۵ و جمع‌بندی نهایی و پیشنهادهایی برای بهبود روش مورد تحقیق در بخش ۶ آورده شده است.

۲. روش k -means برای کاهش رنگ

روش k -means استاندارد، یک الگوریتم خوشه‌بندی است که می‌تواند برای دسته‌بندی رنگ‌های یک تصویر به کار رود. اگر هر پیکسل تصویر رنگی را برداریم شکل از ویژگی‌های رنگی آن پیکسل در نظر بگیریم، روش k -means باید این بردارهای رنگی را از روی شباهتشان به هم به k خوشه تقسیم کند. هرچند ساده‌ترین ویژگی‌های رنگی مورد استفاده، ویژگی‌های فضای RGB یا مانند [۳۲] ویژگی‌های فضای HSV است، اما تحقیقات انجام شده در [۳۳] نشان می‌دهد فضای رنگی CIE-Lab به لحاظ ادراکی بودن، بیشترین بازده را در مسأله خوشه‌بندی رنگی

تصویر را از یک فضای اطلاعاتی با ابعاد چند میلیون رنگ به یک فضای دسته‌بندی شده محدود با ابعاد چند رنگ مهم تبدیل می‌کند. تحلیل ماشینی این فضای محدود کاری ساده‌تر و طبعاً کارتر است که می‌تواند به عنوان گامی ابتدایی در بخش‌بندی تصویر به کار رود [۸].

هدف از کاهش تعداد رنگ‌های یک تصویر، ترکیب رنگ‌های نزدیک و ایجاد تصویری جدید با تعدادی محدود رنگ است که بتواند رنگ‌های تصویری اصلی (الگوها یا اشیاء) را به نمایش درآورد. این فرآیند می‌تواند برای تقسیم تصویر به عناصر اصلی سازنده‌اش مفید باشد.

دسته‌ای از روش‌های کاهش رنگ بر مبنای تقسیم متوالی مکعب سه‌بعدی RGB عمل می‌کنند که از جمله آنها می‌توان به روش‌های Octree [۹]، برش میانه [۱۰]، و الگوریتم‌های مبتنی بر پراش [۱۱] اشاره کرد. دسته‌ی دیگر الگوریتم‌های کاهش رنگ، بر مبنای خوشه‌بندی رنگ‌ها عمل می‌کنند [۱۲]. این الگوریتم‌ها برای پیدا کردن رنگ‌های مهم تصویر از روش‌های خوشه‌بندی مانند کوانتیزاسیون برداری [۱۳]، GNG [۱۴]، FOSART [۱۵]، Fuzzy ART [۱۶]، ACR [۱۷] استفاده می‌کنند. شبکه‌های عصبی نیز برای خوشه‌بندی اطلاعات رنگی مورد استفاده قرار می‌گیرند. به عنوان مثال، شبکه عصبی خودسامانده کوهونن برای کاهش رنگ محبوبیت فراوانی دارد که از جمله پژوهش‌های مرتبط با آن می‌توان از [۱۸-۲۰] نام برد. پیش‌تر، با افزایش تعداد ویژگی‌های رنگی به صورت افزونه و در نظر گرفتن پیکسل‌ها در فضاها رنگی مختلف، در [۲۱] روشی مبتنی بر شبکه عصبی خودسامانده پیشنهاد کردیم که به دسته‌بندی بهتر رنگ‌ها می‌انجامد.

در روش‌های مزبور، پیکسل‌های تصویر مانند یک بردار با مؤلفه‌های ویژگی‌های رنگی (مثلاً درصد مشارکت رنگ‌های اصلی قرمز و سبز و آبی در ساخت آن) در نظر گرفته می‌شوند که باید با هم ترکیب شوند و بردارهایی بسازند که نمایندگان خوبی از رنگ‌های تصویر باشند. به بیان دیگر به‌کاهش رنگ می‌توان به عنوان یک تبدیل از فضای برداری وسیع به فضای برداری محدود نگریست. این فضای برداری محدود همان رنگ‌های مهم تصویر است که دسته‌بندی اشیاء تصویر به کمک آنها می‌تواند قدم مهمی در بخش‌بندی ابتدایی تصویر باشد [۲۰، ۲۲].

یکی از راه‌حل‌های ساده، سریع و کارا برای مسأله خوشه‌بندی استفاده از الگوریتم k -means است [۲۳] که به همراه نسخه فازی آن به نام Fuzzy c-means یا FCM [۲۴]، پرکاربردترین الگوریتم‌های خوشه‌بندی در کاربردهای صنعتی پردازش تصویر و یادگیری ماشین به شمار می‌روند که از جمله پژوهش‌های اخیر می‌توان به الگوریتم‌های

خود قرار گیرند. معیار این صحت، کمینه شدن حاصل جمع خطاها، S در رابطه ۲ است:

$$S = \sum_{i=1}^k \sum_{p \in C_i} d(c_i, p) \quad (2)$$

که در آن $d(c_i, p)$ فاصله اقلیدسی مرکز خوشه i ام تا پیکسل p آن خوشه است.

انتساب بردارهای ورودی برای تصحیح مقادیر بردارهای نماینده حتماً باید تصادفی باشد؛ در غیر این صورت بردارهای نماینده در فضای برداری به خوبی پخش نخواهند شد و مشکل فراموشی پیش خواهد آمد [۳۴].

۴) پس از پایان فرآیند آموزش، بردارهای نماینده در محل صحیح خود قرار گرفته‌اند. اکنون با ارائه پیکسل‌های تصویر به این بردارها و نظیر کردن هر پیکسل به شبیه‌ترین بردار نماینده، می‌توان به خوشه‌بندی صحیحی رسید.

روش فوق به عنوان یکی از متداول‌ترین شیوه‌های خوشه‌بندی و کاهش رنگ تصاویر به کار می‌رود؛ اما همانطور که در بخش ارزیابی نتایج نشان داده خواهد شد، روش k -means استاندارد کارایی مناسبی در کاهش رنگ این تصاویر به هدف بخش‌بندی ابتدایی از خود نشان نمی‌دهد [۲۹] [۳۰]. شکل ۱ این مطلب را به خوبی نشان می‌دهد.



شکل ۱ - تصویر اصلی (راست) و تصویر کاهش رنگ داده شده به کمک الگوریتم k -means (چپ)

به اشتباه با کلاس‌های بزرگتر ترکیب شده‌اند؛ مثلاً برج وسط تصویر با جاده یا ساختمان‌های بالا سمت چپ با سبزه و شاخ و برگ درختان تلفیق شده‌اند.

اهم دلایل عدم کارایی الگوریتم k -means استاندارد برای بخش‌بندی به هدف تشخیص اشیاء در تصاویر بیرونی به شرح زیر است:

دارد. به علاوه چون فاصله اقلیدسی دو رنگ در این فضا متناسب با تفاوت بصری آنهاست، خوشه‌بندی رنگی با تکیه بر فاصله اقلیدسی در این فضا کارا تر می‌شود. به همین لحاظ در این تحقیق از ویژگی‌های رنگی پیکسل‌ها در فضای CIE-Lab استفاده شده است.

الگوریتم k -means برای خوشه‌بندی فوق به شرح زیر است:

۱) در آغاز k پیکسل تصادفی از تصویر به عنوان نمایندگان ابتدایی k خوشه رنگی (در فضای CIE-Lab) انتخاب می‌شوند که باید در ادامه اصلاح شوند.

۲) یکی از پیکسل‌های تصویر به صورت تصادفی انتخاب و به بردارهای نماینده عرضه می‌شود. شباهت بردار ورودی با هر کدام از بردارهای نماینده که بیشتر باشد، بردار ورودی جذب آن شده و در عین حال آن بردار نماینده را به خود شبیه می‌کند. برای این هدف، مقدار جدید بردار نماینده «برنده» برابر میانگین مقدار قبلی و بردار ورودی جذب شده خواهد بود (رابطه ۱). معیار شباهت می‌تواند فاصله اقلیدسی، حاصلضرب داخلی، یا معیارهای دیگر باشد.

۱) (پیکسل جدید، بردار نماینده قدیم) $f =$ بردار نماینده جدید که تابع f در الگوریتم k -means تابع میانگین می‌باشد.

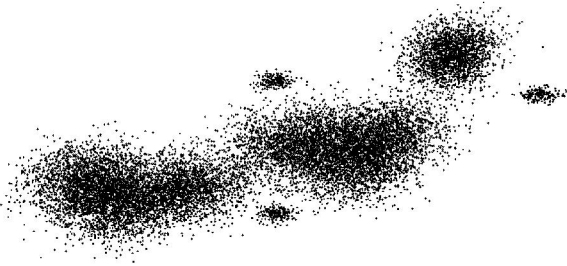
۳) گام ۲ با ارائه بقیه پیکسل‌های تصویر به بردارهای نماینده و تکرار این روند تا جایی ادامه می‌یابد که بردارهای نماینده در جای صحیح

همانطور که مشاهده می‌شود، الگوریتم k -means می‌تواند رنگ‌های تصویر را با حفظ کیفیت بصری ترکیب کند. اما تصویر حاصل از آن برای بخش‌بندی به هدف تشخیص شیء مناسب نیست. مثلاً در شکل بالا دیده می‌شود که آسمان و چمن با وجود اینکه یک شیء واحد هستند، به دلیل وجود جزئیات بافتی و سایه‌های رنگی مختلف، به چند بخش مختلف کم‌اهمیت تقسیم شده‌اند. به علاوه، بعضی اشیاء کوچک

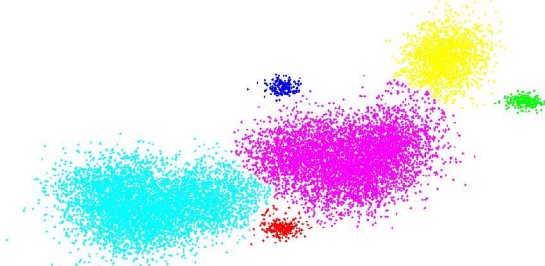
بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه

جزییات بافتی زیاد و نیز اشیاء کوچک و ناهمگن در تصاویر بیرونی باعث ایجاد نویز زیادی می‌شوند که سرعت و نیز دقت همگرایی را کاهش می‌دهد و یک خطای مانا در رابطه (۲) ایجاد می‌کند. شکل ۲ یک نمای دوبعدی از مشکلات الگوریتم k -means در بخش‌بندی را نشان می‌دهد.

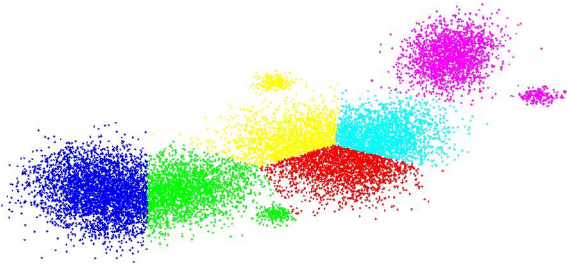
جزییات بافتی، تعداد زیادی بخش کوچک و جدا از هم تولید می‌کند که برای روندهای تشخیص اشیاء مشکل‌ساز می‌باشند (چمن در شکل ۱). تعدادی از اشیاء کوچک در بخش‌های بزرگتر ادغام می‌شوند (برج وسط تصویر در شکل ۱) سایه‌های رنگی متعلق به یک شیء تعدادی بخش بی‌مورد تولید می‌کنند (چمن و آسمان در شکل ۱).



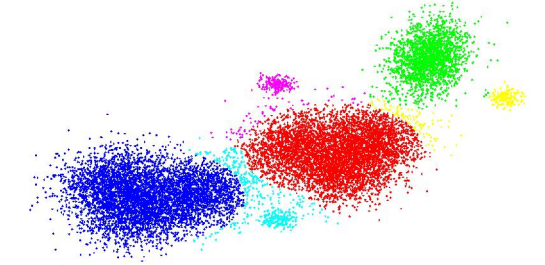
ب) خوشه‌بندی ایده‌آل



الف) یک فضای دوبعدی نوعی



د) خوشه‌بندی به کمک الگوریتم پیشنهادی



ج) خوشه‌بندی به کمک الگوریتم k -means عادی

شکل ۲) کارایی الگوریتم k -means در خوشه‌بندی

۳. بهبود الگوریتم k -means به کمک هرم تصاویر چند

دقتی

حتماً تاکنون صحنه‌ای را که شخصی که تازه به هوش می‌آید در مقابل خود مشاهده می‌کند، دیده‌اید؛ اشیاء ابتدا کاملاً محو هستند و تنها رنگ چند شیء اصلی مشخص است. رفته‌رفته اشیاء مشخص‌تر می‌شوند و حتی اشیاء جدید که قبلاً دیده نمی‌شده‌اند اضافه می‌شوند. این موضوع می‌تواند برای بخش‌بندی ابتدایی تصاویر بیرونی یک ایده مناسب باشد؛ چون چشم انسان نیز در وهله اول برای تشخیص اشیاء مختلف به رنگ آنها بیشتر از جزئیات بافتی توجه می‌کند.

اگر یک فیلتر ملایم‌کننده (مانند فیلتر میانگیر یا گوسی) را پی‌درپی روی یک تصویر اعمال کنیم، تعدادی تصویر محوشده با دقت‌های متفاوت به دست می‌آیند. این تصاویر در بسیاری از کاربردهای پردازش

در شکل ۲-الف تعدادی نقطه را در یک فضای دوبعدی می‌بینید که باید دسته‌بندی شوند. یک خوشه‌بندی ایده‌آل باید این نقاط را به شش دسته (سه دسته بزرگتر و سه دسته کوچکتر) تقسیم کند (شکل ۲-ب). اما تعداد زیاد نقاط در دسته‌های بزرگتر سبب می‌شود تمرکز الگوریتم خوشه‌بندی k -means بر این دسته‌ها بیشتر شود. در نتیجه همانطور که در شکل ۲-ج دیده می‌شود، خوشه‌های بزرگتر به چند دسته کوچکتر تقسیم می‌شوند و خوشه‌های کوچکتر نادیده گرفته می‌شوند. شکل ۲-د خوشه‌بندی به کمک الگوریتمی که در بخش بعد تشریح می‌شود را نشان می‌دهد.

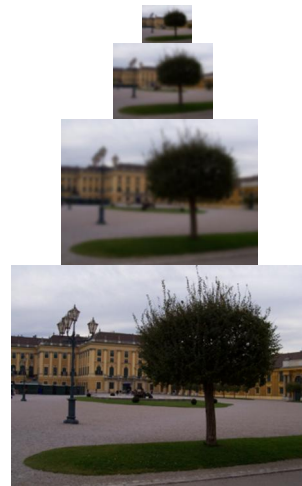
۳.۱. الگوریتم بهبودیافته

در طبقات بالای هرم چنددقتی، جزئیات بافتی و اشیاء کوچک محو شده و رنگ مهمترین ویژگی شیء برای بخش‌بندی خواهد بود. به علاوه سایه‌های مربوط به یک رنگ تا حد زیادی با هم ترکیب می‌شوند. این دو ویژگی یک بستر مناسب برای کاهش رنگ تصاویر بیرونی را فراهم می‌کند. در [۳۱] برای ایجاد تصاویر ۶ رنگ و ۹ رنگ یک روش کاهش رنگ بر اساس الگوریتم *k-means* پیشنهاد کردیم که در هر طبقه هرم تصاویر چنددقتی به کمک خوشه‌بندی رنگی اشیاء مهم را تشخیص داده و در طبقه زیرین آنها را حذف می‌کند تا بخش‌بندی در ادامه فقط روی اشیاء باقی‌مانده متمرکز شود. استفاده از این روش نسبت به خوشه‌بندی کامل و یک‌مرحله‌ای تصویر در طبقات مختلف هرم فوق‌کاری بهتری از خود نشان داده است؛ چون از تأثیر خوشه‌های بزرگ در خوشه‌بندی رنگی در طبقات پایین هرم جلوگیری می‌کند.

مراحل این روش (که از این پس الگوریتم حذف تدریجی نامیده می‌شود) برای ایجاد یک تصویر ۶ رنگ به کمک دو طبقه از هرم چنددقتی به شرح ذیل می‌باشد:

(۱) یک فیلتر ملایم‌کننده 3×3 میانگیرا ۵ بار و ۱۰ بار روی تصویر اصلی اعمال می‌کنیم تا دو نسخه محو شده از تصویر (که به ترتیب B1 و B2 نامیده می‌شوند) به دست آیند. شکل ۴ این تصاویر محو شده را نشان می‌دهد.

تصویر از جمله بخش‌بندی و دسته‌بندی تصاویر مفید هستند [۳۵-۳۷]. چون با هربار اعمال فیلتر ملایم‌کننده روی تصویر، حجم اطلاعات کمتر می‌شود، می‌توان با کاهش ترکیبی نمونه‌ها اندازه تصویر را نیز کاهش داد. بنابراین مجموعه تصاویر چنددقتی می‌توانند یک هرم تشکیل دهند که تصویر اصلی در قاعده آن و تصاویر محوتر (و البته کوچکتر) در طبقات بالاتر قرار می‌گیرند. شکل ۳ یک هرم تصاویر چنددقتی را نشان می‌دهد. اولین کاربرد هرم تصاویر چنددقتی در بخش‌بندی تصاویر به کمک ایجاد پیوندهایی بین پیکسل‌های نظیر در طبقات مختلف هرم در [۳۵] مطرح شد. در سال‌های اخیر بخش‌بندی تصاویر با استفاده از ویژگی‌هایی که از طبقات مختلف این هرم استخراج می‌شوند، موضوع تحقیقات زیادی بوده است که نمونه‌هایی از آنها در [۳۶] فهرست شده‌اند. همچنین در [۳۷] روشی برای خوشه‌بندی رنگی بر اساس تحلیل قله‌های هیستوگرام در طبقات مختلف هرم فوق مطرح شده است.



شکل ۳) هرم تصاویر چنددقتی برای یک تصویر بیرونی



شکل ۴- الف) تصویر اصلی ب) تصویر B1 ج) تصویر B2

(۲) به کمک روش *k-means* تصویر B2 (محوترین نسخه تصویر) را به سه خوشه رنگی تقسیم می‌کنیم. شکل ۵-الف یک تصویر بیرونی و شکل ۵-ب، نسخه سه‌رنگی آن را نشان می‌دهد. این سه خوشه از رنگهای بسیار مهم تصویر با چگالی پیکسلی بالا هستند. نکته مهمی که

این کار باعث می‌شود تا نقاط مزاحم، اشیاء بسیار کوچک و به ویژه جزئیات بافتی که در تصاویر بیرونی بسیار دیده می‌شود و کار بخش‌بندی را با مشکل مواجه می‌کند از بین بروند و با ترکیب سایه‌های رنگی مشابه، تنها دسته‌های رنگی مهم باقی بمانند.

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هستوگرام با دقت دوگانه

هیچکدام از خوشه‌های رنگی جا نمی‌گیرند و باید مجدداً خوشه‌بندی شوند. روش تعیین حد آستانه مزبور را در ادامه خواهیم دید. این نقاط را (که نقاط یتیم نامیده می‌شوند) در شکل ۵-۵ می‌بینید (قسمت سفیدرنگ این تصویر حاوی پیکسل‌های غیریتیم است که در خوشه‌بندی مرحله اول نمایندگان مناسب خود را یافته‌اند).

۴) اکنون به روشی مشابه گام ۲، نقاط یتیم را به سه خوشه جدید تقسیم می‌کنیم تا مجموعاً شش خوشه رنگی در تصویر به دست آوریم. تصویر ۶ کلاسه که به این روش به دست آمده (شکل ۵-۶) برای بخش‌بندی مناسب است؛ به ویژه وقتی آن را با زمانی مقایسه کنیم که از ابتدا تصویر را به ۶ کلاس تقسیم کرده باشیم (شکل ۵-۷). برای مقایسه بهتر از یک جعبه رنگ متفاوت و درخشان استفاده شده است.

باید در اینجا مورد توجه قرار بگیرد این است که تعدادی از پیکسل‌های تصویر واقعاً به این سه کلاس شباهت دارند و در خوشه رنگی آنها قرار می‌گیرند؛ اما نقاطی دیگر هم در تصویر وجود دارند که مربوط به اشیاء دیگر هستند و رنگهای آنها شباهتی به این سه خوشه رنگی ندارد؛ اما به ناچار در این خوشه‌ها قرار گرفته‌اند. در گام بعدی این نقاط به صورت مجزا خوشه‌بندی خواهند شد. مقایسه شکل‌های ۵-ب و ۵-ج نشان می‌دهد حتی این کاهش رنگ سه‌خوشه‌ای به کمک تصاویر محوشده، به خاطر حذف جزئیات بافتی از دید بخش‌بندی بهتر از کاهش رنگ در فضای معمولی عمل می‌کند (هرچند کیفیت بصری پایین‌تری دارد).

۳) اکنون فاصله اقلیدسی نقاط تصویر B1 (نسخه واضح‌تر تصویر) را تا سه نماینده به دست آمده از مرحله قبل محاسبه می‌کنیم. نقاطی که فاصله اقلیدسی‌شان از هر سه نماینده بیش از حد آستانه $th1$ باشد، در



ج



ب



الف



و



هـ



د

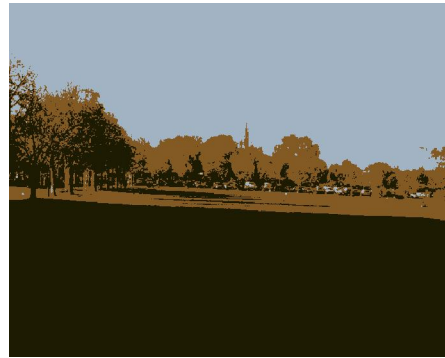
شکل ۵) الف) تصویر اصلی ب) تصویر B2 سه‌رنگی شده ج) تصویر اصلی سه‌رنگی شده د) نقاطی که در ۳ خوشه رنگی جا نگرفته‌اند ه) تصویر شش رنگ به کمک k -means در هر تصاویر چنددقتی و) تصویر شش رنگ به کمک k -means عادی

رنگی مهم که از طبقه بالاتر به دست آمده‌اند، از نسخه واضح‌تر تصویر برای یافتن اشیاء جدید استفاده می‌کند و با جلوگیری از تمرکز خوشه‌های رنگی در رنگهای مهم با چگالی پیکسلی بالا، به خوشه‌های رنگی کوچکتر (اشیاء کوچکتر) میدان بروز می‌دهد (شکل ۲ را مجدداً ببینید).

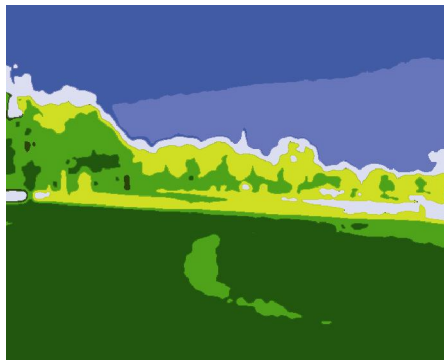
این کیفیت بهتر ناشی از این است که در مرحله میانی، سه خوشه‌ای که از مرحله اول به دست آمده‌اند حذف می‌شوند و در مرحله آخر سه خوشه جدید فقط با تمرکز بر بقیه تصویر به دست می‌آیند؛ در حالی که وقتی از ابتدا تصویر را به ۶ خوشه تقسیم کنیم، نقش خوشه‌هایی با چگالی بالا پررنگ‌تر می‌شود. به بیان دیگر، این روش با حذف خوشه‌های



ب



الف



د



ج

شکل ۶- الف) تصویر اصلی (ب) تصویر ۳ رنگ شده (ج) تصویر ۶ رنگ شده با روش حذف تدریجی (د) تصویر ۶ رنگ شده با k -means

۱) به کمک الگوریتم تقسیم تطبیقی مکعب RGB تعدادی تجمع پیکسلی در فضای رنگی می‌یابیم؛ به گونه‌ای که این تعداد، از تعداد رنگی که معمولاً در یک تصویر بیرونی یافته می‌شود بیشتر باشد.

۲) یک گراف وزن‌دار کاملاً متصل که هر گره آن، ویژگی رنگی میانگین در یک خوشه در فضای رنگی CIE-Lab و وزن هر یال آن فاصله اقلیدسی بین رنگ دو گره و سر آن (که در فضای مزبور متناظر با تفاوت بصری آن‌ها است) باشد، ایجاد می‌کنیم.

۳) یالی که کمترین وزن را دارد، کمترین فاصله اقلیدسی (معادل با کمترین تفاوت بصری) موجود بین دو رنگ را نشان می‌دهد. اگر وزن این یال از یک حد آستانه کمتر باشد، نشان می‌دهد که گره‌های (رنگ‌های) دو سر این یال به دلیل تشابه رنگی زیاد نامزد ترکیب شدن با یکدیگر هستند. بنابراین دو رنگ مزبور را با هم ترکیب کرده و یک رنگ ایجاد می‌کنیم؛ به علاوه در ساختار گراف، دو گره معادل این دو رنگ را ترکیب کرده و رنگ میانگین را به عنوان ویژگی رنگی گره جدید ثبت کرده و وزن یال‌هایی که به یکی از دو گره قبلی مرتبط بوده‌اند را به‌روز می‌کنیم. به کمک ایده هرس گراف که در [۴۱] نمونه‌ای از آن را مطرح

شکل ۶ اثر اجرای این الگوریتم روی تصویری دیگر را نشان می‌دهد.

۳.۲. تعیین تطبیقی تعداد خوشه‌ها

همان‌طور که پیش‌تر اشاره شد، هدف از به‌کارگیری خوشه‌بندی در این تحقیق، ایجاد پیش‌الگوهای مناسب برای بخش‌بندی و تشخیص اشیاء در تصاویر بیرونی است. یکی از مسایل مهم در خوشه‌بندی رنگی، تشخیص تعداد صحیح رنگ‌های تصویر است [۲۰، ۳۸]. ایجاد پیش‌الگوهای زیاد (در نظر گرفتن تعداد زیاد رنگ نهایی در فرآیند کاهش رنگ مانند آنچه در [۳۹] پیشنهاد شده است) می‌تواند باعث سنگین شدن فرآیند بخش‌بندی و نهایتاً منجر به فرابخش‌بندی شود. از سوی دیگر تعداد کم پیش‌الگوها هم باعث از دست رفتن تعدادی از بخش‌ها و بخش‌بندی نادرست خواهد شد [۴۰]. در این تحقیق، مطابق با آنچه در [۴۱] پیشنهاد کردیم، برای تخمین تعداد خوشه‌های رنگی مهم در هر مرحله خوشه‌بندی به صورت زیر عمل می‌کنیم:

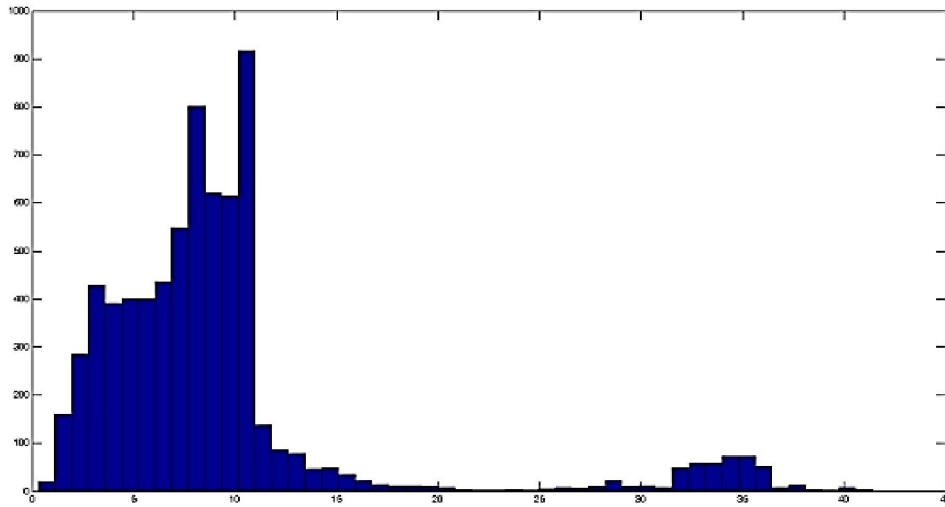
بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه

الگوریتم خوشه‌بندی با حذف تدریجی داده می‌شوند تا خوشه‌های مناسب پیش‌الگوهای بخش‌بندی را ایجاد کند.

۳.۳. تعیین سطوح آستانه شباهت

همانطور که پیشتر گفته شد، در روش حذف تدریجی در مرحله اول تصویر به سه رنگ اصلی خوشه‌بندی می‌شود و در مرحله دوم، نقاطی که به نمایندگان سه رنگ اصلی شبیه نیستند (پیکسل‌های یتیم)، مجدداً خوشه‌بندی می‌شوند. معیار خوشه‌بندی مجدد یک پیکسل این است که فاصله اقلیدسی آن پیکسل تا نماینده دسته‌اش از سطح آستانه th_1 بیشتر باشد. شکل ۷ نمودار هیستوگرام تعداد پیکسل‌ها برحسب فاصله تا نماینده دسته در یک خوشه نوعی را نشان می‌دهد.

کردیم، این کار را تا جایی ادامه می‌دهیم که کمترین وزن یال موجود در گراف (کمترین تفاوت بصری بین دو خوشه رنگی) از یک حد آستانه بیشتر شود (یعنی دو رنگ خیلی شبیه به هم نداشته باشیم). تعداد گره‌های باقیمانده نشان‌دهنده تعداد خوشه‌های رنگی مهم تصویر هستند. حد آستانه را درصدی از فاصله اقلیدسی بیشینه در فضای رنگی CIE-Lab در نظر می‌گیریم. هرچه این درصد بالاتر باشد، به رنگ‌ها بیشتر اجازه ترکیب شدن داده می‌شود؛ یعنی تعداد خوشه‌های رنگی در هر مرحله کمتر می‌شود و پیش‌الگوهای کلی‌تری خواهیم داشت. برای دستیابی به بخش‌بندی جزئی‌تر، باید این درصد کمتر شود. در این تحقیق بهترین پاسخ‌ها با ۵٪ دیده شده است. این تعداد خوشه‌ها به



شکل ۷- نمودار تعداد پیکسل‌ها برحسب فاصله اقلیدسی تا نماینده دسته در یک خوشه نوعی

۳.۳.۱. سطح آستانه ثابت

می‌توان سطح آستانه شباهت هر خوشه را مانند روشی که در [۳۱] پیشنهاد داده‌ایم، یک مقدار ثابت (مثلاً ۹۰٪ بیشینه فاصله اقلیدسی بین پیکسل‌های آن خوشه با مرکز آن) در نظر گرفت؛ رابطه ۳ جزئیات این روش را نشان می‌دهد:

$$th_i = \text{Max}(d_i) \times 90\% \quad (3)$$

th_i سطح آستانه خوشه i ام و d_i مجموعه فاصله اقلیدسی پیکسل‌های خوشه i ام با مرکز آن است.

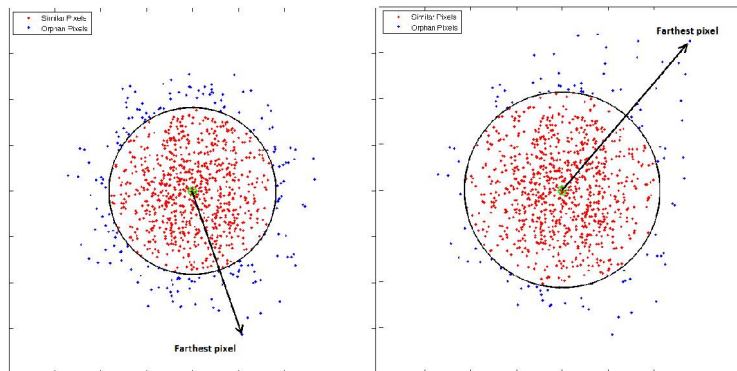
این روش در بسیاری از موارد بهبود خوبی نسبت به روش k -means عادی از نظر دقت بخش‌بندی ایجاد می‌کند؛ اما در بعضی از موارد نیز موفق عمل نمی‌کند؛ دلیل این عدم موفقیت را می‌توان در عدم

همانطور که در این شکل می‌بینید، نمودار دو قله‌ای است؛ قله اول مربوط به تجمع پیکسل‌هایی است که به نماینده دسته شبیه هستند و قله دوم، پیکسل‌هایی را نشان می‌دهد که به نماینده خوشه شبیه نیستند (احتمالاً مربوط به یک شیء جدید). بدیهی است آستانه th_1 باید جایی بین این دو قله باشد.

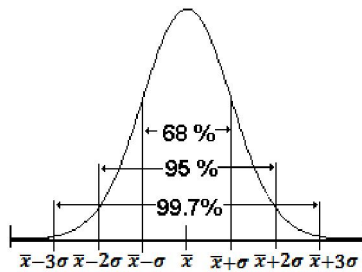
تعیین صحیح حد آستانه شباهت، یک عامل تعیین‌کننده در کیفیت الگوریتم فوق است. به همین لحاظ روشی باید تدوین شود که این حد آستانه را از روی مشخصات آماری تصویر به صورتی تعیین کند که در مقابل تغییرات ویژگی‌های خاص تصاویر بیرونی از جمله تغییرات بافتی و رنگی و نوری مقاوم باشد.

که دور از مرکز یک خوشه قرار گرفته باشند (مثلاً در اثر نویز)، باعث زیاد شدن سطح آستانه می‌شوند. در شکل ۸ دو خوشه نوعی دوبعدی با سطح آستانه ثابت ۵۰٪ با هم مقایسه شده‌اند:

توجه کافی رابطه ۳ به توزیع رنگهای درون یک خوشه جستجو کرد. به بیان دیگر تعیین یک درصد ثابت برای تمام خوشه‌ها در تمام تصاویر (به ویژه در مورد تصاویر بیرونی که دارای تنوع رنگ و نور هستند) نمی‌تواند انتخاب مناسبی باشد. به عنوان مثال، حتی چند پیکسل محدود



شکل ۸- مقایسه حد آستانه شباهت دو خوشه تقریباً مشابه



شکل ۹- نمودار تجمع داده‌ها در یک توزیع نرمال

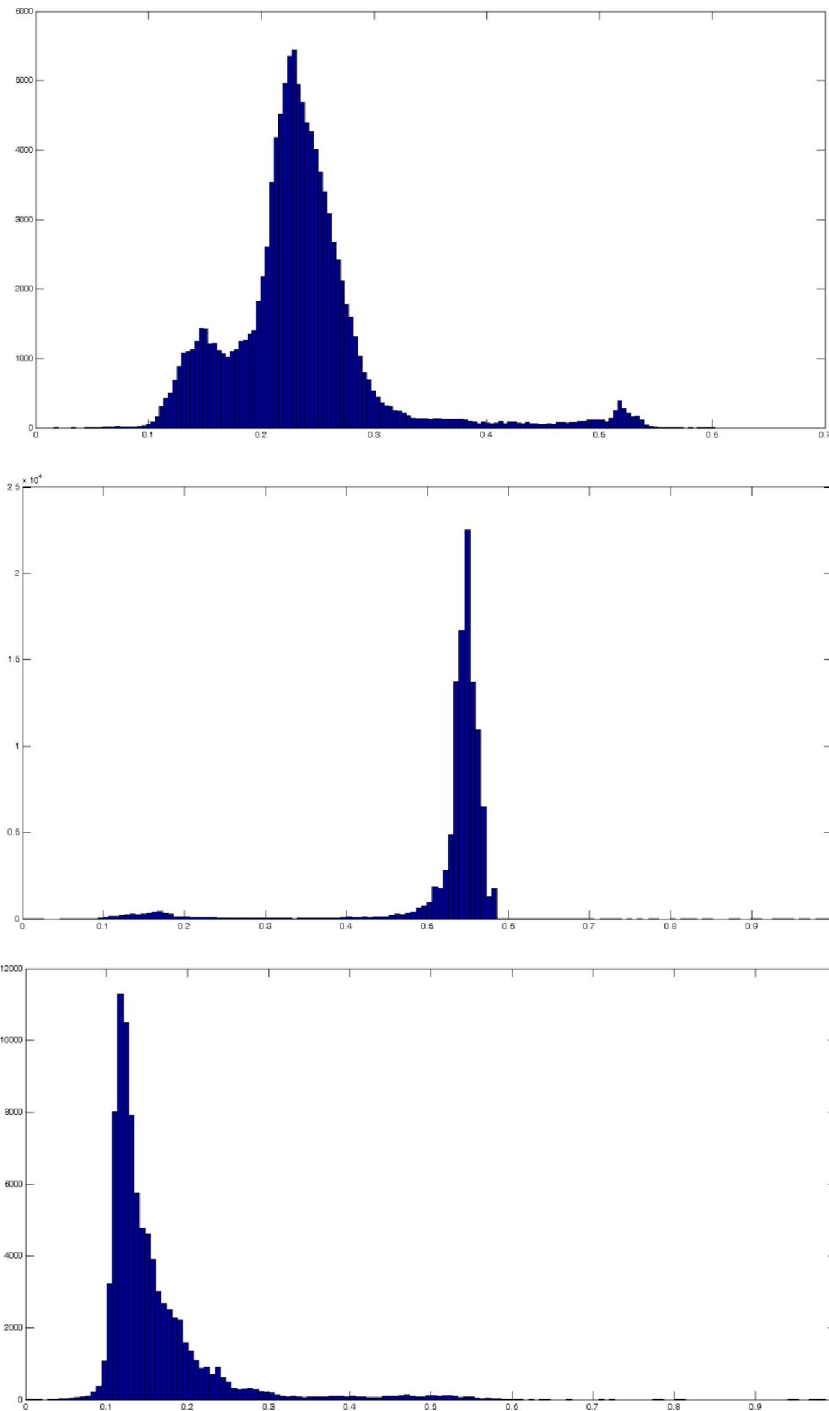
در شکل بالا دو خوشه با توزیع تقریباً مشابه را می‌بینید که حد آستانه شباهت آنها (که با یک دایره نشان داده شده است) اصولاً باید شبیه هم باشند؛ اما چند پیکسل دور از مرکز در خوشه سمت چپ باعث افزایش قابل ملاحظه سطح آستانه شباهت این خوشه شده است. هرچند احتمالاً می‌توان برای یک مجموعه ثابت از تصاویر یک سطح آستانه شباهت مناسب از راه سعی و خطا به دست آورد؛ اما این حد آستانه بهمجموعه دیگری از تصاویر قابل تعمیم نیست. به بیان دیگر الگوریتم فوق در برابر تغییرات شرایط نوری و بافتی مقاوم نیست؛ به همین لحاظ بهتر است حد آستانه شباهت هر خوشه با توجه به توزیع آماری رنگ‌ها در آن خوشه تعیین شود.

۳.۳.۲. تعیین سطح آستانه به کمک هیستوگرام رنگ

همانطور که گفته شد، الگوریتم خوشه‌بندی در مرحله اول تعدادی از رنگهای تصویر را در یک دسته قرار می‌دهد. تعدادی از رنگهای هم‌دسته واقعاً به هم شبیه هستند (احتمالاً مربوط به یک شیء)؛ اما در هر دسته چند رنگ (چند شیء) وجود دارند که شباهتی به رنگهای دیگر آن دسته ندارند و به ناچار در آن دسته قرار گرفته‌اند. این رنگ‌ها باید از این دسته تفکیک و مجدداً خوشه‌بندی شوند تا بتوانند اشیاء دیگر تصویر را مشخص کنند. مطابق آنچه در [۴۲] پیشنهاد شده است، می‌توان توزیع رنگ‌های داخل یک خوشه را (مانند دیگر پدیده‌های طبیعی) «توزیع نرمال» تجمعی حول رنگ میانگین آن دسته فرض کرد. شکل ۹ نمودار یک توزیع نرمال را نشان می‌دهد.

همانطور که در این نمودار دیده می‌شود، حدود ۶۸٪ انرژی توزیع نرمال در فاصله «انحراف معیار ± میانگین» و حدود ۹۵٪ انرژی آن در فاصله «انحراف معیار $2 \times$ میانگین» قرار دارد. بنابراین برای به دست آوردن پیکسل‌های یتیم می‌توان حد آستانه شباهت را مثلاً در فاصله اقلیدسی 2σ (دو برابر انحراف معیار) از میانگین خوشه در نظر گرفت. البته برای مسأله مورد نظر ما در این تحقیق، «میانگین خوشه رنگی» پارامتر برآورد تجمع رنگها در یک خوشه نیست؛ چون می‌تواند توسط رنگ‌های دور از مرکز دسته تحت تأثیر سوء قرار گیرد. شاخص تجمع در این مسأله «تکرار رنگها» و به عبارت دیگر «مد» توزیع آماری رنگ‌ها در یک خوشه است؛ به بیان دیگر رنگ‌های که دور از رنگ‌های پرتکرار خوشه‌شان قرار می‌گیرند، باید به عنوان «یتیم» در نظر گرفته شده و مجدداً خوشه‌بندی شوند. بنابراین هرچند در مورد توزیع نرمال میانگین و مد با هم برابرند، اما در ادامه شاخص مد را به عنوان معیار تجمع در نظر می‌گیریم.

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه



شکل ۱۰-هیستوگرام ته‌رنگ (Hue) در سه خوشه رنگی مربوط به شکل ۵

می‌دهد. این هیستوگرام‌ها حالت حلقوی دارند؛ به عبارت دیگر پیکسل‌هایی که در انتهای نمودار (با مقدار ۱) قرار دارند، با پیکسل‌های ابتدای نمودار (با مقدار صفر) هم‌رنگ می‌باشند. شکل ۱۱ این موضوع را نشان می‌دهد.

توزیع رنگ‌های یک تصویر بیرونی در خوشه‌های رنگی، کمی با توزیع نرمال متفاوت است (هرچند این توزیع‌ها را نمی‌توان به توزیعی بهتر از نرمال نظیر کرد). شکل ۱۰ نمونه توزیع هیستوگرام ته‌رنگ (پارامتر H در فضای رنگی HSV [۳۲]) در سه خوشه شکل ۵ را نشان

شکل ۱۲ رنگ‌های مزبور (داخل مستطیل قرمز رنگ) در مورد سه خوشه رنگی را نشان می‌دهد. پیکسل‌هایی به این رنگ‌ها یتیم هستند و مجدداً خوشه‌بندی می‌شوند تا رنگ‌های جدیدی (مربوط به اشیاء جدید) را معرفی کنند. توجه کنید که در خوشه اول شکل ۱۲ شرط $L < U$ و در خوشه‌های دوم و سوم شرط $U < L$ صادق است.

۴. ارزیابی روشها

همانطور که در شکل‌های ۵ و ۶ دیدیم، روش حذف تدریجی از نظر بصری نسبت به روش عادی نتایج مناسبتری برای بخش‌بندی تولید می‌کند. اما ارزیابی بصری نمی‌تواند ملاکی مناسب برای ارزیابی کیفیت یک روش بخش‌بندی باشد. به همین لحاظ، برای ارزیابی کارایی الگوریتم بهتر است از یک روش کمی نیز استفاده شود. روش‌های کمی ارزیابی الگوریتم‌های بخش‌بندی به سه دسته عمده تقسیم می‌شوند [۴۳]:

الف) روش‌های تحلیلی که بر ویژگی‌های یک الگوریتم بخش‌بندی مانند پیچیدگی و بهره‌وری زمانی و نوع پردازش، فارغ از خروجی آن تمرکز می‌کنند. این روش‌ها برای مقایسه کارایی الگوریتم‌های بخش‌بندی در تشخیص شیء چندان مناسب نیستند.

ب) روش‌های تجربی نظارت شده که کیفیت یک الگوریتم بخش‌بندی را بر اساس اختلاف خروجی آن با بخش‌بندی دستی توسط ناظر می‌سنجند.

ج) روش‌های تجربی نظارت نشده که بر اساس پارامترهای دید انسانی نسبت به کیفیت یک الگوریتم بخش‌بندی اظهار نظر می‌کنند.

بهترین و معمول‌ترین شیوه ارزیابی کمی الگوریتم‌های بخش‌بندی که در این تحقیق نیز از آن استفاده شده است، روش نظارت شده می‌باشد که بیشترین دقت در ارزیابی الگوریتم‌ها را داراست [۴۳]. بنابراین باید تعدادی تصویر بیرونی را توسط ناظر انسانی به ۶ رنگ بخش‌بندی کنیم و سپس تصاویر خوشه‌بندی شده به کمک روش k -means استاندارد و روش‌های حذف تدریجی پیشنهاد شده را با تصاویر مرجع مقایسه کنیم. هر روشی که تصویر حاصل از آن اختلاف کمتری در برچسب پیکسل‌ها با تصویر مرجع داشته باشد روش بهتری است.



شکل ۱۱- نمودار تغییرات تهرنگ (Hue)

نمودارهای شکل ۱۰ نشان می‌دهند که توزیع رنگ‌ها در یک خوشه رنگی حالت شبه‌نرمال داشته و نیز در خوشه‌های مختلف شکلی متفاوت دارد؛ بنابراین فاصله اقلیدسی 2σ از «مد» رنگ‌ها در همه تصاویر انتخاب مناسبی نیست. به علاوه گاهی با توزیع‌های نرمال چندقله‌ای مواجه هستیم که تکیه بر مشخصات توزیع نرمال تک‌قله‌ای را ناموجه می‌سازد (هیستوگرام خوشه اول در شکل ۱۰ را ببینید).

برای رفع این مشکل، در این تحقیق از شیوه‌ای دیگر برای تعیین رنگ‌های یتیمدر یک خوشه استفاده می‌کنیم. مراحل این روش به شرح ذیل است:

بلندترین قله هیستوگرام رنگ یک خوشه را انتخاب می‌کنیم (μ). این رنگ، پرتکرارترین رنگ در این خوشه و به عبارتی «مد» این توزیع آماری است.

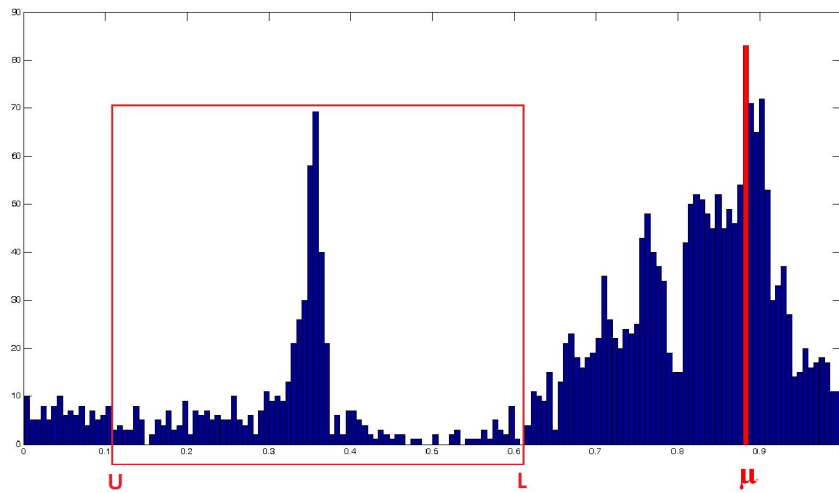
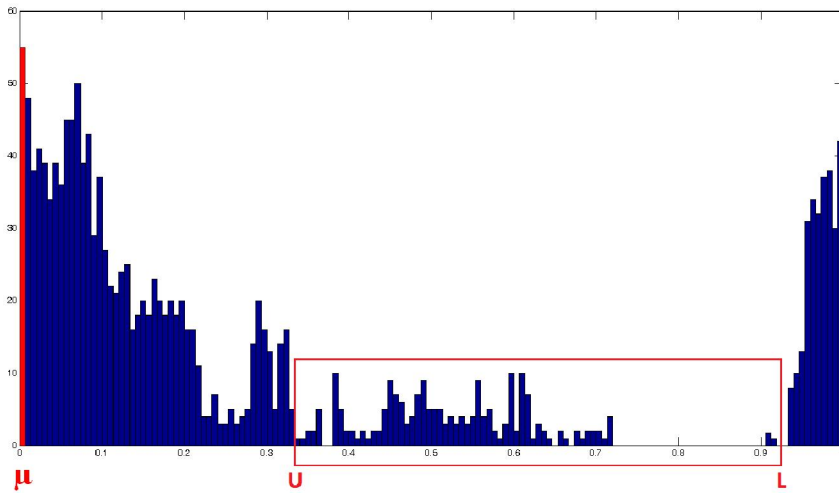
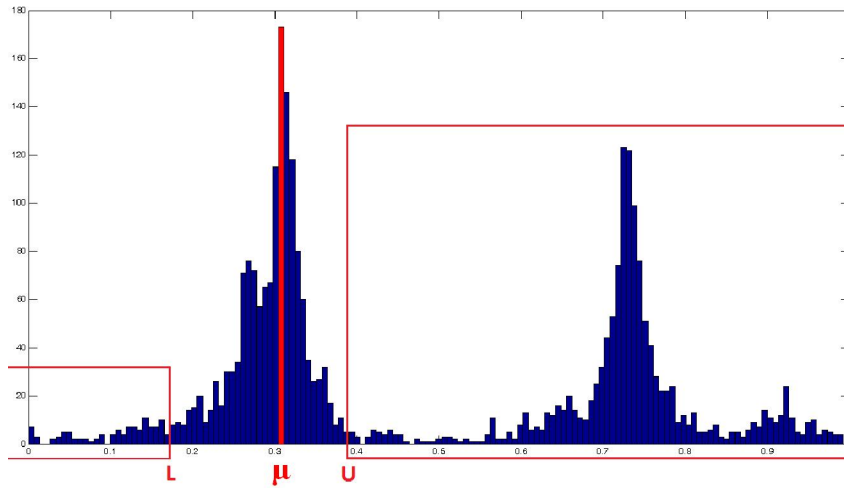
از سمت راست μ به سمت انتهای هیستوگرام حرکت می‌کنیم تا به رنگی برسیم که تعداد پیکسل‌های آن کمتر از ۵٪ پیکسل‌های به رنگ μ باشد. این رنگ را U می‌نامیم. اگر در این حرکت به انتهای هیستوگرام برخوردیم و هنوز به رنگ U نرسیده باشیم، به دلیل حلقوی بودن هیستوگرام تهرنگ «چرخ زده» و جستجو را از ابتدای هیستوگرام پیگیری می‌کنیم.

از سمت چپ μ به سمت ابتدای هیستوگرام حرکت می‌کنیم تا به رنگی برسیم که تعداد پیکسل‌های آن کمتر از ۵٪ پیکسل‌های به رنگ μ باشد. این رنگ را L می‌نامیم. اگر در این حرکت به ابتدای هیستوگرام برخوردیم و هنوز به رنگ L نرسیده باشیم، به دلیل حلقوی بودن هیستوگرام تهرنگ «چرخ زده» و جستجو را از انتهای هیستوگرام پیگیری می‌کنیم.

به کمک رابطه ۴، رنگ‌هایی که پیکسل‌های آن رنگ یتیم هستند را تعیین می‌کنیم:

$$\text{orphan colors} = \begin{cases} (\text{colors below } L) \cup (\text{colors above } U) & \text{if } L < U \\ (\text{colors below } L) \cap (\text{colors above } U) & \text{if } U < L \end{cases} \quad (4)$$

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه



شکل ۱۲-رنگ‌های یتیم در سه خوشه متفاوت (داخل کادر قرمز رنگ)

نرمالیزه شدن تهیه شده است. نقطه ضعف روش نظارت شده، امکان عدم دقت ناظر یا اعمال سلیقه شخصی وی هنگام بخش‌بندی است؛ برای به حداقل رساندن این مشکل، تنها رنگ‌های اصلی توسط ناظر انتخاب می‌شود و یافتن رنگ‌های مشابه رنگ اصلی در تصویر (پیکسل‌هایی که در یک خوشه رنگی قرار می‌گیرند) توسط نرم‌افزار فوتوشاپ و بدون هیچگونه اعمال نظر از سوی ناظر انجام می‌پذیرد. شکل ۱۳ چند نمونه از تصاویری که به این روش بخش‌بندی شده‌اند را نشان می‌دهد.

برای این کار، یک پایگاه داده به نام UIDS توسط «گروه پژوهشی پردازش کاربردی تصویر و سیگنال دانشگاه اصفهان» فراهم آورده شده است که شامل ۵۰ تصویر بیرونی با اندازه‌های 640×640 و 640×480 می‌باشد و توسط یک ناظر، بر اساس شباهت‌های رنگی بین پیکسل‌ها بخش‌بندی شده تا تصاویر مرجع ۶ رنگ (که به اصطلاح استاندارد طلایی یا Ground Truth نامیده می‌شوند) ایجاد شوند. تصاویر این پایگاه داده از محیط‌های طبیعی و در شرایط نوری، رنگی و بافتی متنوع و بدون



شکل ۱۳- تصاویر مرجع ایجاد شده به کمک ناظر انسانی و نرم‌افزار فوتوشاپ

اکنون باید تصاویر ایجاد شده توسط روش k -means استاندارد و دو روش حذف تدریجی (که در ۱-۳-۳ و ۲-۳-۳ پیشنهاد شده‌اند) را با تصویر مرجع مقایسه کنیم تا مشخص شود کدام روش کیفیت بهتری دارد.

در شکل ۱۴ بخش‌های ایجاد شده در تصویر مرجع، تصویر ایجاد شده به کمک روش عادی و تصویر ایجاد شده به کمک روش حذف تدریجی را می‌بینید. بخش‌های مشابه به روشی خودکار و با توجه به بیشینه شباهت بین دو بخش تعیین می‌شوند.

برای ارزیابی هر الگوریتم مطابق با روشی که در [۴۴] پیشنهاد شده به شیوه زیر عمل می‌کنیم:

(الف) تفاضل بین بخش‌های به دست آمده از آن الگوریتم و بخش‌های به دست آمده از تصویر مرجع (تعداد پیکسل‌های مورد اختلاف بین دو بخش مناظر) را می‌یابیم.

(ب) تفاضل‌های فوق (که تعداد آنها بیانگر ضعف الگوریتم در بخش‌بندی صحیح است) را به ازای تمام بخش‌ها با هم جمع می‌کنیم.

جدول ۱ درصد متوسط خطا در پایگاه داده تهیه شده را در مورد روش k -means عادی و دو نسخه روش حذف تدریجی را نشان می‌دهد.

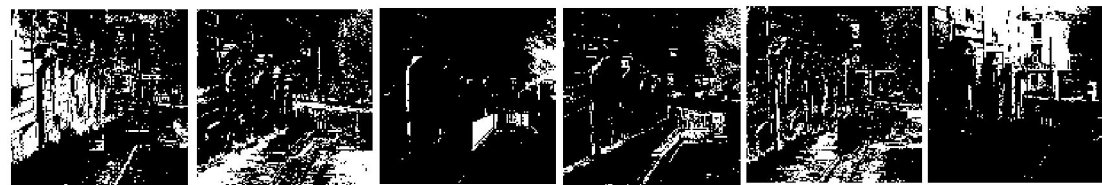
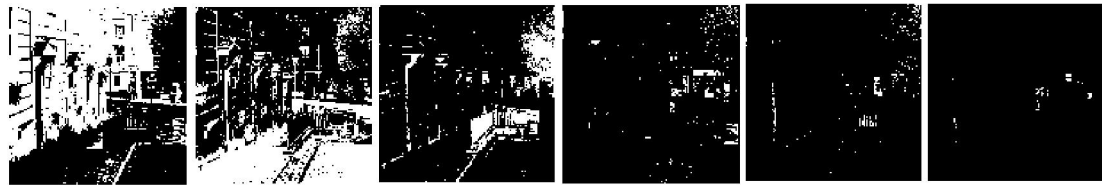
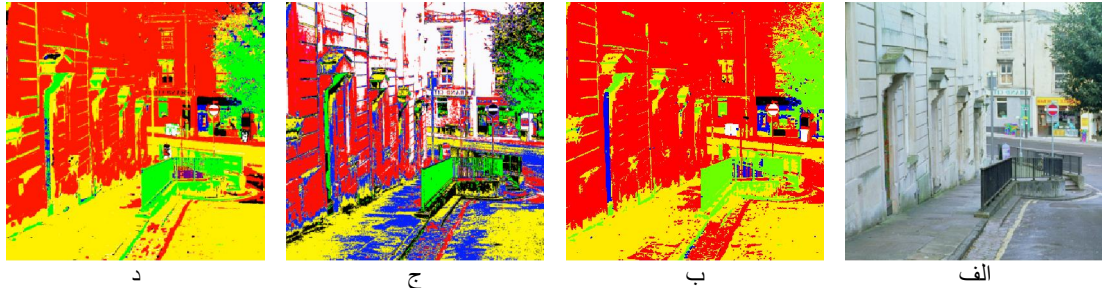
$$ASE = \frac{\sum_{i=1}^K \frac{\sum_{j=1}^M \xi_j^i}{N_i}}{K} \quad (5)$$

در این رابطه K تعداد تصاویر در پایگاه داده مورد بررسی (در اینجا ۵۳)، M تعداد بخش‌های تصویر مرجع (در اینجا ۶)، N_i تعداد پیکسل‌های تصویر i ام و ξ_j^i تعداد پیکسل‌های به اشتباه بخش‌بندی شده در بخش i ام تصویر i ام است.

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه

جدول ۱- مقایسه کیفیت روش‌های خوشه‌بندی برای بخش‌بندی تصاویر به ۶ رنگ

| الگوریتم | kmeans عادی | روش حذف تدریجی ۱ (بخش ۳-۳-۱) | روش حذف تدریجی ۲ (بخش ۳-۳-۲) |
|----------------------|-------------|------------------------------|------------------------------|
| درصد متوسط خطا (ASE) | ٪۴۰/۲۱ | ٪۳۶/۸۲ | ٪۲۹/۱۶ |



شکل ۱۴- الف) تصویر اصلی ب) تصویر مرجع ج) تصویر ۶ رنگ شده به کمک روش k -means عادی د) تصویر ۶ رنگ شده به کمک روش حذف تدریجی ه) بخش‌های به دست آمده از تصویر مرجع و) بخش‌های به دست آمده از تصویر ۶ رنگ شده به روش حذف تدریجی ز) بخش‌های به دست آمده از تصویر ۶ رنگ شده به روش k -means عادی

ارزیابی تحقیقات دیگری از جمله [۴۶] نیز استفاده شده است. جدول ۲ کیفیت روش‌های مورد بحث در این مقاله را در مورد این پایگاه داده نشان می‌دهد. برای آزمایش روش پیشنهادی در یک بستر با دقت دوگانه روی تصویری از این پایگاه داده که به N بخش تقسیم شده باشد، ابتدا تصویر به سه خوشه و سپس پیکسل‌های یتیم به $N-3$ خوشه تقسیم می‌شوند.

مقادیر ذکر شده در جدول ۱ نشان‌دهنده کاهش قابل ملاحظه درصد متوسط خطا به کمک روش حذف تدریجی است. علاوه بر پایگاه داده ذکر شده، روش مورد بررسی این تحقیق به کمک پایگاه داده استاندارد Sowerby Image Data-base (SID) که در [۴۵] معرفی شده نیز مورد ارزیابی قرار گرفت. این مجموعه شامل ۱۰۴ تصویر بیرونی است که توسط ناظر انسانی بر اساس اشیاء موجود در تصاویر به ۵ تا ۷ بخش تقسیم شده‌اند. از این پایگاه داده برای

جدول ۲- مقایسه کیفیت روش‌های خوشه‌بندی برای بخش‌بندی تصاویر به کمک پایگاه داده SID

| الگوریتم | kmeans عادی | روش حذف تدریجی (بخش ۱-۲-۳) | روش حذف تدریجی (بخش ۲-۳) |
|----------------------|-------------|----------------------------|--------------------------|
| درصد متوسط خطا (ASE) | ٪۴۶/۱۲ | ٪۴۴/۸ | ٪۲۶/۷ |

همانطور که در جدول ۲ می‌بینید، روش حذف تدریجی با آستانه شباهت ثابت برای پایگاه داده SID به کیفیت مناسبی نرسیده است. این در حال است که با تغییر درصد حذف پیکسل‌های تیم در رابطه ۳ با سعی و خطا، درصد خطای کمتری به دست خواهد آمد (مثلاً به ازای ۵۰٪ در رابطه ۳ به درصد متوسط خطای ۲۴/۸۳٪ می‌رسیم). مزیت روش پیشنهادی در این مقاله برای تعیین حد آستانه شباهت، تعیین این آستانه از روی ویژگی‌های آماری رنگی تصویر است که آن را نسبت به تغییرات شرایط نوری و رنگی و بافتی تصاویر بیرونی مقاوم می‌کند و

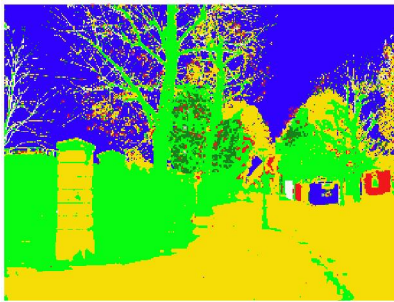
کاربر را از تنظیم الگوریتم به ازای شرایط مختلف تصاویر بی‌نیاز می‌نماید.

۵. تحلیل نتایج

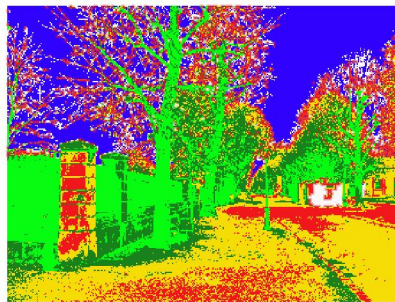
تصاویر به دست آمده نشان می‌دهد کیفیت روش پیشنهادی برای بخش‌بندی ابتدایی تصاویر مناسب است. به عنوان نمونه، در شکل ۱۵ تصویر بخش‌بندی شده به کمک روش پیشنهادی بیشتر از تصویر بخش‌بندی شده به کمک روش عادی، به تصویر مرجع شبیه است.



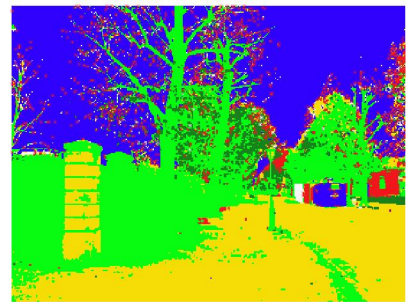
الف



ب



ج



د

شکل ۱۵) مقایسه الگوریتم‌های بخش‌بندی به روش k-means (الف) تصویر اصلی (ب) تصویر مرجع (ج) تصویر حاصل از روش k-means عادی (د) تصویر حاصل از روش حذف تدریجی

شکل ۴ و ۵ می‌بینید که آسمان به بیش از یک بخش تقسیم شده است؛ در حالی که انتظار داریم کل آسمان یک بخش در نظر گرفته شود. این موضوع نه تنها باعث اشتباه در بخش‌بندی اشیاء بزرگتر می‌شود، بلکه به دلیل تأثیرپذیری از تعداد پیکسل‌ها، گاهی باعث می‌شود اشیاء کوچک هم در تصویر بخش‌بندی شده نهایی حضور نداشته باشند. این مشکل

یک ویژگی جالب توجه روش پیشنهادی در شکل‌های ۵ و ۶ دیده می‌شود؛ در روش عادی، تعداد پیکسل‌ها روی نماینده‌ها تأثیر می‌گذارد. به همین دلیل اگر در تصویر چند سایه مربوط به یک رنگ با تعداد پیکسل زیاد موجود باشد (مثلاً بخش‌های مختلف آسمان یا چمن)، این سایه‌ها به عنوان اشیاء مختلف در نظر گرفته می‌شوند. مثلاً در هر دو

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هستوگرام با دقت دوگانه

عنوان یک شیء واحد در نظر گرفته می‌شوند، بلکه بخش‌هایی که در روش عادی از بین می‌روند (مانند اشیاء کوچک)، در روش پیشنهادی خود را نشان می‌دهند.

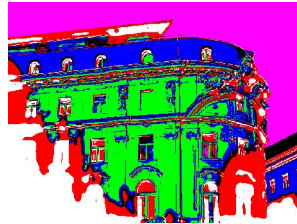
البته این ویژگی روش پیشنهادی در مواردی باعث بروز خطا نیز می‌شود. مثلاً در شکل ۱۶ می‌بینید که بخشی از ساختمان که روی آن سایه افتاده است، رنگی نزدیک به آسمان دارد. این موضوع باعث بروز خطا در روش پیشنهادی نسبت به روش عادی شده است.

یکی از نقاط ضعف عمده روش k -means برای کاهش رنگ به عنوان پیش‌گام بخش‌بندی است؛ به ویژه وقتی تصاویر مورد بررسی از تصاویر بیرونی باشند که از نور محیط تأثیر بسیار می‌پذیرند و گاه یک شیء واحد را در قالب چند رنگ نزدیک به هم نشان می‌دهند.

در روش پیشنهادی، در هر مرحله اشیائی که تعداد پیکسل‌های زیاد (و حتی چند سایه رنگی) دارند به عنوان یک شیء واحد در نظر گرفته شده و حذف می‌شوند و خوشه‌بندی مرحله بعد تنها بر پیکسل‌های باقیمانده متمرکز می‌شود؛ به همین دلیل نه تنها سایه‌های یک رنگ به



ج



ب



الف

شکل ۱۶ (الف) تصویر اصلی (ب) روش عادی (ج) روش پیشنهادی

خوشه‌های رنگی مناسب را در تصاویر ایجاد می‌کند. این در حالی است که در روش k -means در کمتر از ۳۰٪ موارد همگرایی به کمینه عمومی در مرتبه اول اجرای الگوریتم به دست خواهد آمد و با احتساب تکرارهای لازم برای همگرایی همه تصاویر، این زمان بیش از ۵۰ دقیقه خواهد بود. البته باید اذعان داشت که در مواردی که روش k -means به کمینه عمومی همگرا می‌شود، سرعت آن از روش پیشنهادی بیشتر است؛ چون ساختار مرحله به مرحله آن را ندارد. اما همانطور که ذکر شد، دقت الگوریتم پیشنهادی در تشخیص رنگ‌های صحیح از روش k -means بالاتر است.

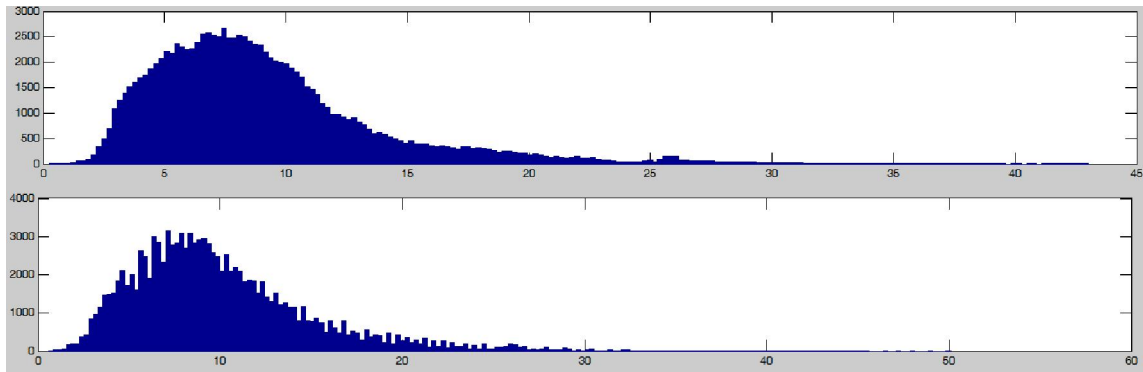
نکته دیگری که باید در مورد روش پیشنهادی مدنظر قرار گیرد این است که اگر سه رنگ ابتدایی (در واقع محل تصادفی اولیه سه بردار نماینده) به اشتباه انتخاب شود، کل الگوریتم با شکست مواجه می‌شود. البته این یک اشکال ذاتی الگوریتم k -means است و در نسخه استاندارد این روش نیز دیده می‌شود.

استفاده از هرم تصاویر چنددقتی باعث بهبود قابل توجهی در نتایج می‌شود؛ به گونه‌ای که بدون استفاده از طبقات این هرم خطای بخش‌بندی پیکسلی حدود ۶٪ افزایش می‌یابد. شکل ۱۷ نشان می‌دهد با استفاده از هرم فوق، تغییرات نمودار هستوگرام فاصله پیکسل‌های یک دسته از مرکز آن دسته ملایم‌تر می‌شود. به بیان دیگر با استفاده از این هرم، رنگهای نزدیک با هم ترکیب می‌شوند و باعث یکنواخت شدن

مزیت دیگر این روش دقت همگرایی و سرعت بیشتر آن نسبت به روش عادی است. همانگونه که پیشتر گفته شد، تکرار الگوریتم خوشه‌بندی k -means تا جایی ادامه می‌یابد که حاصل جمع «مجموع فواصل نقاط هر دسته تا نماینده آن دسته» کمینه شود (رابطه ۲ را ببینید). اگر انتخاب مکان اولیه بردارهای نماینده به درستی صورت نگیرد (که در انتخاب تصادفی تنها به اقبال شما بستگی دارد)، حاصل جمع فوق یا به کمینه‌ای همگرا نمی‌شود یا به یک کمینه محلی همگرا خواهد شد که طبعاً نتیجه مناسبی نخواهد داشت [۴۷]. این مشکل هنگام زیاد بودن تعداد خوشه‌ها، بیشتر بروز می‌کند؛ چون تعداد نماینده‌ها و به تبع آن احتمال واقع شدن تعدادی از نماینده‌ها در کمینه‌های محلی بیشتر می‌شود. به همین لحاظ الگوریتمی که از ابتدا تصویر را به ۶ یا ۹ رنگ تقسیم کند، در بسیاری از مواقع به کمینه عمومی همگرا نخواهد شد (در پایگاه داده ما، در بیش از ۷۰٪ موارد این مشکل ایجاد می‌شود). به علاوه تکرارهای زیاد الگوریتم برای همگرا شدن به یک کمینه عمومی، زمان اجرای آن را بالا می‌برد. در مقابل، روش پیشنهادی در هر مرحله تنها با سه بردار نماینده سروکار دارد؛ به همین لحاظ احتمال درگیر شدن با کمینه‌های محلی و طولانی شدن زمان همگرایی آن ناچیز خواهد بود. در برآورد زمانی کارایی الگوریتم که به کمک کامپیوتری با پردازنده پنتیوم Core i5 با فرکانس ۲ گیگاهرتز و RAM چهار گیگابایتی روی پایگاه داده UIDS انجام شد، الگوریتم پیشنهادی در کمتر از ۱۷ دقیقه

نکته کلیدی الگوریتم پیشنهادی، محو کردن تصاویر نیست؛ چون در کنار مزایای آن، اشیاء کوچک در محوسازی از بین می‌روند. استفاده از هرم تصاویر چنددقتی باعث به‌کارگیری همزمان تصاویر محوشده و تصاویر با وضوح بالا می‌شود تا در کنار توجه به اشیاء کوچک و جزئیات بافتی در موقعیت مناسب، از مزایای تصاویر محو شده نیز استفاده کند.

توزیع پیکسل‌های هم‌رنگ در یک دسته می‌شوند. این موضوع به نوبه خود سبب کاهش اشتباه الگوریتم بخش‌بندی در مواجهه با سایه‌های هم‌رنگ و نیز بافت‌های تصاویر بیرونی خواهد شد. البته باید به این نکته توجه داشت که با وجود اینکه محو کردن تصاویر با از بین بردن جزئیات بافتی و ترکیب سایه‌های رنگی به بخش‌بندی کمک می‌کند، اما



شکل ۱۷- مقایسه نمودار هیستوگرام فاصله‌های پیکسل‌های یک دسته تا مرکز دسته در هرم تصاویر چنددقتی (تصویر بالا) و بدون استفاده از هرم تصاویر چنددقتی (تصویر پایین)

تصاویر بیرونی را به حداقل می‌رساند. کارایی این الگوریتم به کمک یک روش ارزیابی نظارت‌شده روی دو پایگاه داده از تصاویر بیرونی بررسی شد. نتایج به‌دست آمده نشان می‌دهد این روش برای کاهش رنگ تصاویر بیرونی به هدف بخش‌بندی ابتدایی از الگوریتم k -means عادی بهتر عمل می‌کند.

در پایان باید به این نکته توجه داشت که در [۴۸] از آمارگان‌های مرتبه بالاتر نیز برای تشخیص پیکسل‌ها تیم بهره بردیم. اما روش تحلیل هیستوگرام که شکل توزیع رنگ‌ها در خوشه‌ها را در نظر می‌گیرد به مراتب ساده‌تر و کیفیت آن بالاتر است.

۶. نتیجه‌گیری و پیشنهادات

با انجام اصلاحاتی روی روش پیشنهادی، می‌توان به کیفیت بهتری دست یافت. مثلاً یکی از اشکالات به کارگیری هرم تصاویر چنددقتی این است که با اعمال فیلتر ملایم‌کننده روی تصویر، لبه‌های تصویر تغییر شکل و نیز تغییر رنگ می‌دهند که این موضوع باعث می‌شود لبه‌های تصویر به عنوان یک دسته رنگی جدید در نظر گرفته شوند. در شکل ۱۸-الف و ۱۸-ب، یک تصویر بیرونی و نسخه محو شده آن و در شکل ۱۸-ج لبه‌های تغییر رنگ داده مشاهده می‌شوند.

در این مقاله، مشکلات الگوریتم k -means برای کاهش رنگ تصاویر بیرونی به هدف بخش‌بندی ابتدایی و تشخیص شیء در آنها بررسی و روشی برای حل این مشکل پیشنهاد شد. در این روش که بر مبنای بستری متشکل از دو دقت متفاوت تصویر عمل می‌کند، با حذف تدریجی خوشه‌های اشیاء مهم در هر طبقه هرم، تمرکز روش خوشه‌بندی بر بقیه اشیاء افزایش می‌یابد. در این تحقیق، روشی برای تعیین حدود رنگ‌های مهم در هر مرحله پیشنهاد شد که با استفاده از ویژگی‌های آماری خوشه‌های رنگی، وابستگی الگوریتم به شرایط



ج

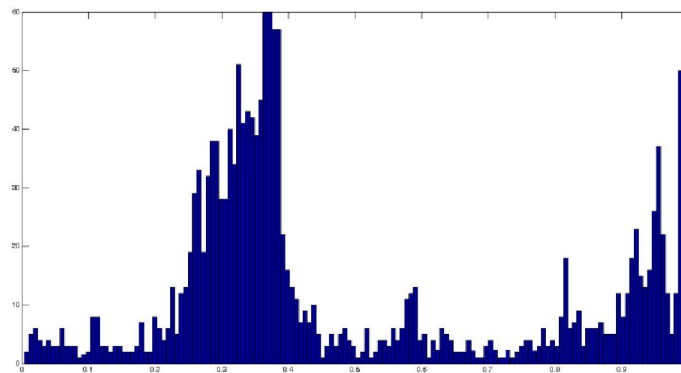
ب

الف

شکل ۱۸-الف) تصویر اصلی (ب) تصویر محو شده (ج) لبه‌های تغییر رنگ داده شده در اثر محو شدن

بخش‌بندی تصاویر رنگی بیرونی به هدف تشخیص اشیاء به کمک هیستوگرام با دقت دوگانه

است. واضح است هرچه این درصد بزرگتر باشد، تعداد پیکسل‌هایی که به مرحله بعد راه می‌یابند بیشتر خواهد بود و این موضوع احتمال تقسیم یک شیء واحد به چند بخش (*over-segmentation*) را افزایش می‌دهد؛ این در حالی است که اگر درصد فوق‌خوبی کوچک باشد، رنگ‌های کمی به مرحله بعد راه می‌یابند و احتمال از نظر دور ماندن بخش‌های کوچک (*under-segmentation*) بیشتر می‌شود [۴۰]. حتی اگر درصد فوق‌خوبی کوچک (مثلاً ۱٪) باشد، گاهی با بررسی هیستوگرام تهرنگ موفق به یافتن رنگ‌های *U* و *L* نمی‌شویم. به بیان دیگر، در هیستوگرام تهرنگ نمی‌توان رنگی یافت که تعداد پیکسل‌های آن کمتر از ۱٪ پیکسل‌های به رنگ μ باشد. نمونه‌ای از این هیستوگرام را در شکل ۱۹ می‌بینید.



شکل ۱۹- یک هیستوگرام تهرنگ «تخت»

نکته دیگر اینکه همزمان با شفاف‌سازی تصویر (عبور به طبقات پایینی هرم تصاویر چنددقتی) ویژگی‌های بافتی خود را بهتر نشان می‌دهند که می‌توان آنها را نیز برای بخش‌بندی ابتدایی تصویر مورد استفاده قرار داد. در واقع طول بردارهای مربوط به ویژگی‌های پیکسل‌ها می‌تواند با عبور به طبقات پایین هرم تصاویر چنددقتی افزایش یابد و شامل ویژگی‌های بافتی نیز شود.

هرچند در این مقاله به هدف بخش‌بندی تصاویر بیرونی الگوریتم خوشه‌بندی *k-means* سفارشی شد، اما روش پیشنهادی در این مقاله می‌تواند برای بهبود نتایج الگوریتم‌های دیگر خوشه‌بندی مانند [۱۵]، [۲۱] نیز مورد استفاده قرار گیرد.

پس از بخش‌بندی ابتدایی باید به کمک تکنیک‌های بازیابی تصاویر، ماهیت بخش‌ها تشخیص داده شود که بخش تکمیل‌کننده این تحقیق خواهد بود [۴۲، ۵۰].

برای رفع این مشکل، می‌توان ابتدا لبه‌های تصویر را به دست آورد و از در نظر گرفتن پیکسل‌های نزدیک به لبه‌ها به عنوان پیکسل یتیم اجتناب نمود؛ چون احتمالاً این پیکسل‌ها هم‌رنگ با زمینه هستند و به دلیل استفاده از فیلترهای ملایم‌کننده تغییر رنگ داده‌اند. میزان نزدیکی به لبه در روش بالا می‌تواند با توجه به قوت لبه تعیین شود.

برای بهبود دقت بخش‌های ایجاد شده، بعد از به کار بردن روش مزبور یک مرحله ادغام نیز می‌تواند انجام بپذیرد تا کلاس‌های ایجاد شده را در صورت لزوم بر حسب شباهت‌های رنگی و بافتی و شکلی و نیز پیوستگی پیکسل‌ها ادغام کند و یا اشیائی که تعداد پیکسل‌های آنها کم است (اشیاء کوچک یا کم اهمیت) را حذف کند. ایده‌هایی برای ادغام در [۴۹] مطرح شده است.

نکته دیگر در مورد درصد استفاده شده در مراحل ۲ و ۳ الگوریتم پیشنهاد شده در این تحقیق برای تعیین رنگ‌های *L* و *U* (در اینجا ۰.۵٪)

چنین می‌نماید که در این خوشه تمام رنگ‌ها وجود دارند! در حالی که این طور نیست؛ در چنین خوشه‌ای پارامترهای اشباع و روشنایی (*S* و *V* در فضای *HSV*) مقداری دارند که تغییر تهرنگ در رنگ‌های آن خوشه تأثیری ندارد و به همین لحاظ در هیستوگرام تهرنگ تمام رنگ‌ها دیده می‌شوند. مثلاً اگر روشنایی صفر باشد، پارامترهای *H* و *S* هر مقداری داشته باشند تفاوتی ندارد و آن رنگ، سیاه دیده می‌شود (نوار بالایی شکل ۱۱ را ببینید). بنابراین اگر هیستوگرام تهرنگ یک خوشه تخت باشد (رنگی که شرایط رنگ‌های *U* و *L* را ارضا کند وجود نداشته باشد)، رنگ‌های آن خوشه همگی به یک رنگ دیده می‌شوند و به همین لحاظ هیچ‌یک از آنها به عنوان رنگ یتیم در نظر گرفته نمی‌شوند. در هر حال تعیین بهینه مقدار درصد مذکور با توجه به شرایط تصویر می‌تواند به کیفیت الگوریتم پیشنهادی کمک کند.

پوشیدنی برای کمک به افراد نابینا و کم‌بینا در مسیریابی بیرون منزل و کمک به آن‌ها در مواجهه با چالش‌هایی نظیر عبور از خیابان و چهارراه بهره برد.

مراجع

- [1]. W. W. Mayol, "Wearable Visual Robots," Ph.D, Computer Science, University of Oxford, 2004.
- [2]. M. Everingham, B. T. Thomas, and T. Troscianko, "Wearable mobility aid for low vision using scene classification in a Markov random field model framework," *International Journal of Human Computer Interaction, special issue on mediated reality*, vol. 15, pp. 231-244, 2003.
- [3]. R. C. González and R. E. Woods, *Digital Image Processing*: Pearson/Prentice Hall, 2008.
- [4]. R. Manduchi, "Learning Outdoor Color Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1713-1723, 2006.
- [5]. J. Batlle, A. Casals, J. Freixenet, and J. Martí, "A review on strategies for recognizing natural objects in colour images of outdoor scenes," *Image and Vision Computing*, vol. 18(6-7), pp. 515-530, 2000.
- [6]. Y.-W. Tai, J. Jia, and C.-K. Tang, "Soft Color Segmentation and Its Applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 1520-1537, 2007.
- [7]. H. D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color Image Segmentation: Advances & Prospects," *Pattern Recognition*, vol. 34, pp. 2259-2281, 2001.
- [8]. H. B. M'hadheb, A. Douik, M. M. Fendri, and M. Annabi, "Reduction of color variability in color image segmentation," in *IEEE International Conference on Electronics, Circuits and Systems*, 2006.
- [9]. I. Ashdown, "Octree color quantization," in *Radiosity: A Programmer's Perspective*, ed: Wiley New York 1994.
- [10]. P. Heckbert, "Color image quantization for frame buffer display," *SIGGRAPH Comput. Graph.*, vol. 16, pp. 297-307, 1982.
- [11]. S. J. Wan, P. Prusinkiewicz, and S. K. M. Wong, "Variance based color image quantization for frame buffer display," *Color Res. Applicat*, vol. 15(1), pp. 52-58, 1990.
- [12]. P. Scheunders, "A comparison of clustering algorithms applied to color image quantization," *Pattern Recognition Letters*, vol. 18, pp. 1379-1384, 1997.
- [13]. N. Vlajic and H. C. Card, "Vector quantization of images using modified adaptive resonance algorithm for hierarchical clustering," *IEEE Transactions on Neural Networks*, vol. 12, pp. 1147-1162, 2001.
- [14]. B. Fritzke, "A Growing Neural Gas Network Learns Topologies," *Advances in Neural Information Processing Systems*, 1995.
- [15]. A. Baraldi and P. Blonda, "A survey of fuzzy clustering algorithms for pattern recognition. II," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 29, pp. 786-801, 1999.
- [16]. G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 3, pp. 698-713, 1992.

از نتایج این پژوهش می‌توان در بخش‌بندی و تشخیص صحیح اشیاء در تصاویر بیرونی و در کاربردهای مختلفی از جمله طراحی روبات‌های خودکار، کنترل ترافیک، کاربردهای امنیتی و نیز طراحی کامپیوترهای

- [17]. N. Papamarkos, A. E. Atsalakis, and C. P. Strouthopoulos, "Adaptive color reduction," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 32, pp. 44-56, 2002.
- [18]. G. Cheng, J. Yang, K. Wang, and X. Wang, "Image Color Reduction Based on Self-Organizing Maps and Growing Self-Organizing Neural Networks," in *The Sixth International Conference on Hybrid Intelligent Systems*, 2006, p. 24.
- [19]. K. Zagoris, N. Papamarkos, and I. Koustoudis, "Color Reduction Using the Combination of the Kohonen Self-Organized Feature Map and the Gustafson-Kessel Fuzzy Algorithm," in *The 5th international conference on Machine Learning and Data Mining in Pattern Recognition*, Leipzig, Germany, 2007, pp. 703-715.
- [20]. A. Atsalakis, N. Papamarkos, and I. Andreadis, "On estimation of the number of image principal colors and color reduction through self-organized neural networks," *International Journal of Imaging Systems and Technology*, vol. 12, pp. 117-127, 2002.
- [21]. J. Rasti, A. Monadjemi, and A. Vafaei, "Color reduction using a multi-stage Kohonen Self-Organizing Map with redundant features," *Expert Systems with Applications*, vol. 38, pp. 13188-13197, 2011.
- [22]. S. Kiranyaz, S. Uhlmann, and M. Gabbouj, "Dominant Color Extraction Based on Dynamic Clustering by Multi-dimensional Particle Swarm Optimization," in *The Seventh International Workshop on Content-Based Multimedia Indexing*, 2009, pp. 181-188.
- [23]. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*: Wiley, 2001.
- [24]. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*: Kluwer Academic Publishers, 1981.
- [25]. M. Y. Choong, W. L. Khong, W. Y. Kow, L. Angeline, and K. T. K. Teo, "Graph-Based Image Segmentation Using K-Means Clustering and Normalised Cuts," in *The Fourth International Conference on Computational Intelligence, Communication Systems and Networks*, 2012, pp. 307-312.
- [26]. Y. C. Hu and M. G. Lee, "K-means-based color palette design scheme with the use of stable flags," *Journal of Electronic Imaging*, vol. 16, pp. 033003-1 to 033003-11, 2007.
- [27]. S. N. Sulaiman and N. A. M. Isa, "Adaptive fuzzy-K-means clustering algorithm for image segmentation," *IEEE Transactions on Consumer Electronics*, vol. 56, pp. 2661-2668, 2010.
- [28]. P. Ng and C.-M. Pun, "Skin Color Segmentation by Texture Feature Extraction and K-mean Clustering," in *The Third International Conference on Computational Intelligence, Communication Systems and Networks*, 2011, pp. 213-218.
- [29]. R. Figueiredo, L. Schnitman, and F. d. Souza, "Using Neural Network and K-means Clustering for Image Segmentation in Outdoor Scenes," in *The 2nd International Congress on University-Industry Cooperation*, Perugia, Italy, 2007.
- [30]. R. Huang, N. Sang, D. Luo, and Q. Tang, "Image Segmentation via Coherent Clustering in Lab Color Space," *Pattern Recognition Letters*, vol. 32, pp. 891-902, 2011.

[۳۱]. جواد راستی، سید امیرحسین منجمی و عباس وفايي، «کاهش رنگ تصاویر بیرونی به هدف بخش‌بندی ابتدایی با استفاده از خوشه‌بندی سلسله‌مراتبی با

حذف تدریجی در هرم گوسی»، ششمین کنفرانس ماشین بینایی و پردازش تصویر، دانشگاه اصفهان، آبان ۱۳۸۹.

- [32]. A. Roy, S. K. Parui, D. Nandi, and U. Roy, "Color image segmentation using a semi-wrapped gaussian mixture model," in *The 4th international conference on Pattern recognition and machine intelligence*, Moscow, Russia, 2011, pp. 148-153.
- [33]. M. Recky and F. Leberl, "Windows Detection Using K-means in CIE-Lab Color Space," in *The 20th International Conference on Pattern Recognition*, 2010, pp. 356-359.
- [34]. S. Haykin, *Neural Networks: A Comprehensive Foundation*: Prentice Hall PTR, 1994.
- [35]. H. J. Aantonisse, "Image segmentation in pyramids," *Computer Graphics and Image Processing* vol. 19, pp. 367-383, 1982.
- [36]. R. Marfil, L. Molina-Tanco, A. Bandera, J. A. Rodriguez, and F. Sandoval, "Pyramid segmentation algorithms revisited," *Pattern Recognition*, vol. 39, pp. 1430-1451, 2006.
- [37]. G. Ramella and G. S. Baja, "Color Quantization by Multiresolution Analysis," in *The 13th International Conference on Computer Analysis of Images and Patterns*, Germany, 2009, pp. 525-532.
- [38]. A. Atsalakis and N. Papamarkos, "Color reduction and estimation of the number of dominant colors by using a self-growing and self-organized neural gas," *Engineering Applications of Artificial Intelligence*, vol. 19, pp. 769-786, 2006.
- [39]. S. Makrogiannis, G. Economou, and S. Fotopoulos, "A region dissimilarity relation that combines feature-space and spatial information for color image segmentation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 35, pp. 44-53, 2005.
- [40]. Y. J. Zhang, *Advances in Image And Video Segmentation*: IRM Press, 2006.
- [41]. J. Rasti, A. Monadjemi, and A. Vafaei, "A Graph-Based Vision System for Automatic Object Detection in Outdoor Scenes," in *The 22nd International DAAAM Symposium*, Vienna, Austria, 2011, pp. 0167-0168.
- [42]. A. Bosch, X. Munoz, and J. Freixenet, "Segmentation and description of natural outdoor scenes," *Image and Vision Computing*, vol. 25, pp. 727-740, 2007.
- [43]. H. Zhang, J. E. Fritts, and S. A. Goldman, "A Co-Evaluation Framework for Improving Segmentation Evaluation," in *SPIE Defense and Security Symposium - Signal Processing, Sensor Fusion, and Target Recognition XIV*, 2005, pp. 420-430.
- [44]. A. Alonso-Betanzos, B. Arcay-Varela, and A. Castro-Martinez, "Analysis and evaluation of hard and fuzzy clustering segmentation techniques in burned patient images," *Image and Vision Computing*, vol. 18, pp. 1045-1054, 2000.
- [45]. D. Collins, W. A. Wright, and P. Greenway, "The sowerby image database," presented at the The 7th IEEE International Conference of Image Processing and Its Applications, Manchester, England, 1999.
- [46]. X. He, R. S. Zemel, and M. Carreira-Perpi, "Multiscale conditional random fields for image labeling," in *IEEE computer society conference on Computer vision and pattern recognition*, Washington, D.C., USA, 2004, pp. 695-703.
- [47]. A. Likas, M. Vlassis, and J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognition* vol. 36, pp. 451-461, 2003.
- [۴۸]. جواد راستی، «ارائه یک روش بخش‌بندی مبتنی بر الگوریتم‌های هوشمند به منظور تشخیص اشیاء در تصاویر بیرونی»، پایان‌نامه دکترا، گروه مهندسی کامپیوتر، دانشگاه اصفهان، ۱۳۹۱.
- [49]. F. Y. Shih and S. Cheng, "Automatic seeded region growing for color image segmentation," *Image and Vision Computing*, vol. 23, pp. 877-886, 2005.
- [50]. R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, pp. 1-60, 2008.

تقویت محور مرکزی ساختارهای لوله‌ای و کاربرد آن در استخراج محور مرکزی سیاهرگ

پورتال

امیرحسین فروزان^{۱*} رضا آقایی‌زاده ظروفی^{**} یوشی‌نبو ساتو^{***} ماساتوشی هوری^{***}

* استادیار، دانشکده مهندسی پزشکی، دانشگاه شاهد، تهران

** استاد، دانشکده مهندسی برق و کامپیوتر، دانشگاه تهران، تهران

*** دانشیار، دانشکده پزشکی و رادیولوژی، دانشگاه اساکا، ژاپن

تاریخ دریافت: ۱۳۹۱/۰۳/۱۵

تاریخ پذیرش: ۱۳۹۱/۱۰/۱۵

چکیده

در این مقاله با ارائه توصیف جدیدی از ویژگی نقاط محور مرکزی ساختارهای لوله‌ای، روشی برای تقویت این ساختارها پیشنهاد شده است. در این روش، در یک چارچوب چندمقیاسی و با استفاده از بردارهای ویژه ماتریس هسین نقاط تصویر، فاصله هر نقطه را از لبه‌های تصویر به دست می‌آوریم. برای نقاطی که روی محور مرکزی قرار دارند این فاصله از دوسر هر راستای دلخواه متقارن است. در این مرحله با نمونه برداری فاصله هر نقطه از لبه‌های تصویر در راستاهای مختلف، به نقاطی که تقارن بیشتری دارند مقدار بیشتری نسبت می‌دهیم. در مرحله بعد برای تقویت محور مرکزی لوله‌ها، از یک فیلتر براساس روش Pock استفاده می‌کنیم. ارزیابی روش پیشنهادی با استفاده از تصاویر فانتوم دوبعدی و سه‌بعدی و داده‌های پزشکی به صورت کیفی و کمی با معیارهای حداکثر خطای تعیین محور مرکزی و نرخ آشکارسازی انجام گرفته است که مزیت این روش را به روش‌های موجود نشان می‌دهد.

کلید واژگان: استخراج ساختارهای لوله‌ای، استخراج محور مرکزی سیاهرگ پورتال، آنالیز تصاویر سی‌تی‌اسکن کبد، پردازش تصاویر پزشکی

۱. مقدمه

کبد بزرگترین غده بدن و بخشی از دستگاه گوارش است که ذخیره-سازی خون و فیلتر کردن آن، فعالیت‌های متابولیکی، عملکرد ترشحي و دفع فضولات را به‌عهده دارد. در کبد چهار ساختار آوندی وجود دارند که عبارتند از: سرخرگ هپاتیک، سیاهرگ هپاتیک، سیاهرگ پورتال و آوندهای صغراوی. بیماری‌های مربوط به کبد از جمله سرطان کبد یکی از عوامل اصلی مرگ و میر در جهان است. برداشتن تومورهای سرطانی، از متداول‌ترین روش‌های درمانی به‌شمار می‌رود که قبل از انجام جراحی، لازم است پزشک از موقعیت تقریبی تومورها در کبد و فاصله آنها نسبت به شاخه‌های اصلی ساختارهای آوندی آگاه باشد تا عمل جراحی با کمترین آسیب به ساختارهای آوندی انجام شود [۱]. استفاده از روش‌های تصویربرداری به‌عنوان یک روش غیرتهاجمی در تشخیص بیماری‌های کبد، طراحی روش درمان، ارزیابی تاثیر روش درمانی و آموزش اهمیت بسیاری دارد. استفاده از اسکنرهای سی‌تی برای تصویربرداری کبد در مقایسه با سایر مدولیت‌ها متداول‌تر است [۲]. این داده‌ها شامل صدها اسلایس می‌باشد که استخراج ساختارهای آوندی از آنها به‌روش دستی و توسط پزشک متخصص عملاً امکان‌پذیر نیست. بنابراین طراحی روشی برای استخراج آوندهای کبدی، انجام آنالیزهای کمی و نمایش سه‌بعدی نتایج تأثیر اساسی در زمینه روش‌های درمانی کبد خواهد داشت [۳].

تقویت محور مرکزی ساختارهای لوله‌ای و کاربرد آن در استخراج محور مرکزی سیاه‌رگ پورتال

جدول ۱ - مقادیر ویژه متناظر با ساختارهای مختلف در یک تصویر سه‌بعدی (با

$$\text{فرض } |\lambda_1| \leq |\lambda_2| \leq |\lambda_3| \quad [16]$$

| case | Corresponding pattern | λ_3 | λ_2 | λ_1 |
|------|-------------------------------|---------------|---------------|---------------|
| 1 | Noisy, no preferred direction | Noisy | Noisy | Noisy |
| 2 | Bright plate structure | High negative | Low | Low |
| 3 | Dark plate structure | High positive | Low | Low |
| 4 | Bright tubular structure | High negative | High negative | Low |
| 5 | Dark tubular structure | High positive | High positive | Low |
| 6 | Bright Blob structure | High negative | High negative | High negative |
| 7 | Dark blob structure | High positive | High positive | High positive |

طبق جدول ۱، اگر قدرمطلق مقادیر ویژه ماتریس $H^\sigma(X)$ (به صورت صعودی مرتب شده باشند) $(\lambda_i, 1 \leq i \leq 3)$ و بخواهیم یک ساختار لوله‌ای روشن را که در زمینه تاریک قرار دارد شناسایی کنیم، باید بدنبال واکنش‌هایی باشیم که در آن λ_1 تقریباً برابر صفر است و λ_2 و λ_3 دارای مقادیر بزرگ و منفی هستند. این ویژگی‌ها در مرکز لوله کاملاً برقرار می‌باشند و برای واکنش‌هایی که از مرکز لوله دورتر هستند، این شرایط تغییر می‌کند. روش‌های مبتنی بر ماتریس هسین از نظر محاسباتی کارآمد هستند ولی بدلیل درنظر گرفتن اطلاعاتی محلی، به نوبت حساس می‌باشند [۸].

فرانچی و همکاران با استفاده از مقادیر ویژه، فیلتری را برای تقویت ساختارهای لوله‌ای معرفی کرده‌اند که دارای سه مولفه برای تشخیص ساختارهای خطی از ساختارهای حبابی و صفحه‌ای و حذف نویز است (معادله (۲) و (۳)). در معادله (۲) باید شرط $\lambda_2 > 0$ or $\lambda_3 > 0$ برقرار باشد و یک تخمین خوب برای c ، برابر نصف حداکثر مقدار اندازه هسین تصویر است [۶].

$$V_0 = \left(1 - \exp\left(-\frac{R_A^2}{0.5}\right) \right) \exp\left(-\frac{R_B^2}{0.5}\right) \left(1 - \exp\left(-\frac{S^2}{2c^2}\right) \right). \quad (2)$$

$$R_A = \frac{|\lambda_2|}{|\lambda_3|}, R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}}, S = \sqrt{\sum_{j \leq 3} \lambda_j^2}. \quad (3)$$

آگام (Agam) و همکاران فیلتری برای تقویت ساختارهای لوله‌ای براساس مقادیر ویژه ماتریس کورلیشن بردارهای گرادیان طراحی کردند [۹]. چون فیلتر طراحی شده بر مبنای مشتقات مرتبه اول می‌باشد، نسبت

در این مقاله با ارائه توصیف جدیدی از ویژگی نقاط محور مرکزی لوله‌ها، روشی جهت تقویت ساختارهای لوله‌ای پیشنهاد شده‌است و کاربرد آن در بخش‌بندی سیاه‌رگ پورتال مورد بررسی قرار گرفته‌است. در بخش دوم مروری بر روش‌های متداول در تقویت محور مرکزی ارائه شده‌است. بخش سوم به توضیح روش پیشنهادی اختصاص یافته‌است. در بخش چهارم نتایج مربوط به پیاده‌سازی روش پیشنهادی با استفاده از تصاویر فانتوم دویعدی و سه‌بعدی و داده‌های پزشکی و ارزیابی آن ارائه شده‌است و بخش پنجم به ارزیابی نتایج می‌پردازد. بخش ششم به جمع‌بندی مقاله و معرفی کارهای آینده در این تحقیق اختصاص یافته‌است.

۲. مروری بر روش‌های تقویت محور مرکزی ساختارهای لوله‌ای

در مراجع [۴] و [۵] مروری بر تحقیقات قبلی در زمینه آنالیز ساختارهای لوله‌ای ارائه شده‌است. این تحقیقات با هدف تقویت ساختارهای لوله‌ای و محور مرکزی آنها، استخراج محور مرکزی و بخش‌بندی ساختارهای لوله‌ای انجام شده‌است. یکی از نکاتی که در طراحی الگوریتم‌ها مورد توجه محققین قرار گرفته‌است استفاده از تحلیل چند مقیاسی (Multiscale) ساختارها با استفاده از ماتریس هسین واکنش‌های تصویر می‌باشد [۶] و [۷-۱۳-۱۶]. اگر $I(X)$ ماتریس سه-بعدی تصویر و $I^\sigma(X) = I(X) * G(X|0, \sigma)$ تصویر هموار شده با هسته گوسی $G(X|0, \sigma)$ باشد، مقادیر ویژه ماتریس هسین واکنش‌های تصویر، اطلاعاتی درباره ساختارهای موجود در تصویر (ساختارهای حبابی، لوله‌ای، و صفحه‌ای) ارائه می‌دهند (جدول ۱) و بردارهای ویژه راست‌های را نشان می‌دهند که ساختار مرتبه دوم تصویر روی آنها قابل تجزیه است [۶]. ماتریس هسین هموار شده تصویر سه-بعدی در نقطه X توسط معادله (۱) تعریف می‌شود.

$$H^\sigma(X) = \begin{bmatrix} \frac{\partial^2 I^\sigma(X)}{\partial x^2} & \frac{\partial^2 I^\sigma(X)}{\partial x \partial y} & \frac{\partial^2 I^\sigma(X)}{\partial x \partial z} \\ \frac{\partial^2 I^\sigma(X)}{\partial y \partial x} & \frac{\partial^2 I^\sigma(X)}{\partial y^2} & \frac{\partial^2 I^\sigma(X)}{\partial y \partial z} \\ \frac{\partial^2 I^\sigma(X)}{\partial z \partial x} & \frac{\partial^2 I^\sigma(X)}{\partial z \partial y} & \frac{\partial^2 I^\sigma(X)}{\partial z^2} \end{bmatrix} \quad (1)$$

دارای انحنای زیاد هستند و نیز نقاط متعلق به انشعابات از جمله نکات قوت این روش می‌باشد. غیریکنواختی پاسخ فیلتر طراحی شده و عدم اتصال نتایج به‌دست آمده، از جمله اشکالات روش پیشنهادی است. این روش بر روی تعداد اندکی از داده‌های کلینکی آزمایش شده است [۱۲].

مانی‌سینگ^۲ و همکاران، روشی برای تقویت ساختارهای رنگی پیشنهاد کرده‌اند که در آن از فیلترهای نفوذ استفاده شده و تانسور نفوذ بر اساس فیلتر بهبود یافته پیشنهادی توسط فرانچی به‌دست آمده است (معادله (۴)) [۱۸].

$$D = Q\Lambda'Q^T. \quad (4)$$

در معادله (۴)، Λ' یک ماتریس قطری است که دارای مؤلفه‌های زیر است (معادله (۵)).

$$\lambda'_1 = 1 + (\omega - 1) \cdot V^{\frac{1}{s}}, \quad \lambda'_2 = \lambda'_3 = 1 + (\varepsilon - 1) \cdot V^{\frac{1}{s}}. \quad (5)$$

در معادله (۵)، S پارامتر مقیاس است و ω و ε مقادیر ثابت هستند. فیلتر بهبود یافته فرانچی^۳ (V) در معادله (۶) تعریف شده است.

$$V_s(\lambda) = \begin{cases} 0 & \text{if } \lambda_2 \geq 0 \text{ or } \lambda_3 \geq 0 \\ (1 - e^{-\frac{A^2}{2\alpha^2}}) \cdot e^{-\frac{B^2}{2\beta^2}} \cdot (1 - e^{-\frac{S^2}{2\gamma^2}}) \cdot e^{-\frac{2C^2}{|\lambda_1| \lambda_2^2}} & \text{otherwise} \end{cases} \quad (6)$$

در مرجع [۷]، Pock و همکاران برای تقویت محور مرکزی ساختارهای لوله‌ای، تابع محور مرکزی را مبتنی بر اندازه‌گیری مقدار گردایان حول محیط یک دایره ارائه کرده‌اند. یک فازور که بر اساس بردارهای ویژه ماتریس هسین تعریف می‌شود تابع را روی لبه رگ حرکت می‌دهد و متوسط وزن‌دار اندازه گردایان، خروجی تابع محور مرکزی خواهد بود [۷].

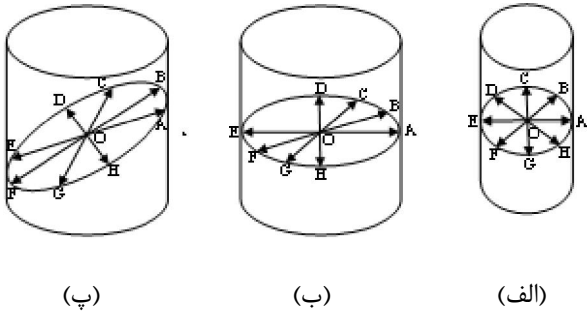
تحلیل روش‌های تقویت و استخراج ساختارهای لوله‌ای نشان می‌دهد که استفاده از ماتریس هسین در یک چارچوب چندمقیاسی برای تقویت ساختارهای لوله‌ای و محور مرکزی آنها نتایج خوبی خواهد داشت و معمولاً این روش‌ها به مداخله کاربر نیاز ندارند. برای کاهش حساسیت به نویز، استفاده از هسته گوسی برای هموار کردن تصویر پیشنهادی می‌شود ولی هموارسازی در مقیاس‌های بزرگ موجب می‌شود که شعاع رگ با تقریب محاسبه شود. در این مقاله، روشی برای تقویت محور مرکزی ساختارهای لوله‌ای با کاربرد در آنالیز سیاهرگ‌های

به فیلترهایی که بر اساس مشتقات مرتبه دوم هستند حساسیت کمتری به نویز دارد. در این روش پس از یافتن نقاط کاندید رگ بر اساس مقادیر ویژه و خطوط watershed، پنجره‌ای با ابعاد $3 \times 3 \times 3$ در نظر گرفته می‌شود که ابعاد این پنجره شروع به افزایش نموده و تا وقتی نسبت تعداد نقاط کاندید به ابعاد پنجره از اندازه یک مقدار آستانه کمتر نشده باشد، ابعاد پنجره افزایش می‌یابد. در این روش راستای رگ، برداری انتخاب می‌شود که بر تمام بردارهای گرادیان در آن نقطه عمود باشد.

تراک (Truc) و همکاران از بانک فیلترهای جهت‌دار برای تقویت ساختارهای لوله‌ای استفاده کرده‌اند [۱۷]. هریک از فیلترهای طراحی شده روی تصویر اعمال می‌شود و خروجی نهایی ماکزیمم خروجی فیلترها می‌باشد. روش پیشنهادی تراک (Truc) و همکاران روی تصاویر دوبعدی آنژیوگرافی و تصاویر شبکه چشم آزمایش شده است [۱۷].

چیان (Qian) و همکاران از پروفایل شدت‌روشنایی در مختصات قطبی استفاده کرده تا ویژگی انواع ساختارهای آوندی از جمله نقاط انشعاب را به‌دست آورند [۱۲]. پروفایل تغییرات شدت‌روشنایی در مختصات قطبی نشان می‌دهد که در نقاط محور مرکزی، دو ناحیه باند باریک (در راستای طولی رگ) وجود دارد. شدت‌روشنایی در راستای طولی رگ کمترین مقدار تغییرات و در صفحه عمود بر محور رگ بیشترین تغییرات را دارد. نواحی باند باریک دارای شدت‌روشنایی زیاد و تغییرات اندک هستند درحالی‌که نواحی باند پهن دارای تغییرات شدت‌روشنایی زیاد و متوسط شدت‌روشنایی کم هستند. بنابراین در مورد تعلق یک نقطه به ساختار رگ بر اساس پروفایل تغییرات شدت‌روشنایی همسایگی محلی آن نقطه تصمیم‌گیری می‌شود. محققین با ترسیم پروفایل برای یک داده مصنوعی که دارای انواع ساختارهای لوله‌ای، انشعابات و ساختارهای غیررگی است، نتیجه گرفته‌اند که اگر یک نقطه به رگ تعلق داشته باشد، در پروفایل آن، ناحیه باند باریکی یافت می‌شود که تغییرات شدت‌روشنایی در آن ناحیه کم و متوسط شدت‌روشنایی نسبت به نقاط مجاور بیشتر خواهد بود. نقاط متعلق به رگ شامل نقاط روی محور مرکزی، نقاط اطراف محور مرکزی، نقاط واقع در انشعابات، نقاط انتهایی رگ‌ها، رگ‌هایی با شدت‌روشنایی غیریکنواخت، رگ‌هایی با ضخامت غیریکنواخت، رگ‌های دارای گرفتگی، رگ‌های دارای انحنای زیاد، و نقاط متعلق به رگ‌های مماس بایکدیگر می‌باشند. روش پیشنهادی چیان^۱ و همکاران بر روی تصاویر مصنوعی و تصاویر کلینکی آزمایش شده است و نشان داده شده است که تقویت نقاطی از رگ که

² Manniesing³ Frangi¹ Qian



شکل ۱. توصیف محور مرکزی در روش پیشنهادی. محور مرکزی به‌عنوان مکان هندسی نقاطی که فاصله آن تا دو سر سطح لوله در هر راستای دلخواه با هم برابر است. (الف) برای یک لوله با مقطع دایروی، فاصله محور مرکزی تا دو سر سطح لوله در تمام راستاها با هم برابر است (OA=OB=OC=OD=OE=OF=OG=OH). (ب) برای یک لوله با مقطع بیضوی، فاصله محور مرکزی تا دو سر سطح لوله در هر راستاها با هم برابر است (OA=OE, OB=OF, OC=OG, OD=OH). (ج) توصیف پیشنهادی برای محور مرکزی در مقاطع غیر متعامد با راستای لوله نیز برقرار است (OA=OE, OB=OF, OC=OG, OD=OH).

اگر فرض کنیم نقاط ماکزیمم محلی در جهت‌های $\vec{V}(\theta) +$ و $\vec{V}(\theta) -$ به ترتیب $X_{\max+}$ و $X_{\max-}$ باشند، تابع $f_0(X|\theta)$ را توسط معادلات (۷) و (۸) تعریف می‌کنیم.

$$f_0(X|\theta) = \exp\left[-\frac{|dist(X, X_{\max+}) - dist(X, X_{\max-})|}{\sigma_d}\right] \quad (۷)$$

$$\arg\max_X [|\nabla I(X)|], \quad \vec{V}(\theta) \cdot \vec{X}_{\max^\pm} = \pm 1. \quad (۸)$$

در معادله (۷)، پارامتری σ_d که پهنای پاسخ فیلتر را تعیین می‌کند. نمونه‌برداری از تابع $f_0(X|\theta)$ در زاویه θ_i که طبق معادله (۹) تغییر می‌کند، انجام می‌شود.

$$\theta_i = \frac{\pi i}{N}, \quad i \in, \quad 0 \leq i < N, \quad N = \lfloor 2\pi r_{\max} \rfloor \quad (۹)$$

در معادله (۹)، ثابت r_{\max} حداکثر شعاع لوله‌ها در تصویر است. برای کاهش حساسیت خروجی تابع $f_0(X|\theta)$ به نویز، تابع $f(X|\theta_i)$ با معادله (۱۰) تعریف می‌شود.

$$f(X|\theta_i) = \begin{cases} 0 & \text{if } \frac{\nabla I(X_{\max+}) + \nabla I(X_{\max-})}{2} \leq \nabla I(X) + \delta \\ f_0(X|\theta_i) & \text{otherwise} \end{cases} \quad (۱۰)$$

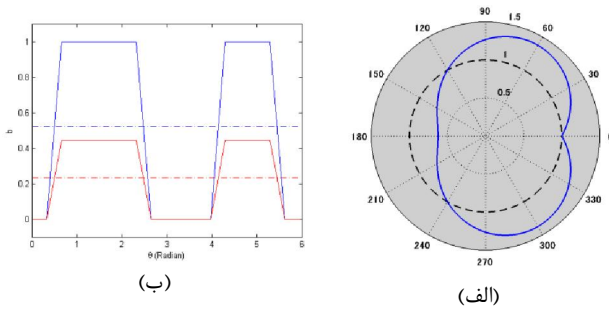
در معادله (۱۰)، $\nabla I(\cdot)$ گرادیان تصویر و δ پارامتری است که حساسیت تابع $f(X|\theta_i)$ به نویز را تعیین می‌کند. در این تحقیق مقدار

پورتال، پیشنهاد می‌شود که یک روش جدید چندمقیاسی مبتنی بر نمونه-برداری و تحلیل تقارن پروفایل رگ در راستاهای مختلف است. این روش برای داده‌های دوبعدی و سه‌بعدی پیاده‌سازی شده و نتایج آن با نتایج آخرین تحقیقات در زمینه روش‌های تقویت ساختارهای لوله‌ای مقایسه شده‌است. نتیجه این مقایسه نشان می‌دهد که این روش برتری کاملی نسبت به روش‌های موجود دارد.

۳. روش پیشنهادی

اگر نقطه X_0 متعلق به یک ساختار لوله‌ای از تصویر $I(X)$ باشد و مقادیر ویژه ماتریس هسین $H(X_0)$ در این نقطه بصورت صعودی $\lambda_1 \leq \lambda_2 \leq \lambda_3$ مرتب شده و بردارهای ویژه متناظر آنها به ترتیب \vec{v}_1, \vec{v}_2 و \vec{v}_3 باشند، بردارهای \vec{v}_1 و \vec{v}_2 صفحه‌ای را توصیف می‌کنند که بر مقطع رگ عمود است و بردار \vec{v}_3 در امتداد محور طولی رگ قرار خواهد داشت [۶]. نوآوری روش پیشنهادی برای تقویت محور مرکزی ساختارهای لوله‌ای، توصیف جدیدی از نقاط محور مرکزی می‌باشد. در روش پیشنهادی، محور مرکزی مکان هندسی نقاطی از لوله است که فاصله آن تا دو لبه لوله در هر راستا برابر می‌باشد. این توصیف برای لوله‌هایی با مقاطع دایروی (شکل ۱-الف و ب) و بیضوی (شکل ۱-پ) و برای هر راستای دلخواه برقرار است. در این توصیف تابعی تعریف می‌کنیم که به هر نقطه X از نقاط تصویر مقداری نسبت می‌دهد که متناسب با معکوس فاصله آن نقطه از محور مرکزی لوله است. اگر بردارهای ویژه \vec{v}_1 و \vec{v}_2 از راستای خود منحرف شوند، توصیف نقاط محور مرکزی به گونه‌ای است که مقدار تابع در آن نقاط تغییر نخواهد کرد.

نقطه X را در تصویر $I(X)$ در نظر می‌گیریم و مقادیر و بردارهای ویژه آن را به ترتیب $\lambda_1 \leq \lambda_2 \leq \lambda_3$ و $\vec{v}_1, \vec{v}_2, \vec{v}_3$ می‌نامیم. در نقطه X ، بردار $\vec{V}(\theta) = \sin(\theta)\vec{v}_1 + \cos(\theta)\vec{v}_2$ را که در آن $\theta \in [0, \pi]$ قرار دارد تشکیل می‌دهیم. می‌خواهیم فاصله این نقطه را در جهت‌های $\vec{V}(\theta) \pm$ تا دو لبه لوله به‌دست آوریم. پروفایل قدرمطلق گرادیان تصویر $|\nabla I_\sigma|$ در مرکز لوله دارای یک می‌نیمم محلی و در لبه‌های لوله دارای ماکزیمم محلی است.



شکل ۲. الف) مقطع غیر متقارن یک لوله (مقطع آبی رنگ)، ب) پاسخ فیلتر تقویت کننده محور مرکزی بدون استفاده از ضریب ایزوتروپیک (منحنی آبی) و با استفاده از ضریب ایزوتروپیک (منحنی قرمز).

۴. نتایج

ارزیابی روش پیشنهادی، با استفاده از تصاویر فانتوم دوبعدی و سه بعدی و تصاویر فاز دوم سی تی اسکن کبد انجام شده است. روش پیشنهادی در این مقاله با روش های اخیر در این زمینه شامل روش های Lee [۱۱]، Qian [۱۲]، Pock [۷]، Bauer [۱۹] و Truc [۱۷] مقایسه شده است. پیاده سازی در محیط برنامه نویسی MATLAB و C++ و با استفاده از یک کامپیوتر مبتنی بر سیستم عامل ویندوز با پردازنده Intel[®] Core™-i5 2.27GHz و 4GB-DRAM انجام شده است. مجموعه داده های پزشکی، فاز دوم تصاویر سی تی اسکن ناحیه شکم و متعلق به دانشکده پزشکی دانشگاه اساکا - ژاپن است. رزولوشن داده ها ۰/۶۲۵ × ۰/۶۲۵ میلی متر مربع و فاصله بین هر اسلایس ۱/۲۵ میلی متر می باشد. در هر مجموعه ۱۵۹ اسلایس و اندازه هر اسلایس ۵۱۲ × ۵۱۲ است.

یکی از مواردی که در استخراج محور مرکزی ساختارهای لوله ای اهمیت دارد حالتی است که در آن دو لوله در مجاورت یکدیگر و در بعضی نقاط مماس بایکدیگر هستند. در مرجع [۱۱]، Lee و همکاران با طراحی دو لوله مستقیم و منحنی به شعاع ۳ واکسل که در وسط تصویر به یکدیگر مماس هستند، ارزیابی روش خود را انجام داده اند. در این تحقیق برای مقایسه روش پیشنهادی با روش Lee از این داده سه بعدی استفاده شد. در شکل ۳-الف محور مرکزی استخراج شده به وسیله روش پیشنهادی برای این داده، نشان داده شده است. در شکل ۳-ب حداکثر خطای محور مرکزی برحسب واکسل نشان داده شده و با حداکثر خطای محور مرکزی به روش Lee مقایسه شده است.

δ ثابت و برابر با ۰/۰۵ قرار داده شده است. برای کاهش پاسخ فیلتر در لبه های تصویر، تابع $f(X)$ را با متوسط گیری وزن دار از نمونه های تابع $f(X|\theta_i)$ توسط معادلات (۱۱) و (۱۲) تعریف می کنیم.

$$f(\mathbf{x}) = \frac{1}{N} \sum_{i=0}^{N-1} w[f(X|\theta_i)] f(X|\theta_i). \quad (11)$$

$$w[f(X|\theta_i)] = \exp\left[-\frac{(1-f(X|\theta_i)/f_0(X))^2}{2\sigma_w^2}\right]. \quad (12)$$

در معادله (۱۲)، پارامتر حساسیت پاسخ فیلتر است که در این تحقیق این پارامتر ثابت و برابر ۰/۰۲ قرار داده شده است. تابع تعریف شده توسط معادله (۱۱)، پاسخ اولیه ای برای تقویت محور مرکزی ایجاد می کند که وظیفه آن تقویت تمام ساختارهای لوله ای در تصویر است. اکنون خروجی این مرحله به عنوان ورودی برای فیلتر پیشنهادی توسط Pock و همکاران [۷] استفاده می شود (معادله (۱۳) - (۱۵)).

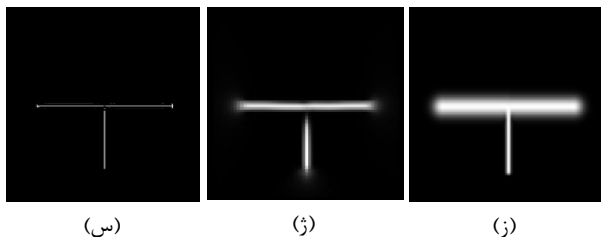
$$M^\sigma(\mathbf{x}) = \frac{1}{N} \sum_{i=0}^{N-1} w(b_i) b_i. \quad (13)$$

$$w(b_i) = \exp\left[-\frac{1}{2} \frac{(1 - \frac{b_i}{M_0^\sigma})^2}{0.04^2}\right], \quad (14)$$

$$M_0^\sigma(\mathbf{x}) = \frac{1}{N} \sum_{i=0}^{N-1} b_i, \quad b_i = |\nabla f(\mathbf{x} + \sqrt{3}\sigma \vec{V}(\theta_i))|. \quad (15)$$

مقدار M^σ خروجی تابع تقویت محور مرکزی خواهد بود. برای کاهش خروجی فیلتر تقویت کننده محور مرکزی برای ساختارهای لوله ای که مقطع آنها متقارن نیست، ضریب ایزوتروپیک $w(b_i)$ ، طبق معادله (۱۴) تعریف می شود. در شکل ۲، اثر ضریب ایزوتروپیک برای یک ساختار لوله ای که مقطع آن متقارن نیست، نشان داده شده است. در شکل ۲-الف، مقطع لوله به رنگ آبی نشان داده شده است. در شکل ۲-ب، پاسخ فیلتر تقویت کننده محور مرکزی بدون استفاده از ضریب ایزوتروپیک (منحنی آبی) و با استفاده از ضریب ایزوتروپیک (منحنی قرمز) با یکدیگر مقایسه شده اند. همانطور که در شکل ۲-ب ملاحظه می شود، استفاده از ضریب ایزوتروپیک می تواند خروجی فیلتر را برای لوله هایی با مقاطع نامتقارن کاهش دهد.

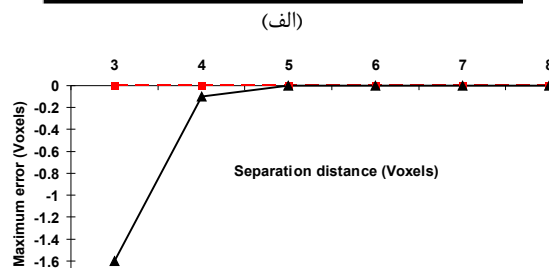
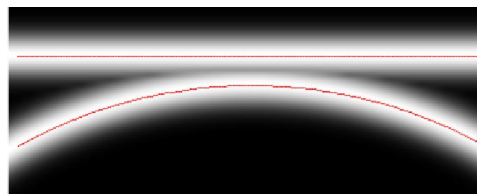
تقویت محور مرکزی ساختارهای لوله‌ای و کاربرد آن در استخراج محور مرکزی سیاهرگ پورتال



شکل ۴. مقایسه کیفی روش پیشنهادی با روش باوئر و همکاران. تصاویر فانتوم متعلق به کریسیان و فارن‌بک [۲۰ و ۲۱] می‌باشند. ستون سمت راست: تصاویر ورودی، ستون وسط: تقویت محور مرکزی توسط روش باوئر، ستون سمت چپ: تقویت محور مرکزی با روش پیشنهادی.

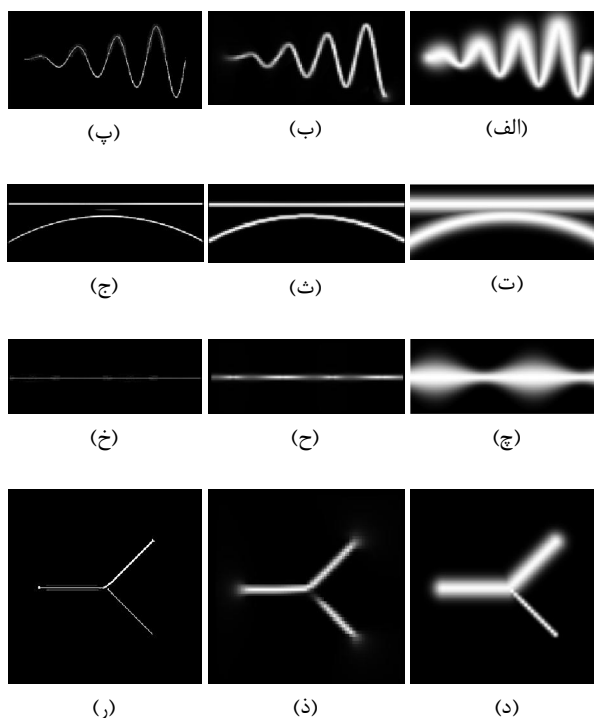
فانتوم عمومی دیگری که توسط محققین مورد استفاده قرار گرفته است متعلق به ماهادوان و همکاران است [۶]. این فانتوم که یک تصویر دوبعدی است شامل طیف گسترده‌ای از ساختارهای لوله‌ای می‌باشد. اگر قطعات موجود در تصویر اصلی (شکل ۵-الف) را از بالا به پایین و از چپ به راست شماره‌گذاری کنیم، قطعه یک ساختار رگ‌های انشعابی را شبیه‌سازی می‌کند و قطعات چهار تا هفت انشعابات با قطرهای مختلف را نشان می‌دهند. قطعه سه از چندین لوله تشکیل شده که در جهت‌های مختلف قرار گرفته‌اند. قطعات دو، ۱۲ و ۱۴ به گونه‌ای طراحی شده‌اند که در طول لوله، شعاع آن دارای تغییرات باشد. همچنین چالش‌هایی نظیر لوله‌های موازی که در کنار یکدیگر قرار دارند (قطعات هشت، نه و ۱۱)، لوله‌هایی با شعاع بسیار کم (قطعه ۱۰)، لوله‌های منقطع (قطعه ۱۳)، و لوله‌هایی که شدت روشنایی آنها در طول لوله تغییر می‌کند (قطعه ۱۵) در این فانتوم ایجاد شده‌اند [۱۰]. در مرجع [۱۲]، Qian و همکاران معیار نرخ آشکارسازی^۴ که در معادله (۱۶) و (۱۷) تعریف شده است را برای مقادیر مختلف واریانس نویز تصویر دوبعدی ماهادوان و همکاران و در نرخ ثابت False alarm=0.5% محاسبه نموده‌اند [۱۲]. نتیجه این ارزیابی در جدول ۲ نشان داده شده است.

در شکل ۵ ارزیابی کمی روش پیشنهادی با استفاده از تصویر دوبعدی مرجع [۱۲] که به وسیله نویز گوسی با انحراف معیار ۱۰٪ آغشته شده (شکل ۵-ب)، نشان داده شده است. در شکل ۵-پ نتیجه تقویت ساختارهای لوله‌ای با روش پیشنهادی و در شکل ۵-ت تا (خ)، نمای نزدیکتر قسمت‌هایی از شکل ۵-پ نشان داده شده است.



شکل ۳. (الف) نمایش MIP تصویر دو لوله مستقیم و منحنی که در وسط تصویر مماس به یکدیگر می‌باشند. محور مرکزی لوله‌ها که با روش پیشنهادی بدست آمده در تصویر با رنگ قرمز نشان داده شده است. (ب) منحنی حداکثر خطای محور مرکزی در روش پیشنهادی و روش Lee.

باوئر و همکاران ارزیابی روش پیشنهادی را به صورت کمی و با استفاده از تصاویر فانتوم کریسیان و فارن‌بک [۲۰ و ۲۱] انجام داده‌اند. در شکل ۴، مقایسه بین روش پیشنهادی و روش باوئر انجام شده است [۱۹].



⁴ Detection rate

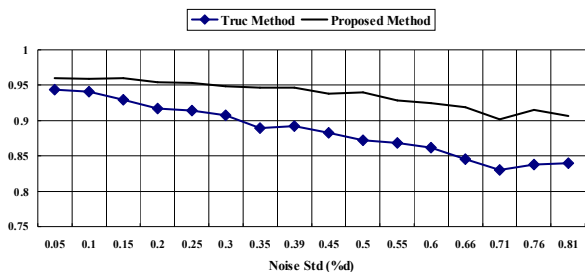
جدول ۲- مقایسه نرخ آشکارسازی روش پیشنهادی و روش Qian.

| انحراف معیار نویز تصویر | روش Qian | روش پیشنهادی |
|-------------------------|----------|--------------|
| 0.01 | 0.92 | 0.99 |
| 0.05 | 0.83 | 0.96 |
| 0.10 | 0.73 | 0.85 |
| 0.15 | 0.65 | 0.83 |
| 0.20 | 0.59 | 0.83 |
| 0.25 | 0.53 | 0.81 |
| 0.30 | 0.47 | 0.79 |
| 0.40 | 0.38 | 0.78 |
| 0.50 | 0.29 | 0.76 |

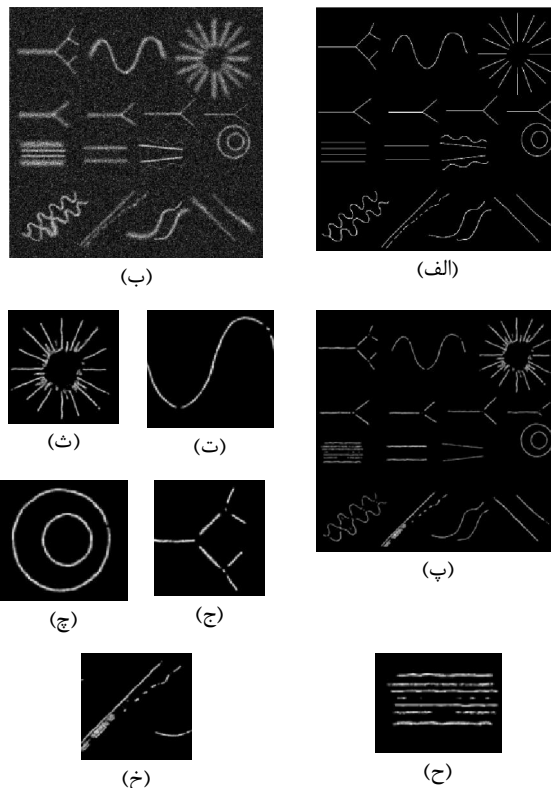
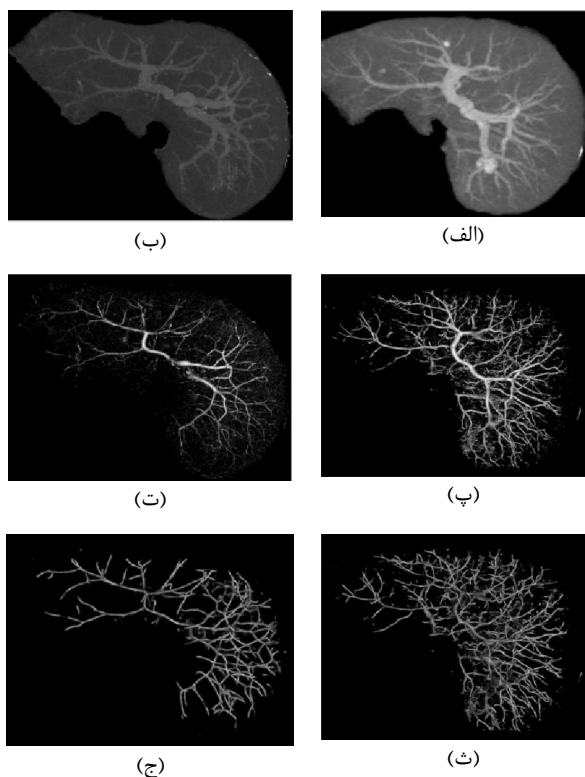
$$Detection\ Rate = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (16)$$

$$False\ Alarm = 1 - \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (17)$$

انجام داده‌اند [۱۷]. آنها با محاسبه سطح زیر منحنی ROC (AUC) برای مقادیر مختلف نویز تصویر، ارزیابی کمی روش خود را انجام داده‌اند. در شکل ۶، نتیجه مقایسه روش پیشنهادی با روش تراک و همکاران با محاسبه سطح زیر منحنی ROC نشان داده شده است.



شکل ۶- مقایسه کمی روش تراک و همکاران با روش پیشنهادی. مقایسه سطح در تقویت ساختارهای لوله‌ای برای روش پیشنهادی و روش تراک زیر منحنی ROC در مقادیر مختلف نویز تصویر. تصویر مورد استفاده فانتوم دوبعدی ماهادوان می‌باشد.

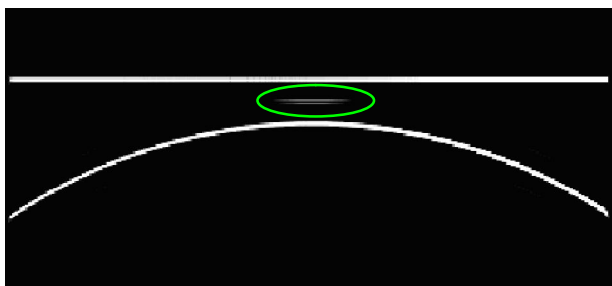


شکل ۷- تقویت محور مرکزی سیاهرگ پورتال در تصاویر فاز دوم سی تی اسکن، (ت) و (ج) Pock کبد. (الف) و (ب) تصویر اصلی، (پ) و (ت) نتیجه روش نتیجه روش پیشنهادی.

شکل ۵- ارزیابی روش پیشنهادی با استفاده از تصویر دوبعدی مرجع [۱۰]. (الف) تصویر مرجع، (ب) تصویر آغشته به نویز گوسی با انحراف معیار ۰.۱۰، (پ) نتیجه تقویت محور مرکزی با روش پیشنهادی، (ت) تا (خ) نمای نزدیکتر از قسمت‌هایی از تصویر قسمت (پ).

تراک و همکاران روش خود را روی فانتوم دوبعدی ماهادوان و همکاران [۱۰] اجرا کرده‌اند و ارزیابی آن را به صورت کمی و کیفی

ارزیابی انجام شده در روش باوئر از نوع کیفی است و با نمایش تصویر MIP انجام شده است. طبق نتایج شکل ۴، روش باوئر در اتصالات ساختارها (شکل ۴-ذ)، (ژ)) و برای لوله‌هایی که قطر آنها متغیر است (شکل ۵-۱۸-ح))، نتیجه خوبی ندارد و این اشکال برای غالب روش‌های تقویت ساختارهای لوله‌ای وجود دارد [۱۹]. در حالی که بررسی نتایج شکل ۴-خ)، (ز)، و (س) نشان می‌دهد که روش پیشنهادی در این تحقیق، محور مرکزی لوله‌هایی با قطر متغیر را به خوبی استخراج می‌کند و نقاط اتصال در ساختار لوله‌ای را بهتر از روش باوئر تقویت می‌نماید. در مورد لوله‌های مارپیچ (شکل ۴-الف)) و لوله‌های مماس (شکل ۴-ت))، روش باوئر محور مرکزی را به خوبی و بدون تداخل با اجزاء مجاور تقویت کرده است که دلیل آن عدم استفاده از چارچوب چند مقیاسی و فیلترهای گوسی با انحراف معیار زیاد است (شکل ۴-ب) و (ث)). مقایسه این نتایج با نتایج شکل ۴-پ) و (ج) نشان می‌دهد که روش پیشنهادی در این تحقیق، محور مرکزی این لوله‌ها را بهتر از روش باوئر تقویت کرده است. البته در مورد شکل ۴-ج)، پاسخ روش پیشنهادی در ناحیه بین دو لوله (ناحیه مشخص شده در شکل ۸)، به دلیل این است که در ناحیه مرکزی بین دو لوله مماس، شرط تقارن فاصله از دو لبه لوله در تعداد کمی از راستاها برقرار است که در نتیجه پاسخ ضعیفی از الگوریتم تقویت محور مرکزی در این ناحیه ایجاد می‌شود.



شکل ۸- بررسی پاسخ روش پیشنهادی تقویت محور مرکزی برای لوله‌های مماس. ناحیه مشخص شده در شکل پاسخ ضعیف روش پیشنهادی به ناحیه بین دو لوله را نشان می‌دهد.

اهمیت نتایج شکل ۵-پ) تا (خ) در این است که فانتوم مورد استفاده شامل طیف گسترده‌ای از ساختارهای لوله‌ای از جمله رگ‌های انشعابی، انشعابات با قطرهای مختلف، انشعابات در جهت‌های مختلف، لوله‌هایی با قطر متغیر در طول لوله، لوله‌های موازی، لوله‌هایی با شعاع بسیار کم، لوله‌های منقطع و لوله‌هایی که شدت‌روشنایی آنها در طول لوله تغییر می‌کند، می‌باشد. در مرجع [۱۲]، Qian و همکاران با استفاده

در ارزیابی روش پیشنهادی از تصاویر پزشکی نیز استفاده شده است. در شکل ۷، مقایسه نتیجه تقویت محور مرکزی سیاه‌رگ پورتال با روش پیشنهادی و روش Pock انجام شده است.

۵. ارزیابی روش پیشنهادی

روش ارائه شده توسط Lee [۱۱] یک روش از نوع مسیریابی است که نقاط شروع مسیریابی نقاطی هستند که دارای حداکثر شدت‌روشنایی در تصویر می‌باشند. بنابر گفته محققین، روش پیشنهادی برای یافتن نقاط شروع در تصاویر با نویز کم نتایج خوبی خواهد داشت، بنابراین اجرای این روش روی تصاویر سی‌تی‌اسکن کبد که نویز قابل توجهی دارد، نتیجه ندارد. همچنین این محققین روش پیشنهادی خود را روی تصاویر سی‌تی‌اسکن رگ‌های قلب موش با ابعاد $256 \times 1000 \times 1000$ و اندازه هر واکسل $20/9 - 8/2$ میکرومتر اجرا کرده‌اند که در مقایسه با تصاویر پزشکی مورد استفاده در این مقاله، دارای رزولوشن بالاتری است. مقایسه کیفی روش پیشنهادی با روش Lee در شکل ۳-ب)، دقت روش پیشنهادی را برای استخراج محور مرکزی نسبت به روش Lee نشان می‌دهد. محور مرکزی به‌وسیله روش پیشنهادی در این مقاله بدون خطا به دست آمده است. برای کاهش خطا، ابعاد داده با روش درونیابی به چهار برابر اندازه اصلی افزایش یافته است. در شکل ۳-الف)، محور مرکزی استخراج شده به رنگ قرمز روی تصویر اصلی نشان داده شده است که نشان می‌دهد روش پیشنهادی می‌تواند به خوبی محور مرکزی ساختارهای لوله‌ای که در مجاورت هم و به‌صورت مماس با هم قرار دارند استخراج نماید. نویز تصاویر سی‌تی‌اسکن قابل چشم‌پوشی نیست که یکی از علل شکست روش‌های تقویت رگ‌ها در کبد به‌شمار می‌رود. بنابراین روش لی برای تصاویر سی‌تی‌اسکن با رزولوشن بالا طراحی شده و نمی‌توان از آن برای استخراج سیاه‌رگ پورتال استفاده کرد.

روش باوئر (Bauer) و همکاران [۱۹] برای حذف چارچوب چندمقیاسی پیشنهاد شده است، تا بدون استفاده از هموارسازی توسط هسته‌های گوسی با انحراف معیار زیاد بتواند لوله‌هایی را که در مجاورت هم قرار دارند به‌خوبی آشکار کند. برای تصاویری که نویز در آنها زیاد است، استفاده از تصویر میدان برداری گرادینان به جای تصویر هموار شده می‌تواند علاوه بر حذف نویز تصویر، مرزهای لوله‌ها را حفظ کند و تصویر نتیجه برای آشکارسازی لوله‌ها مناسب‌تر است. البته باوئر پیشنهاد کرده است که اگر نویز تصویر زیاد باشد از هموارسازی گوسی به‌وسیله هسته گوسی با انحراف معیار یک استفاده می‌کنند.

مقایسه این نتایج با نتایج روش Pock نشان می‌دهد که نویز خروجی روش پیشنهادی کمتر و قطر ساختارهای تقویت شده برای همه رگ‌ها یکسان است.

۶. جمع بندی و پیشنهادات

تقویت رگ‌ها در تصویر سی‌تی‌اسکن کبد برای طراحی درمان و ارزیابی روش درمانی اهمیت بسیاری دارد. روش استخراج محور مرکزی پیشنهادی در این مقاله با استفاده از نقاط قوت روش‌های موجود در تقویت و استخراج ساختارهای لوله‌ای و ترکیب آن با توصیف جدیدی از ویژگی نقاط محور مرکزی توانسته است نسبت به روش‌های موجود نتایج بهتری بدست آورد. خاصیت انتگرال‌گیری موجود در این روش مقاومت بیشتر در برابر نویز را فراهم کرده است. این روش برای طیف گسترده‌تری از ساختارهای لوله‌ای مستقیم، خمیده، با پروفایل مقطع دایروی و بیضوی می‌تواند به کار رود. اجرای این روش روی مجموعه‌ای از تصاویر فانتوم دوبعدی و سه‌بعدی و داده‌های مربوط به فاز دوم سی-تی‌اسکن کبد و مقایسه آن با روش‌های موجود، مزیت روش پیشنهادی را نشان می‌دهد.

در آینده کاربرد روش پیشنهادی در تقویت سایر ساختارهای لوله‌ای در تصاویر پزشکی، از جمله استخراج مسیرهای هوایی، استخراج رگ‌های قلبی و استخراج رگ‌ها در تصاویر کره چشم می‌تواند علاوه بر ارزیابی روش پیشنهادی، زمینه‌های جدیدی را برای کاربردهای آن فراهم آورد.

مراجع

- [1]. E. N. Marieb, Human Anatomy and Physiology. The Benjamin/Cummings Publishing Company, 1992, pp. 796-797.
- [2]. A. M. R. Agur, Atlas of Anatomy. Canada: Williams and Wilkins, 1991, p.79.
- [3]. J. H. D. Fasel et al., "Segmental Anatomy of the Liver: Poor Correlation with CT," Radiology, vol. 206, no. 1, pp. 151-156, 1998.
- [4]. D. Lesage et al., "A Review of 3D Vessel Lumen Segmentation Techniques: Models, Features and Extraction Schemes," Medical Image Analysis, vol. 13, no. 6, pp. 819-845, 2009.
- [5]. C. Kirbas and F. Quek, "A Review of Vessel Extraction Techniques and Algorithms," ACM Computing Surveys, vol. 36, no. 2, pp. 81-121, 2004.
- [6]. A.F. Frangi et al., "Multiscale vessel enhancement filtering," In Proc. 1st MICCAI, 1998, pp. 130-137.
- [7]. T. Pock et al., "Multiscale Medialness for Robust Segmentation of 3D Tubular Structures," 10th computer vision winter workshop, 2005, pp. 93-102.

از ویژگی پروفایل به ویژه در انشعابات، در مجاورت با سایر ساختارها و در نقاط انتهایی، روشی را برای آشکارسازی ساختارهای لوله‌ای پیشنهاد کرده‌اند. این روش توسط تصاویر فانتوم دوبعدی و تصاویر سی‌تی‌اسکن رگ‌های قلب موش با اندازه واکسل ۱۰۰ میکرومتر به روش کمی و کیفی ارزیابی شده است. بررسی نتایج شکل ۴ و مقایسه آن با نتایج مرجع [۱۲] نشان می‌دهد که خروجی روش پیشنهادی نسبت به روش Qian دارای نویز کمتری است. البته روش پیشنهادی نقاط انشعاب (شکل ۴-ج) را تقویت نمی‌کند و نواحی بین لوله‌های مجاور را به دلیل تقارن فاصله این نقاط از دو سر ساختارهای لوله‌ای مجاور تقویت کرده است. مقایسه روش پیشنهادی و روش Qian که در جدول ۲ انجام شده است، نشان دهنده برتری روش پیشنهادی برای تقویت ساختارهای لوله‌ای در مقایسه با روش Qian است.

روش تراک و همکاران شامل مجموعه‌ای از فیلترهای دوبعدی است که ساختارهای لوله‌ای را در زوایای مختلف تقویت کرده و خروجی نهایی، متوسط پاسخ فیلترها در نقاط تصویر است. مقایسه کمی انجام شده روش پیشنهادی با روش تراک (شکل ۶) برتری کامل روش پیشنهادی در این رساله را نشان می‌دهد. سطح زیرمنحنی روش پیشنهادی در این تحقیق در محدوده انحراف معیار نویز [۰/۸]، حداقل ۰/۹۰ است.

ارزیابی روش پیشنهادی در مورد تصاویر پزشکی (شکل ۵-ث) و (ج) نشان می‌دهد که این روش توانسته است بخوبی محور مرکزی سیاهرگ پورتال را از تصاویر فاز دوم سی‌تی‌اسکن کبد استخراج نماید.

- [8]. S. Wörz and K. Rohr, "Segmentation and Quantification of Human Vessels Using a 3-D Cylindrical Intensity Model." IEEE Trans. on Image Processing, vol. 16, no.8, pp. 1994-2004, 2007.
- [9]. G. Agam, S.G. Armato III, and C. Wu, "Vessel-tree reconstruction in thoracic CT scans with application to nodule detection," IEEE Transactions on Medical Imaging, vol. 24, no. 4, 2005.
- [10]. V. Mahadevan, H. Narasimha-Iyer, B. Roysam, H.L. Tanenbaum, "Robust model-based vasculature detection in noisy biomedical images. IEEE Trans. Inf. Technol. Biomed., vol. 8, no. 3, pp. 360-375, 2004.
- [10]. J. Lee, P. Beighley, E. Ritman, N. Smith, "Automatic segmentation of 3D micro-CT coronary vascular images," Med. Imag. Anal., vol. 11, no. 6, pp. 630-647, 2007.
- [11]. X. Qian, M.P. Brennan, D.P. Dione, et al., "A non-parametric vessel detection method for complex vascular structures," Med. Imag. Anal., vol. 13, issue 1, pp. 49-61, 2009.
- [12]. K. Krissian, G. Malandain, N. Ayache, R. Vaillant and Y. Troussset, "Model-based multiscale detection of 3D vessels," in Proc. CVPR '98, 1998, pp. 722-727.
- [13]. C. Lorenz, I-C. Carlsen, T. M. Buzug, C. Fassnacht and J. Weese, "Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical

- images,” in Proc. First Joint Conference on CVRMed and MRCAS, Lecture Notes in Computer Science 1205, Springer-Verlag, Berlin, 1997, pp. 233–242.
- [14]. T. M. Koller, G. Gerig, G. Szekely and D. Dettwiler, “Multiscale detection of curvilinear structures in 2D and 3D image data,” in Proc. 5th Int. Conf. Computer Vision, Boston, MA, 1995, pp. 864-869.
- [15]. Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig and R. Kikinis, “Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images,” *Med. Imag. Anal.*, vol. 2, no. 2, pp. 143-168, 1998.
- [16]. P.T.H. Truc, Md.A.U. Khan, Young-Koo Lee, Sungyoung Lee and Tae-Seong Kim, “Vessel enhancement filter using directional filter bank,” *Comput. Vision and Imag. Understanding*, vol. 113, no.1, pp.101-112, 2009.
- [17]. R. Manniesing, M. A. Viergever and W. Niessen, “Vessel enhancing diffusion: A scale space representation of vessel structures,” *Med. Imag. Anal.*, vol. 10, no. 6, pp. 815-825, 2006.
- [18]. C. Bauer and H. Bischof, “A novel approach for detection of tubular objects and its application to medical image analysis,” in Proc. of DAGM., 2008, pp. 163–172.
- [19]. K. Krissian, G. Malandain, N. Ayache, R. Vaillant and Y. Troussel, “Model-based detection of tubular structures in 3D images,” *Computer Vision and Image Understanding*, vol. 80, no. 2, pp. 130–171, 2000.
- [20]. K. Krissian and G. Farneback, “Building reliable clients and servers,” in *Medical Imaging Systems Technology: Methods in Cardiovascular and Brain Systems*, World Scientific Publishing Co., Singapore, 2005.

پردازش تصاویر ورق های فولادی به منظور آشکارسازی عیوب به کمک موجک گابور

مصطفی صادقی^{۱*} مسعود شفیعی^{**}

* کارشناس ارشد، دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی واحد زواره، اصفهان

** استاد، دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، تهران

تاریخ دریافت: ۱۳۹۱/۰۴/۱۵

تاریخ پذیرش: ۱۳۹۱/۱۰/۲۵

چکیده

در مراحل مختلف تولید فولاد، خرابی‌هایی متعددی بر سطح ورق ظاهر می‌شود. صرف نظر از دلایل ایجاد خرابی‌ها، تشخیص دقیق انواع آن‌ها به طبقه بندی صحیح ورق فولاد کمک می‌کند و در نتیجه در صد بالایی از فرآیند کنترل کیفیت را به خود اختصاص می‌دهد. کنترل کیفیت ورق‌های فولادی به منظور بهبود کیفیت محصول و حفظ بازار رقابتی از اهمیت بالایی برخوردار می‌باشد. در این مقاله ضمن بررسی اجمالی تکنیک‌های پردازش تصویر مورد استفاده، با به‌کارگیری پردازش تصویر به کمک موجک گابور دو بعدی راه حل سریع و با دقت بالا برای آشکار سازی عیوب بافتی ورق‌های فولادی ارائه شده است. در ابتدا با استفاده از موجک گابور ویژگی‌های بافتی قابل توجهی را از تصاویر استخراج می‌کند که هم دربرگیرنده‌ی جهات مختلف و هم فرکانس‌های مختلف می‌باشد. سپس با استفاده از روش آماری، تصاویری که دربردارنده‌ی عیوب به طور واضح‌تری هستند انتخاب شده و محل وقوع عیب تعیین می‌گردد. با ارائه‌ی نمونه‌های آزمایشی میزان دقت و سرعت عمل روش به‌کار گرفته شده نشان داده شده است.

کلید واژگان: پردازش تصویر، بازرسی اتوماتیک، کنترل کیفیت،

بخش‌بندی عیوب، موجک گابور و بعدی

۱. مقدمه

کنترل کیفیت یکی از مسائل مهم در صنعت تولید ورق‌های فولادی به شمار می‌آید. تشخیص عیوب سطحی، درصد بالایی از فرآیند کنترل کیفیت را به خود اختصاص می‌دهد. تولید کنندگان فولاد می‌خواهند خرابی‌هایی غیر منتظره را بشناسند تا مانع از اتفاق افتادن آن‌ها بطور

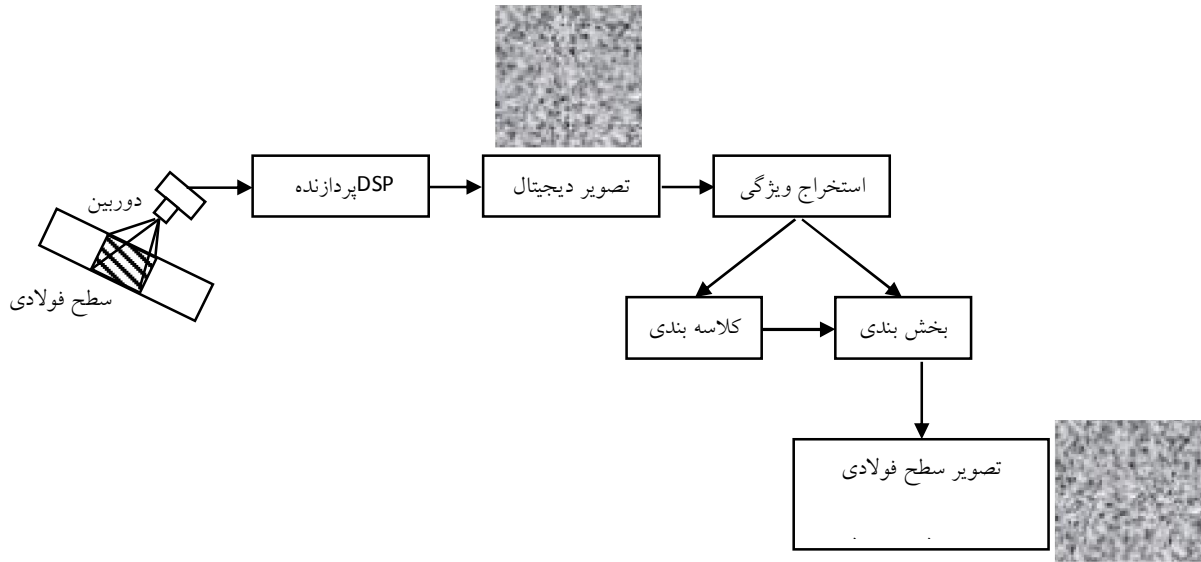
مداوم بشوند. و مطمئن باشند که کیفیت محصولات مطابق با نیاز مصرف کننده می‌باشد [۱]. از طرفی خرابی‌هایی ایجاد شده باید در گزارش‌های کیفی آماری ثبت شود. بدون یک سیستم بازرسی سطوح شناسایی خرابی سطوح، ثبت و بررسی علل خرابی می‌تواند مدت زیادی طول بکشد. در طول این تأخیر ممکن است که مشکل دوباره تکرار شده و باعث کاهش کیفیت تولید شود.

امروزه در اکثر خطوط تولید ورق، کنترل کیفیت توسط نیروهای خبره و به صورت دستی انجام می‌گیرد. بررسی محصولات کارخانه به طور دیداری سرعت کار پروسه، زمان مصرفی، و درآمد حاصله را کم می‌کند. علاوه بر این ممکن است شخص ناظر در محیط کارخانه آسیب ببیند. از طرفی با توجه به هزینه سنگین بازرسی انسانی، سرعت خط تولید و تکراری بودن کار ایجاد یک سیستم تشخیص خرابی و بازرسی سطوح در بهبود و کارایی تولید فولاد بسیار موثر می‌باشد [۲].

با اینکه امروزه بعضی از کارخانجات تجاری، از تکنیک‌های بررسی و عیب یابی سطح ورق فولاد استفاده می‌کنند، ولی بررسی، تشخیص و طبقه بندی خرابی‌های سطحی همزمان با تولید محصول هنوز هم بطور گسترده در سطح دنیا بررسی می‌شود. عدم وجود یک سیستم کنترل کیفیت خودکار باعث کاهش راندمان کاری، عدم دقت کافی و افزایش هزینه‌ها می‌گردد. در شکل ۱ نمونه ای از یک سیستم بازرسی اتوماتیک سطوح فولادی در کارخانه نشان داده شده است.

پردازش تصویر فناوری غالب امروزی در زمینه‌ی بررسی بافت‌های مختلف و شناسایی تنوع موجود در آن است. توانمندی این فناوری به ویژه در دو زمینه تشخیص و طبقه‌بندی الگو، باعث بهره‌گیری از آن در کنترل کیفیت صنایعی مانند پارچه، کاغذ و سرامیک شده است.

پردازش تصاویر ورق‌های فولادی به منظور آشکارسازی عیوب به کمک موجک گابور



شکل ۱- نمونه ای از یک سیستم بازرسی اتوماتیک سطوح فولادی در کارخانه

عیوب سطحی ورق فولاد پرداخته است. روش ماتریس هم رخداد کارایی ضعیفی در عیوب بافت پارچه در مقایسه با تکنیک‌های دیگری از قبیل روش مبتنی بر فیلتر از خود نشان داده است. در [۸-۹] با استفاده از فیلتر گابور ویژگی‌های بافتی قابل توجهی از تصاویر استخراج شده که هم دربرگیرنده‌ی جهات مختلف و هم دربرگیرنده‌ی فرکانس‌های مختلف است. در [۱۰] نیز روش‌های ارائه شده است که در تمامی آن‌ها عمل استخراج ویژگی از تصاویر مورد نظر با استفاده از فیلتر گابور انجام شده است. فیلتر گابور به دلیل بهینه بودن در هر دو حوزه‌ی مکانی و فرکانسی، امکان بهره‌گیری از محاسن پردازش سیگنال را در هر دو حوزه فراهم می‌کند.

در این مقاله در بخش ۲ به بررسی کلی خرابی‌هایی سطح ورق فولاد پرداخته و سپس از بین این خرابی‌ها چهار نوع خرابی متداول را برای بررسی انتخاب می‌کنیم. سپس در بخش ۳ به بررسی روش تحلیل بافت و معرفی موجک گابور و سپس به استخراج ویژگی از تصاویر به کمک موجک گابور پرداخته می‌شود. مدل پیشنهادی و نتایج به دست آمده از آزمون‌های مختلف نیز در بخش ۴ معرفی می‌گردد. در بخش ۵ با نشان دادن نتیجه‌ی آزمایش‌های انجام گرفته بر روی مجموعه تصاویر، به دو صورت جدول نتایج و تصویر حاصل از آزمایش، کارایی روش پیشنهادی نشان داده می‌شود. در نهایت در بخش ۶ نتیجه‌های حاصل شده از روش پیشنهادی در این مقاله، ارائه می‌گردد.

محققین زیادی خرابی‌هایی ایجاد شده بر سطح فولاد را به روش‌های گوناگون مورد بررسی قرار داده‌اند. برخی از آن‌ها از روش پیدا کردن لبه استفاده نموده‌اند [۳].

در این روش لبه‌های موجود در تصویر با استفاده از یک الگوریتم پیدا کردن لبه تشخیص داده می‌شود و از مقایسه‌ی تصویر لبه‌های موجود با مشخصات ذخیره شده در حافظه، عیب مورد نظر تشخیص داده می‌شود. مشکل اصلی این روش این است که خرابی‌هایی مختلف می‌توانند اشکال هندسی یکسانی داشته باشند ولی از لحاظ ساختمانی کاملاً با هم متفاوت باشند. به همین دلیل این روش در مورد خیلی از خرابی‌هایی هم شکل قابل استفاده نیست. از جمله روش‌های دیگری که در تعیین محل عیب استفاده می‌شود می‌توان به روش‌هایی مانند فیلتر لاپلاس، فیلتر گرادیان اشاره کرد [۴-۶]. اعمال این فیلترها بر روی تصویر باعث ظاهر شدن لبه‌های عیب و جزئیات بی اهمیت صفحه با یک شدت می‌شود که با اعمال روش‌های ریخت شناسی، لبه‌های عیب نیز حذف می‌گردند. بنابراین نتایج این روش‌ها بر روی عیب مورد نظر در این مقاله قابل قبول نمی‌باشد.

مجموعه کارهایی که تا کنون انجام گرفته است بیشتر با تمرکز بر روی ویژگی‌های رنگ، شکل و بافت سعی در ارائه روش‌های مناسب آشکارسازی خرابی‌ها نموده‌اند. در روش‌های مبتنی بر تحلیل بافت، هدف اصلی فراهم کردن معیاری برای تشخیص خواص بافتی تصویر مانند زبری، نرمی، همجنسی، همواری و ... است. در [۷] با به کار بردن یکی از روش‌های تحلیل بافت به نام ماتریس هم رخداد به آشکار سازی

۲. پیش‌زمینه

حوزه‌ی مکان تعریف شود، با تصویر مورد نظر کانولوشن^۱ شده و تصویر جزئی را ایجاد می‌کند و چنانچه در حوزه‌ی فرکانس تعریف شود، با گرفتن تبدیل سریع فوریه از تصویر مورد نظر آن را به حوزه‌ی فرکانس برده و سپس با موجک گابور حوزه‌ی فرکانس ضرب می‌گردد. با انتقال حاصل ضرب انجام شده به حوزه‌ی مکان، تصویر جزئی به دست می‌آید. از آنجا که عمل کانولوشن در حوزه‌ی مکان دیرتر از ضرب در حوزه‌ی فرکانس انجام می‌شود ما در این مقاله از موجک گابور در حوزه‌ی فرکانس استفاده می‌کنیم تا با سرعت بالاتری ویژگی‌های تصاویر را استخراج کنیم. از آنجایی که بیشتر خرابی‌ها در سطح فولاد در زمان تولید و به طور تصادفی ایجاد می‌شوند، نمی‌توان یک موجک گابور خاص را برای تشخیص هر نوع خرابی به کار برد. به همین دلیل، در این مقاله از بانک موجک‌های گابور جهت استخراج ویژگی استفاده می‌شود. وجود فرکانس‌ها و جهات متفاوت در بانک موجک‌های گابور، موجب می‌شود که ویژگی‌های استخراج شده حاوی مقدار زیادی اطلاعات در مورد بافت تصویر باشند و بتوانند خرابی‌ها را در جهات و فرکانس‌های مختلف را به خوبی تشخیص دهند. موجک گابور دو بعدی مورد استفاده در این مقاله در رابطه (۱) نشان داده شده است.

$$(u, v) = e^{-\pi \left[\frac{u_p^2}{\sigma_x^2} + \frac{v_p^2}{\sigma_y^2} \right]} \cdot e^{-2\pi j(x_0 u + y_0 v)} \quad (1)$$

مقادیر u_p و v_p با استفاده از رابطه (۲) به دست می‌آیند:

$$\begin{aligned} u_p &= (u - w_x) \cos(\theta) + (v - w_y) \sin(\theta) \\ v_p &= -(u - w_x) \sin(\theta) + (v - w_y) \cos(\theta) \end{aligned} \quad (2)$$

که در این فرمول مقادیر W_x و W_y فرکانس‌های مرکزی فیلتر در جهات X و Y می‌باشد. همچنین x_0 و y_0 میزان جابجایی افقی و عمودی در حوزه‌ی مکان می‌باشد. در آزمایش‌های انجام شده، مقدار $X_0=0$ ، $Y_0=0$ و $W_x=W_y$ و $\sigma_x = \sigma_y$ قرار داده شده است. از آنجا که تا کنون هیچ روش تحلیلی جهت بهینه کردن بانک فیلترهای گابور ارائه نشده، لازم است از یک روند طولانی آزمایش و ارزیابی جهت بهترین پیکربندی از این بانک فیلترها استفاده شود. در آزمایش‌های انجام شده پس از بررسی و آزمون‌های مختلف، این نتیجه حاصل شد که استفاده از یک بانک فیلتر با دوازده فیلتر گابور با فرکانس‌های $\left\{ \frac{\Omega_m}{8}, \frac{\Omega_m}{4}, \frac{\Omega_m}{2} \right\}$ و زوایای $(0, 45, 90, 135)$ می‌تواند افت را به خوبی استخراج کند.

با توجه به ماهیت فناوری ماشین بینایی و پردازش تصویر، در اولین مرحله استفاده از این فناوری، باید یک تصویر مناسب از شی مورد نظر به وجود آید. تهیه این تصویر مناسب مستلزم شناخت کافی از خصوصیات ظاهری شی است. لذا در بررسی عیوب سطح ورق فولاد لازم است در ابتدا دسته بندی بر اساس خصوصیات ظاهری عیوب انجام شود. این دسته بندی مقدمه ای راهگشا برای بررسی روش‌های نور پردازی و تصویر برداری و همچنین دستیابی به الگوریتم‌های پردازشی مناسب می‌باشد [۷]. در مجتمع فولاد مبارکه اصفهان- ایران حدود ۲۱۰ خرابی کدگذاری شده، از لحاظ تشابه ظاهری خصوصیات نورپردازی و تصویربرداری و الگوریتم‌هایی پردازشی به گروه‌های ده‌گانه‌ای تقسیم شده و به تفصیل مورد بررسی قرار گرفته است. بر اساس نتایج این مطالعه و با در نظر گرفتن پارامترهایی از قبیل زیاد بودن تعداد دفعات اتفاق افتادن عیب (رایج بودن)، تنوع شکل عیوب و اهمیت تفکیک آن‌ها، ۴ خرابی حفره، خراش، چین‌های عرضی و زنگ‌زدگی، در این مقاله جهت بررسی انتخاب شده است [۹-۱۰]. شکل ۲ چهار تصویر واقعی از چهار نمونه سطوح فولاد دارای خرابی‌های رایج را نشان می‌دهد. برای انجام تحقیق به عکس‌های متعددی از این عیوب نیاز است. عکس‌های استفاده شده متعلق به دانشگاه کانپور هند می‌باشد [۳].

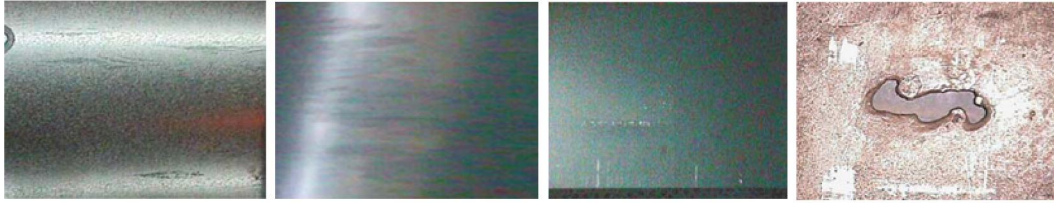
۳. تحلیل بافت و استخراج ویژگی به کمک موجک

گابور

فناوری غالب امروزی برای تشخیص و دسته‌بندی اشیا بر مبنای ویژگی ظاهری، فناوری پردازش تصویر است. به‌طور کلی عملیات پردازش تصویر در دو مرحله‌ی عمده‌ی استخراج ویژگی‌ها و کلاسه‌بندی انجام می‌گیرد. در مرحله‌ی استخراج ویژگی با تعیین ویژگی‌های مورد نظر و انتخاب پارامترها و روش استخراج، ابتدا از تصویر خام این ویژگی‌ها تفکیک و بهینه می‌شود. در مرحله کلاسه‌بندی و در نهایت بخش‌بندی، نواحی دارای بافت مشابه مشخص شده، مرزهای بین بافت‌های مختلف تعیین می‌گردد.

از پرکاربردترین روش‌های مورد استفاده در استخراج ویژگی‌های بافتی، استفاده از موجک گابور می‌باشد. موجک گابور، ویژگی‌های بافتی قابل توجهی را از تصاویر استخراج می‌کند که هم دربرگیرنده‌ی جهات مختلف و هم فرکانس‌های مختلف است. چنانچه موجک گابور در

¹ Convolution



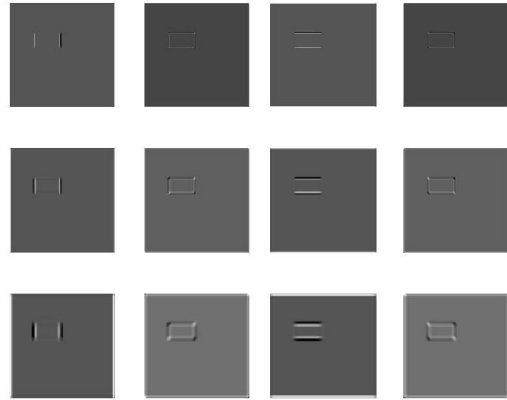
شکل ۲- چهار تصویر واقعی از چهار نمونه سطوح فولاد

عمل نرمال سازی صورت می‌گیرد. چنانچه مقدار هر پیکسل در یک تصویر جزئی $f(X,Y)$ باشد که در آن x و y مختصات پیکسل در تصویر جزئی را نشان می‌دهد. نرمال سازی به صورت فرمول ۳ انجام می‌گیرد.

$$f_{norm}(x,y) = \frac{f(x,y) - \min(f)}{\max(f) - \min(f)} \quad (3)$$

که در آن \min و \max به ترتیب نمایانگر کمترین و بیشترین مقادیر تصویر جزئی f و $F_{norm}(x,y)$ نرمال شده‌ی $f(x,y)$ می‌باشد. پس از انجام این نرمال‌سازی، مقادیر هر یک از تصاویر جزئی در بازه‌ی $[0,1]$ قرار می‌گیرد.

با اعمال ۵ سطح خاکستری ۰، ۱ و ۲ و ۳ و ۴ با ۴ سطح آستانه ۰/۲ و ۰/۴ و ۰/۶ و ۰/۸ روی تصاویر جزئی به خوبی میزان پراکندگی و تمرکز داده‌ها قابل تشخیص است. تصاویر جزئی میزان تمرکز و پراکندگی داده‌ها را به صورت قابل لمس‌تری نشان می‌دهند. آن دسته از تصاویر که میزان پراکندگی داده‌ها در آن‌ها کمتر باشد یا به عبارتی متمرکزتر باشد، به طور مشخص‌تری حاوی ناحیه‌ی دارای خرابی می‌باشد زیرا ناحیه‌ی دارای خرابی باعث می‌شود که ویژگی‌های این پیکسل‌ها، شفاف‌تر از بقیه‌ی تصویر باشند. در تصاویری که عیب در آن‌ها مشهود نیست معمولاً تفاضل انرژی محاسبه شده نزدیک به صفر است و در محدوده‌ی پراکندگی محاسبه شده قرار نمی‌گیرند. از آنجا که در بعضی از تصاویر جزئی عیب به صورت بهتری قابل مشاهده است و در بعضی دیگر عیب وجود ندارد یا نویزهای موجود در تصویر به اشتباه به صورت عیب در نظر گرفته شده‌اند با ارائه‌ی الگوریتمی این دو دسته تصویر را از هم جدا کرده و تصاویری که به صورت شفاف‌تری نشان‌دهنده‌ی عیب می‌باشند را جهت شناسایی عیب با هم ترکیب کرده و تصویر حاصل از ترکیب را به تصویر دودویی تبدیل کردیم. در الگوریتم پیشنهادی تصاویر دارای پراکندگی زیاد حذف شده و تصاویر دارای پراکندگی داده‌ای کم و تمرکز بیشتر که خرابی‌ها را به طور مشخص‌تر نشان می‌دهند جهت ترکیب انتخاب می‌شوند. پس از ترکیب صحیح تصاویر جزئی با



شکل ۳- پاسخ بانک موجک گابور در فرکانس‌ها و زوایای اعلام شده به یک تصویر معیوب می‌باشد.

Ω_m بیش‌ترین فرکانس تصویر بوده و برابر نصف ابعاد تصویر در حوزه‌ی فرکانس می‌باشد. باید توجه داشت که افزایش تعداد موجک‌ها از حدی بیشتر کارایی را چندان افزایش نمی‌دهد و افزایش بار محاسباتی را نیز به همراه دارد.

۴. روش پیشنهادی

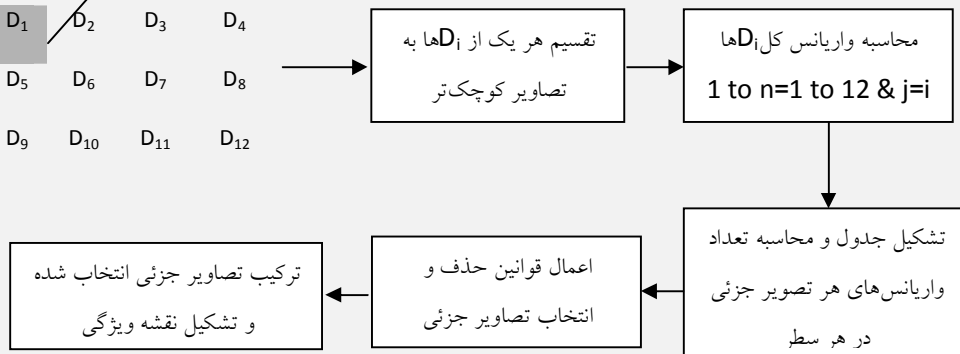
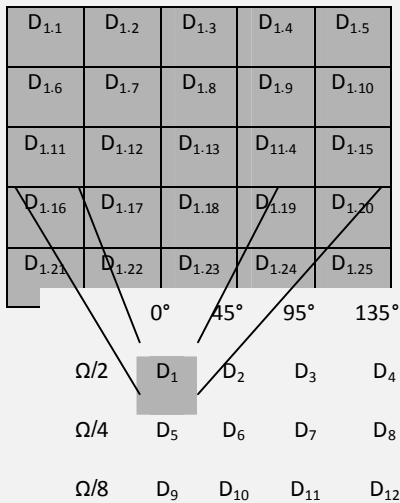
در این مقاله روشی پیشنهاد می‌شود که در آن بخش‌بندی عیوب بر مبنای محاسبه‌ی میزان پراکندگی در تصاویر جزئی است. به این صورت که ابتدا استخراج ویژگی توسط بانک فیلترهای گابور با دوازده فیلتر که در قسمت قبل بیان شد، انجام می‌شود.

پس از استخراج ویژگی از تصاویر دارای خرابی و تکمیل تصاویر جزئی و محاسبه‌ی انرژی، تصاویر جزئی که بطور مشخص‌تری حاوی ناحیه‌ی دارای خرابی هستند جهت ترکیب انتخاب و تصاویر جزئی که فاقد نواحی معیوب هستند را حذف می‌کنیم. از آنجا که مقادیر انرژی بدست آمده از هر پیکسل از تصاویر جزئی در محدوده‌ی وسیعی از اعداد حقیقی پراکنده شده‌اند با محدود کردن این مقادیر به چند مقدار محدود بهتر می‌توان میزان تمرکز و پراکندگی داده‌ها را تشخیص داد. از طرفی با انجام این عمل حجم داده‌ها را کاهش داده و سرعت پردازش بالا می‌رود. ابتدا روی کلیه‌ی مقادیر موجود در هر یک از تصاویر جزئی

تصویری دیگر هیچ یک از تصاویر جزئی انتخاب نشوند. مراحل کلی انجام الگوریتم حذف تصاویر جزئی با روش پیشنهادی در شکل ۴ نشان داده شده است.

یکدیگر، ناحیه‌ی معیوب کاملاً مشخص شده و درصد نواحی سالمی که به اشتباه معیوب شناخته می‌شوند بسیار ناچیز است. در یک تصویر دارای خرابی ممکن است کل تصاویر جزئی انتخاب شوند و در

تبدیل D_i به قسمت‌های کوچک‌تر



شکل ۴- مراحل کلی حذف تصاویر جزئی با روش پیشنهادی بهینه

مطلوب جهت ترکیب به شمار می‌روند، تعداد زیادی از این ۲۵ واریانس به دست آمده صفر شده و مابقی حاوی مقادیر غیر صفر هستند. (۴)

جهت تشخیص و انتخاب تصاویر مناسب جهت ترکیب و در نهایت ایجاد تصویر دودویی برای هر تصویر هفت سطح واریانس تعیین می‌کنیم:

سپس هر ۲۵ مقدار واریانس را در آن سطوح محاسبه می‌کنیم. با انجام این آزمایش این نتیجه حاصل شد که تصاویری جهت ترکیب انتخاب شوند که حاوی سطح صفر بیشتر از ۵۰٪ کل تصویر می‌باشند. به منظور ترکیب تصاویر جزئی انتخاب شده در این مقاله از روش جمع ساده مطابق فرمول ۵ استفاده شده است.

$$\text{Simple Addition: Comb (a,b,c)=a+b+c} \quad (5)$$

برای انتخاب تصاویر مناسب جهت ترکیب، ابتدا هر یک از تصاویر جزئی به تصاویر کوچک‌تر تقسیم می‌شود. در آزمایش‌های انجام شده تصاویر مورد بررسی دارای اندازه‌ی $250 * 250$ پیکسل است. با انجام آزمایش‌های مختلف به این نتیجه رسیدیم که چنانچه هر یک از این تصاویر جزئی به ۲۵ تصویر با اندازه‌ی $50 * 50$ پیکسل تقسیم شود. و واریانس کل هر یک از این تصاویر محاسبه شود، در تصویر جزئی که پراکندگی داده‌ای کم دارند و به نوعی تصاویر مطلوب جهت ترکیب به-شمار می‌روند، تعداد زیادی از این ۲۵ واریانس بدست آمده صفر شده و مابقی حاوی مقادیر غیر صفر هستند. آزمایش‌ها در نرم‌افزار متلب پیاده-سازی شده‌اند. جهت محاسبه‌ی واریانس در این نرم‌افزار از رابطه (۴) استفاده شده است.

در تصاویر جزئی که پراکندگی داده‌ای کم دارند و به نوعی تصاویر

پردازش تصاویر ورق‌های فولادی به منظور آشکارسازی عیوب به کمک موجک گابور

جدول ۱. نمایش واریانس‌های مربوط به تصاویر کوچک‌تر هر یک از تصاویر جزئی در شکل ۴

| D ₁ | | | | | D ₂ | | | | | D ₃ | | | | | D ₄ | | | | |
|----------------|-------|-------|-------|-------|-----------------|-------|-------|-------|-------|-----------------|-------|-------|-------|-------|-----------------|-------|-------|-------|-------|
| 0 | 0.056 | 0 | 0 | 0 | 0.017 | 0.056 | 0.014 | 0.006 | 0.015 | 0 | 0.033 | 0 | 0 | 0 | 0.003 | 0.002 | 0.052 | 0.005 | 0.004 |
| 6 | 6 | 0 | 0 | 0 | 4 | 6 | 4 | 2 | 5 | 4 | 4 | 0 | 0 | 0 | 7 | 1 | 7 | 1 | 4 |
| 0 | 0.033 | 0 | 0 | 0 | 0.006 | 0.002 | 0.006 | 0.002 | 0.012 | 0 | 0.002 | 0 | 0 | 0 | 0.013 | 0.033 | 0.012 | 0.002 | 0.006 |
| 0.027 | 4 | 0 | 0 | 0 | 8 | 1 | 5 | 4 | 7 | 0 | 1 | 0 | 0 | 0 | 9 | 4 | 5 | 1 | 8 |
| 2 | 1 | 0.002 | 0.089 | 0.089 | 6 | 6 | 0.1 | 0.036 | 0.009 | 1 | 0.002 | 0.002 | 0.033 | 0 | 0.073 | 0.002 | 0.051 | 0.033 | 0.001 |
| 0.056 | 0.002 | 2 | 2 | 0 | 0.037 | 0.056 | 0.1 | 0.036 | 0.009 | 1 | 9 | 1 | 4 | 0 | 1 | 9 | 2 | 4 | 8 |
| 6 | 0.056 | 6 | 6 | 0.056 | 6 | 6 | 0.1 | 0.036 | 0.009 | 6 | 0.056 | 0.002 | 0.056 | 0.056 | 0.033 | 0.002 | 0.033 | 0.002 | 0.002 |
| 0.001 | 0.033 | 0.007 | 0.002 | 0 | 6 | 6 | 0.1 | 0.036 | 0.009 | 6 | 1 | 6 | 1 | 6 | 4 | 1 | 4 | 1 | 9 |
| 6 | 6 | 1 | 9 | 0 | 0.033 | 0.056 | 0.041 | 0.019 | 0.005 | 0 | 0.033 | 0.002 | 0.056 | 0 | 0.057 | 0.002 | 0.012 | 0.029 | 0.003 |
| | | | | | 3 | 6 | 2 | 1 | 2 | 5 | 4 | 9 | 6 | 0 | 1 | 5 | 7 | 7 | |
| D ₅ | | | | | D ₆ | | | | | D ₇ | | | | | D ₈ | | | | |
| 0 | 0.001 | 0 | 0 | 0 | 0.001 | 0.002 | 0.002 | 0.002 | 0.002 | 0 | 0.033 | 0 | 0 | 0 | 0.002 | 0.056 | 0.008 | 0.007 | 0.005 |
| 5 | 5 | 0 | 0 | 0 | 6 | 1 | 4 | 7 | 6 | 0 | 4 | 0 | 0 | 0 | 8 | 6 | 5 | 5 | 9 |
| 0 | 0.033 | 0 | 0 | 0 | 0.003 | 0.001 | 0.003 | 0.001 | 0.006 | 0 | 0.002 | 0 | 0 | 0 | 0.016 | 0.002 | 0.059 | 0.011 | 0.015 |
| 0.043 | 4 | 0 | 0 | 0 | 0.003 | 5 | 8 | 8 | 0.006 | 0 | 1 | 0 | 0 | 0 | 5 | 1 | 1 | 3 | 8 |
| 3 | 1 | 0.002 | 0.022 | 0.105 | 0.009 | 0.056 | 0.08 | 0.032 | 0.002 | 0.001 | 0.002 | 0.153 | 0.03 | 0 | 0.033 | 0.056 | 0.002 | 0.009 | 0.01 |
| 0.056 | 0.001 | 8 | 5 | 0 | 9 | 6 | 0.08 | 3 | 3 | 7 | 9 | 4 | 0.03 | 0 | 5 | 6 | 9 | 0.009 | 0.01 |
| 6 | 0.056 | 6 | 6 | 1 | 0.056 | 0.002 | 0.001 | 0.002 | 0.056 | 0.056 | 0.002 | 0.002 | 0.056 | 0.002 | 0.016 | 0.002 | 0.056 | 0.002 | 0.056 |
| 0.001 | 0.002 | 0.001 | 0.002 | 0 | 6 | 1 | 5 | 1 | 6 | 6 | 1 | 9 | 6 | 1 | 6 | 1 | 6 | 1 | 6 |
| 5 | 5 | 5 | 9 | 0 | 0.017 | 0.056 | 0.017 | 0.006 | 0.006 | 0.000 | 0.001 | 0.154 | 0.052 | 0 | 0.018 | 0.056 | 0.005 | 0.003 | 0.002 |
| | | | | | 3 | 6 | 4 | 4 | 9 | 2 | 5 | 8 | 1 | 0 | 7 | 6 | 7 | 8 | 9 |
| D ₉ | | | | | D ₁₀ | | | | | D ₁₁ | | | | | D ₁₂ | | | | |
| 0 | 0.056 | 0 | 0 | 0 | 0.001 | 0.033 | 0.001 | 0.000 | 0.001 | 0 | 0.001 | 0 | 0 | 0 | 0.000 | 0.001 | 0.001 | 0.000 | 0.003 |
| 6 | 6 | 0 | 0 | 0 | 1 | 4 | 7 | 6 | 3 | 0 | 5 | 0 | 0 | 0 | 6 | 5 | 1 | 8 | 9 |
| 0 | 0.033 | 0 | 0 | 0 | 0.001 | 0.002 | 0.002 | 0.001 | 0.001 | 0 | 0.056 | 0.000 | 0 | 0 | 0.004 | 0.002 | 0.012 | 0.003 | 0.003 |
| 0.033 | 4 | 0 | 0 | 0 | 5 | 1 | 0.002 | 2 | 6 | 0 | 6 | 6 | 0 | 0 | 6 | 1 | 0.012 | 2 | 8 |
| 4 | 1 | 0.002 | 0.056 | 0.002 | 0.056 | 0.033 | 0.056 | 0.002 | 0.002 | 0.002 | 0.033 | 0.056 | 0.002 | 0.001 | 0.056 | 0.056 | 0.056 | 0.002 | 0.056 |
| 0.056 | 0.002 | 6 | 9 | 1 | 6 | 4 | 6 | 1 | 9 | 1 | 4 | 6 | 1 | 5 | 6 | 6 | 6 | 1 | 6 |
| 6 | 0.056 | 6 | 6 | 0.000 | 0.003 | 0.002 | 0.003 | 0.005 | 0.001 | 0.000 | 0.001 | 0.120 | 0.022 | 0 | 0.009 | 0.056 | 0.000 | 0.002 | 0.003 |
| 0.001 | 0.033 | 0.001 | 0.001 | 0 | 5 | 9 | 2 | 8 | 1 | 9 | 5 | 5 | 6 | 0 | 3 | 6 | 4 | 3 | 2 |
| 8 | 4 | 0 | 3 | 0 | 0.003 | 0.002 | 0.004 | 0.002 | 0.000 | 0.000 | 0.033 | 0.150 | 0.051 | 0 | 0.002 | 0.002 | 0.001 | 0.001 | 0.000 |
| | | | | | 8 | 1 | 3 | 7 | 8 | 2 | 4 | 7 | 2 | 0 | 0.002 | 1 | 6 | 2 | 3 |

جدول ۲- تقسیم‌بندی ۲۵ واریانس هر یک از تصاویر جزئی جدول ۱ در سطوح مختلف

| Levels | Detail Images | | | | | | | | | | | |
|--------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|-----------------|
| | D ₁ | D ₂ | D ₃ | D ₄ | D ₅ | D ₆ | D ₇ | D ₈ | D ₉ | D ₁₀ | D ₁₁ | D ₁₂ |
| No. of Vars=0 | 14 | 0 | 15 | 0 | 14 | 0 | 14 | 0 | 15 | 0 | 13 | 0 |
| 0<No. of Vars<0.0001 | 1 | 0 | 1 | 0 | 1 | 0 | 2 | 0 | 0 | 0 | 0 | 4 |
| 0.0001<No. of Vars<0.001 | 3 | 4 | 4 | 0 | 3 | 0 | 1 | 4 | 1 | 2 | 3 | 4 |
| 0.001<No. of Vars<0.01 | 3 | 6 | 1 | 7 | 3 | 12 | 3 | 9 | 2 | 14 | 4 | 11 |
| 0.01<No. of Vars<0.1 | 2 | 9 | 2 | 9 | 2 | 4 | 2 | 7 | 2 | 3 | 2 | 1 |
| 0.1<No. of Vars<1 | 1 | 3 | 2 | 6 | 1 | 6 | 2 | 3 | 4 | 4 | 2 | 2 |
| 1<No. of Vars | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 1 | 0 | 1 | 0 | 2 |

معیوب زیاد است. نکته قابل توجه میزان بالای تشخیص درست نواحی سالم در این روش برای کل تصاویر مورد آزمایش است. این روش به خوبی توانسته مانند چشم انسان نواحی سالم را از نواحی معیوب جدا کند و به معنای کامل، روشی بدون نظارت است. در مجموع با ارائه روش تشخیص عیب پیشنهادی، هم سرعت نسبی عملیات تشخیص عیب افزایش یافته و هم درصد تشخیص درست نواحی سالم و معیوب تا حد زیادی نسبت به بقیه روش‌ها افزایش پیدا کرده است و در کل روش مناسبی در تشخیص نواحی معیوب به صورت بدون نظارت می‌باشد. این روش بر روی کلیه تصاویر مورد آزمایش جواب قابل قبولی داشته و چنانچه ویژگی‌های استخراج شده به گونه‌ای باشند که عمل انتخاب تصاویر جزئی را بتوان با چشم انجام داد، این روش نیز قادر خواهد بود مشابه چشم انسان به انتخاب تصاویر جزئی صحیح بپردازد.

پس از ترکیب تصاویر جزئی انتخاب شده با یکدیگر تصویری حاصل می‌شود که به شکل مشخص در برگیرنده‌ی عیوب می‌باشد. پس از انجام این پیش پردازش‌ها عمل تطبیق تصویر دودویی تغییر یافته با تصویر اصلی انجام می‌شود. نتایج حاصل از اجرای این الگوریتم را در جدول شماره ۱ نشان داده‌ایم. همان‌گونه که مشاهده می‌شود این روش به خوبی توانسته بر روی همه‌ی تصاویر به نحو مطلوب عمل کند.

۵. مجموعه‌ی داده و آزمایش‌ها

نتایج بخش‌بندی با روش پیشنهادی بر روی ۵ تصویر در جدول ۱ نشان داده شده است. همان‌گونه که مشاهده می‌شود این روش به خوبی توانسته بر روی همه تصاویر به نحو مطلوب عمل کند. در این روش میزان اشتباه نواحی سالم بسیار کم و در عین حال میزان تشخیص نواحی

به سایر روش‌های موجود هم سرعت نسبی عملیات تشخیص عیب افزایش یافته و هم درصد تشخیص درست نواحی معیوب و درصد تشخیص درست نواحی سالم تا حد زیادی نسبت به سایر روش‌ها افزایش پیدا کرده است و در کل روش مناسبی در تشخیص نواحی معیوب می‌باشد.

در مجموع مزایای روش پیشنهادی و به ویژه نوع خاص آن یعنی روش پیشنهادی بهینه را می‌توان به صورت زیر بیان کرد:

- روش به صورت بدون نظارت است.
- از درصد بالای تشخیص درست نواحی سالم و معیوب برخوردار است.
- برخلاف بیشتر روش‌ها که می‌توانند تنها نوع خاصی از عیوب را از نظر بافت و نوع تشخیص دهند، درصدهای تشخیص عیوب با این روش در مورد همه تصاویر در حد قابل قبولی است.
- از سرعت بالاتری نسبت به روش بخش‌بندی با کلاسه‌بند برخوردار است. زیرا به صورت ناحیه‌ای عمل می‌کند و نه به شکل پیکسلی.

تصویر ۵ نتایج حاصل از انجام روش پیشنهادی بر روی یک تصویر معیوب را نشان می‌دهد.

۶. نتیجه‌گیری

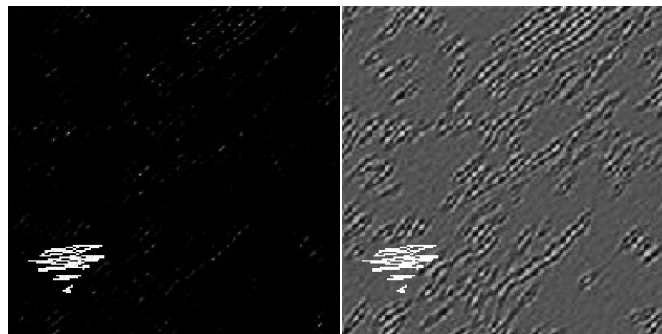
از مزایای این روش می‌توان گفت که در این روش بر خلاف روش‌های موجود هیچ گونه تغییری بر روی مقادیر واریانس‌های بدست آمده داده نمی‌شود و این کاملاً سازگار با بحث تشخیص عیب است. زیرا هر گونه تغییری در مقادیر ممکن است سبب از بین رفتن عیب گردیده و یا سیستم نواحی سالم را به عنوان نواحی معیوب در نظر بگیرد.

این روش بر روی کلیه تصاویر مورد آزمایش جواب قابل قبولی داشته و چنانچه ویژگی‌های استخراج شده به گونه ای باشد که عمل انتخاب تصاویر جزئی را بتوان با چشم انجام داد. این روش نیز قادر است همانند چشم انسان به انتخاب تصاویر جزئی صحیح پردازد.

نتایج بدست آمده از آزمایشات انجام شده هم به صورت درصد و هم به صورت تصویری نشان داده شده است که بیانگر این است که این روش دارای قدرت بالاتری در تشخیص عیب می‌باشد. این روش نسبت

جدول ۳. نتایج مربوط به روش پیشنهادی

| Dataset | Image No | Simple Addition | |
|------------------|----------|----------------------------|-----------------------------|
| | | درصد تشخیص درست نواحی سالم | درصد تشخیص درست نواحی معیوب |
| Modified Dataset | Image1 | 100 | 45.98 |
| | Image2 | 28.95 | 48.99 |
| | Image3 | 100 | 39.98 |
| | Image4 | 100 | 75.98 |
| | Image5 | 100 | 78.98 |



شکل ۵. تصویر سمت راست تصویر دارای عیب ورودی به سیستم و تصویر سمت چپ تصویر بدست آمده با روش پیشنهادی

مراجع

- [1]. Yazdchi, M., Yazdi, M., Golibagh, A., "Steel Surface Defect Detection Using Texture Segmentation Based on Multifractal Dimension", The 1st International Conference on Digital Image Processing (ICDIP 2009), pp. 346-350, 2009, IEEE.
- [2]. Li. Jingting, Ying Wang, Oiang Zhang, Wei Chen, "Method of counting thin steel plates based on digital image processing", Image Analysis and Signal Processing (IASP), 2011, IEEE .
- [3]. M. Sharifzadeh, S. Alirezaee, R. Amirfattahi, S. Sadri. , "Detection of Steel Defect Using the Image Processing Algorithms", Multitopic Conference, INMIC 2008. , Karachi, Pakistan, Dec. 23-24, 2008, IEEE.
- [4]. C.G,Spinola, J.Canero, G. Moreno-Aranda, J.M,Bonelo, M.Martin-Vazquez," Real-time image processing for edge inspection and defect detection in stainless steel production lines ", Imaging Systems and Techniques (IST), 2011, IEEE.
- [5]. M. Sadeghi, M. Shafiee and H. R. Hosseini, "Image processing of steel surfaces For detection of observable defects", 2nd electrical engineering conference, Azad university, Najafabad, 2009.
- [6]. M. Sadeghi, M. Shafiee, F. Memarzadeh Zavareh, A. Mahdeian. "Image Processing of Steel Plates Using 2D Wavelet", International Conference on Computer Science and Network Technology, Harbin, china, Dec. 24-26, 2011, IEEE.
- [7]. M. Sadeghi, M. Shafiee , M. Shafieirad. "A New Approach to Improve Defect Detection of Steel Sheets Using Gabor Wavelet". 3rd International Conference on Signal Processing Systems, Yantai, China. Aug. 27-28, 2010, IEEE.
- [8]. M. Sadeghi, Sh.Valadiesomesaraiet, A. Mahdeiani. " Application of Two Dimensional Wavelet for Defect Detection in Steel Process". 2nd International Conference on Control, Instrumentation, and Automation (ICCIA), Shiraz, I. R. Iran. 2011, IEEE.
- [9]. C.Spinola,. M.J.Cañero-Nieto, J.M.Martin-Vazquez.M.J. Bonelo. " Image processing for surface quality control in stainless steel production lines". Imaging Systems and Techniques (IST), Thessaloniki, 1-2 July ,2010, IEEE.
- [10]. M. Sadeghi, M. Shafiee, F. Memarzadeh Zavareh, M. Memarzadeh Zavareh. " Using image processing in grading tile with Gabor wavelet, International Conference on Computer Science and Network Technology, Changchun, china, Dec. 29-31, 2012, IEEE.

راهنمای نگارش

مقالات فصلنامه علمی - پژوهشی فناوری

اطلاعات و ارتباطات ایران

الف - عنوان مقاله: گویای اصلی‌ترین نکته و یافته مقاله

(طول عنوان نباید از ۱۵ واژه بیشتر باشد).

ب - مشخصات نویسندگان: شامل نام و نام خانوادگی، شغل، وابستگی سازمانی، نشانی پستی و الکترونیکی.

ج - چکیده: متنی حاوی حداکثر ۲۰۰ واژه که بیانگر نکات اصلی مقاله باشد.

د - کلید واژگان: شامل حداکثر ۵ واژه اصلی که بتواند در نمایه‌گذاری مقاله، ویژگی‌های آن را توصیف کند.

ه - بدنه مقاله: شامل مقدمه، تبیین مسئله اصلی تحقیق، روش کاربردی برای حل مسئله و نتایج آن. این بخش باید چنان کامل باشد که خواننده با مطالعه آن به اقناع علمی دست یابد از سویی لازم است نگارندگان از تطویل کلام بپرهیزند و بخشهای غیر ضروری را از مقاله حذف کنند.

و - مراجع: شامل همه کتابها، مقاله‌ها، گزارشها و صفحه‌های وب مورد استناد در مقاله.

یادآوری ۱: شکلها، جدولها و مراجع باید به ترتیبی که در متن آمده‌اند شماره‌گذاری شود. (عنوان شکل در زیر آن و عنوان جدول در بالای آن ذکر می‌شود).

یادآوری ۲: شیوه نگارش مراجع و ارجاع به آن در متن، مطابق استاندارد نشریات IEEE است.

یادآوری ۳: در مقاله‌های فارسی لازم است همه اصطلاحها، اسامی و اعداد کاربردی در متن مقاله، جدولها و شکلها به زبان فارسی نوشته شود و در صورت نیاز معادل فارسی آن در زیرنویس صفحه آورده شود. مبنای انتخاب معادل فارسی، مصوبات فرهنگستان زبان و ادب فارسی است.

یادآوری ۴: شیوه نگارش فارسی فصلنامه براساس «دستور خط فارسی» مصوب فرهنگستان زبان و ادب فارسی است و نگارندگان لازم است اصول نگارش آن را در مقاله رعایت کنند.

فصلنامه علمی-پژوهشی «فناوری اطلاعات و ارتباطات ایران»، نشریه‌ای علمی است که با هدف ایجاد ارتباط میان پژوهشگران و اندیشمندان عرصه فناوری اطلاعات و اعتلای سطح دانش نظری این حوزه علمی منتشر می‌شود. این فصلنامه در پی آن است که زمینه‌ای مناسب را برای ارائه آخرین یافته‌های علمی محققان در حوزه‌های مختلف فناوری اطلاعات در داخل و خارج از ایران پدید آورد.

فصلنامه پذیرای مقاله‌هایی است که حاصل پژوهشهای اصیل نظری، کاربردی و توسعه‌ای در زمینه‌های تخصصی فناوری اطلاعات و فناوری ارتباطات باشد، از این رو از همه پژوهشگران و متخصصان دعوت می‌شود تا نتایج تحقیقات خود را برای انتشار به فصلنامه ارسال کنند. در این خصوص اشاره به نکات زیر ضروری است.

۱- مقاله‌های ارسالی به فصلنامه باید دربرگیرنده نتایج نو در هر یک از حوزه‌های تخصصی فصلنامه باشد و نباید قبلاً در هیچ نشریه دیگری به چاپ رسیده یا به طور همزمان به سایر نشریه‌های داخلی و خارجی برای داوری ارسال شده باشد. شایان ذکر است انتشار مقاله کاملی که خلاصه نتایج آن پیشتر در مجامع علمی ارائه و چاپ شده باشد، بلامانع است.

۲- مقاله‌های ارسالی به فصلنامه باید حاوی مطالب جدید و واجد نتایج، روشها، مفاهیم، کاربردها یا ترکیبی از این موارد باشد.

۳- فصلنامه از مقاله‌های مروری، تحلیلی و توصیفی (با دعوت از صاحب نظران برجسته) و نیز نامه به سردبیر (در نقد، تحلیل و تشریح مقاله‌ای خاص) استقبال می‌کند.

۴- مقاله‌ها باید بتوانند اطلاعات کافی را به خواننده منتقل کنند، از این رو لازم است حاوی بخشهای زیر باشند:

(Corresponding Author) درمقاله ارسالی معلوم

- باشد تا امکان ایجاد ارتباط با نگارندگان فراهم آید.
- ۷- حق نشر مقاله‌ها برای فصلنامه محفوظ است؛ از این رو ضروری است نویسنده عهده‌دار مکاتبات به نمایندگی از همه نویسندگان «مجوز حق انتقال نشر» را امضا کند، در غیر این صورت فصلنامه از چاپ مقاله پذیرفته شده معذور است.
- ۸- هر مقاله پس از بررسی سردبیر به سه داور ارسال می‌شود. جمع‌بندی نظریات داوران برعهده سردبیر فصلنامه است. فصلنامه حق پذیرش، ویرایش یا رد مقاله‌ها را برای خود محفوظ می‌دارد.
- ۹- فصلنامه متعهد می‌شود در ازای چاپ هر مقاله، تعداد پنج نسخه از بازچاپ مقاله (یا پنج نسخه از مجله) را برای نویسنده عهده‌دار مکاتبات ارسال کند. ضمناً دسترسی به نسخه الکترونیکی مقاله برای همه نویسندگان مقاله امکان‌پذیر خواهد بود.

۵- مقاله‌های دریافتی به هر دو زبان فارسی و انگلیسی قابل پذیرش است. فصلنامه شماره‌های ویژه‌ای را به مقاله‌های انگلیسی زبان اختصاص خواهد داد. به هر حال همه مقاله‌ها (خواه فارسی، خواه انگلیسی) باید دارای چکیده دو زبانه باشد.

۶- لازم است نگارندگان مقاله خود را در محیط MS-Word به صورت دو ستونه، با فاصله خطوط ۱ (Single) و با قلم بی‌لوتوس (B Lotus) ۱۱ (برای مقاله فارسی) و Times New Roman 11 (برای مقاله انگلیسی) و با رعایت حاشیه بالا ۳ سانتی‌متر، پایین ۱ سانتی‌متر، چپ ۲.۵ سانتی‌متر، راست ۲.۵ سانتی‌متر و با توجه به نکات بند ۴، آماده و به دفتر نشریه ارسال کنند. (در ضمن در صورت امکان، PDF آن را نیز ارسال شود).

یادآوری: ضروری است نشانی کامل پستی و الکترونیکی و نیز شماره تلفن نویسنده عهده‌دار مکاتبات

معرفی انجمن فناوری اطلاعات و ارتباطات ایران

اهداف

- انجمن فناوری اطلاعات و ارتباطات ایران در سال ۱۳۸۳ با اهداف زیر تشکیل گردید:
- ایجاد ارتباط با پژوهشگران و افراد خبره در حوزه‌های مرتبط با فناوری اطلاعات (IT)، فناوری ارتباطات (CT) و فراهم نمودن زمینه‌های شناسایی و همکاری مشترک بین آنها.
 - همکاری با نهادهای اجرائی، علمی و پژوهشی مرتبط با فناوری اطلاعات و ارتباطات، اقدام در جهت پیوند و همکاری دستگاههای اجرائی با مجموعه‌های علمی - تحقیقاتی.
 - توسعه فرهنگ استفاده مطلوب از توانمندیهای فناوری اطلاعات و ارتباطات.
 - ارزیابی و تحلیل عوامل رشد کشورهای مشابه دیگر و استفاده مطلوب از تجربیات دیگران در توسعه کشور درحوزه ارتباطات و فناوری اطلاعات.

برنامه‌ها

- انجمن به منظور تحقق اهداف خود، برنامه‌های زیر را در اولویت فعالیتهایش قرار داده است.
- ایجاد و تقویت ارتباط علمی، پژوهشی و آموزشی در سطوح ملی، منطقه‌ای و بین‌المللی.
 - برگزاری همایشهای علمی در سطح ملی، منطقه‌ای و بین‌المللی.
 - ایجاد تسهیلات برای نشر نتایج آخرین دستاوردهای علمی و پژوهشی مرتبط با نیازهای کشور درحوزه ارتباطات و فناوری اطلاعات (انتشار کتب و نشریات علمی).
 - برگزاری کارگاههای تخصصی درحوزه ارتباطات و فناوری اطلاعات.

برنامه‌ریزی فعالیتهای

- انتشار خبرنامه انجمن
- تشکیل کمیته‌های تخصصی به منظور تحقق اهداف انجمن
- اقدامات وسیع برای معرفی انجمن در همایشهای علمی، سازمانهای اجرائی و دانشگاهها
- انتشار نشریه علمی - پژوهشی ICT
- همکاری در برگزاری همایشهای علمی از قبیل چهارمین کنفرانس فناوری اطلاعات و دانش (IKT)
- تشکیل جلسات با مسئولین مملکتی ذیربط جهت شناساندن اهداف انجمن و مشخص شدن نحوه همکاری

کمیته‌های تخصصی انجمن

- کمیته پذیرش و اطلاع‌رسانی
- کمیته آموزش
- کمیته پژوهش
- کمیته مالی و پشتیبانی
- کمیته ارتباط صنعت دانشگاه
- کمیته جامعه اطلاعاتی

از کلیه اندیشمندان، اساتید دانشگاه، پژوهشگران و سایر علاقه‌مندان در حوزه ارتباطات و فناوری اطلاعات دعوت می‌شود با عضویت در انجمن، ما را در تحقق اهداف والای آن یاری نمایند.

نشانی دبیرخانه: تهران - خیابان حافظ - دانشگاه صنعتی امیرکبیر - ساختمان ابوریحان - طبقه ششم - اتاق ۶۱۲

تلفن: ۰۲۱-۶۶۴۸۵۸۵۶، تلفکس: ۰۲۱-۶۶۴۹۵۴۳۳

هزینه اشتراک

| | | |
|----------------------|------------------------|--------------------------|
| هزینه اشتراک یکساله: | پست عادی: ۱۲.۰۰۰ تومان | پست پیشتاز: ۱۶.۰۰۰ تومان |
| هزینه اشتراک دوساله: | پست عادی: ۲۰.۰۰۰ تومان | پست پیشتاز: ۲۸.۰۰۰ تومان |

اطلاعات متقاضی

• حقیقی:

نام و نام خانوادگی: سمت:

رشته و مقطع تحصیلی: نام محل تحصیل:

• حقوقی:

نام شرکت / سازمان: نام متقاضی:

سمت:

تعدد نسخه درخواستی از هر شماره:

نسخه شروع اشتراک از شماره:

اشتراک جدید تمدید اشتراک

(شماره اشتراک قبلی:))

نشانی

آدرس:

استان: شهر: کدپستی:

تلفن: همراه: پست الکترونیک:

از علاقه‌مندان به اشتراک فصلنامه فناوری اطلاعات و ارتباطات ایران درخواست می‌نماییم برگ اشتراک را به دقت و با خط خوانا تکمیل نموده و وجه اشتراک را براساس تعرفه، به شماره حساب ۲۶۶۱۱۱۴۲۸ بانک تجارت، شعبه امیراکرم، کد ۲۶۶۱ باجه دانشگاه امیرکبیر، به نام انجمن فناوری اطلاعات و ارتباطات ایران واریز نموده و اصل فیش بانکی و فرم اشتراک را به یکی از دو شیوه زیر برای ما ارسال نمایند.

۱- از طریق دورنگار با شماره ۶۶۴۹۵۴۳۳-۰۲۱

۲- از طریق پست به آدرس: تهران، دانشگاه صنعتی امیرکبیر، ساختمان ابوریحان، طبقه ششم، اتاق ۶۱۲

- اعضای انجمن از ۲۰ درصد تخفیف برخوردار می‌شوند.
- خواهشمند است، مشترکان محترم در صورت تغییر نشانی، امور مشترکین را مطلع فرمایند.
- باتوجه به اینکه تنها عدم وصول مجلاتی که به صورت پیشتاز ارسال می‌شوند قابل پیگیری است لذا توصیه می‌شود از خدمات پست پیشتاز استفاده شود.



«فرم درخواست عضویت حقیقی»

لطفاً در این قسمت چیزی ننویسید

شماره عضویت:

۱- مشخصات فردی:

نام خانوادگی:

نام:

نام خانوادگی (به انگلیسی):

نام (به انگلیسی):

روز ماه سال

جنس: مرد زن تاریخ تولد:

شماره شناسنامه: محل تولد:

شماره ملی:

سمت:

۲- سوابق تحصیلی:

آخرین مدرک تحصیلی: دکتری کارشناسی ارشد سایر (نام ببرید):

رشته تحصیلی: فناوری اطلاعات مهندسی برق مهندسی کامپیوتر مهندسی صنایع علوم اجتماعی اطلاع‌رسانی و کتابداری سایر رشته‌های مرتبط (نام ببرید):

۴- نوع عضویت:

پیوسته وابسته دانشجویی

۵- حق عضویت:

حق عضویت سالانه

| نوع عضویت | پیوسته | وابسته | دانشجویی |
|--|--------------|--------------|-------------|
| مبلغ | ۱۰۰.۰۰۰ ریال | ۱۰۰.۰۰۰ ریال | ۵۰.۰۰۰ ریال |
| شماره حساب: ۲۶۶۱۱۱۴۲۸ بانک تجارت، شعبه امیراکرم، کد ۲۶۶۱ باجه دانشگاه امیرکبیر | | | |

۶- نشانی محل کار (تحصیل):

تلفن ثابت: تلفن همراه: نامبر:

آدرس پست الکترونیکی: آدرس سایت:

امضا:

تاریخ:

* عضویت پیوسته: مؤسسان انجمن و کلیه افرادی که حداقل دارای درجه کارشناسی ارشد در یکی از رشته‌های فناوری اطلاعات، مهندسی برق، علوم و مهندسی کامپیوتر، مهندسی صنایع، علوم اجتماعی، اطلاع‌رسانی و کتابداری و سایر رشته‌های مرتبط باشند می‌توانند به عضویت پیوسته انجمن درآیند.

* عضویت وابسته: اشخاصی که دارای درجه کارشناسی هستند و مدت ۵ سال به نحوری در یکی از رشته‌های مذکور در بند فوق (عضویت پیوسته) شاغل باشند.

* عضویت دانشجویی: کلیه دانشجویانی که در رشته‌های مندرج شده در بند فوق (عضویت پیوسته) به تحصیل اشتغال دارند.

«فرم درخواست عضویت حقوقی»

لطفاً در این قسمت چیزی ننویسید

شماره عضویت:

| | | | | | | | | | |
|--|--|--|--|--|--|--|--|--|--|
| | | | | | | | | | |
|--|--|--|--|--|--|--|--|--|--|

انجمن فناوری اطلاعات و ارتباطات ایران انجمنی علمی است که به منظور گسترش و پیشبرد و ارتقای علم و توسعه کمی و کیفی نیروهای متخصص و بهبود بخشیدن به امور آموزشی و پژوهشی در زمینه‌های مربوط به ارتباطات و فناوری اطلاعات تشکیل شده است. مؤسساتی که در زمینه‌های علمی و پژوهشی ارتباطات و فناوری اطلاعات فعالیت دارند می‌توانند با تکمیل فرم، درخواست عضویت حقوقی خود را اعلام نمایند.

۱- مشخصات مؤسسه:

نام دانشگاه/شرکت/سازمان: تعداد کارکنان:
نوع فعالیت:

۲- حق عضویت:

مبلغ پیشنهادی:
حق عضویت سالانه برای عضویت حقوقی حداقل یک میلیون ریال می‌باشد که مؤسسات دارای امکانات مالی مناسب می‌توانند مبلغ بیشتری جهت توسعه انجمن پرداخت نمایند (شماره حساب: حساب جاری ۲۶۶۱۱۱۴۲۸ بانک تجارت، شعبه امیراکرم، کد ۲۶۶۱ باجه دانشگاه امیرکبیر).

۳- نشانی محل مؤسسه:

تلفن ثابت: تلفن همراه:
آدرس پست الکترونیکی: آدرس سایت:
نمابر:

۴- مشخصات رئیس / مدیرعامل:

نام خانوادگی: نام:
نام خانوادگی (به انگلیسی): نام (به انگلیسی):
سمت:

امضا:

تاریخ:

| اعضای حقوقی انجمن در سال ۹۱ | |
|--|---|
| <ul style="list-style-type: none"> • مدیریت مخابرات شهرستان بابل سید محمد میری هر یکند شماره عضویت: ۹۱۴۱۲۵۱۸۰ | <ul style="list-style-type: none"> • مدیریت مخابرات آمل حمیدرضا سلطانی شماره عضویت: ۹۱۴۱۲۵۱۷۹ |
| <ul style="list-style-type: none"> • شرکت فنی مهندسی ضریب توان چاوش خسرو طلوعی شماره عضویت: ۹۱۴۱۷۵۱۸۲ | <ul style="list-style-type: none"> • مدیریت مخابرات شهرستان قائم شهر سید ذکریا اسدنیا شماره عضویت: ۹۱۴۱۲۵۱۸۱ |
| | <ul style="list-style-type: none"> • شرکت ارتباطات سیار ایران همراه اول وحید صدوقی شماره عضویت: ۹۱۴۲۱۵۱۸۳ |
| اعضای پیوسته انجمن در سال ۹۱ | |
| <ul style="list-style-type: none"> محمد شرفی - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۷۶ | <ul style="list-style-type: none"> کامران عبداللهی - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۷۵ |
| <ul style="list-style-type: none"> محسن مهدی‌زاده - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۷۸ | <ul style="list-style-type: none"> انور محمدی - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۷۷ |
| <ul style="list-style-type: none"> محمد پرویزی - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۸۰ | <ul style="list-style-type: none"> میترا دهقان مهابادی - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۷۹ |
| <ul style="list-style-type: none"> مریم قربان‌علیزاده - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۸۲ | <ul style="list-style-type: none"> محمد ابراهیمی نسب - بانک دی شماره عضویت: ۹۱۶۲۱۶۶۸۱ |
| <ul style="list-style-type: none"> داوود ادیب - بانک دی شماره عضویت: ۹۱۵۷۱۵۶۸۳ | <ul style="list-style-type: none"> انوشیروان مرآت - شرکت صنایع ارتباطی آوا شماره عضویت: ۹۱۵۷۱۵۶۸۴ |
| <ul style="list-style-type: none"> سید جلیل حسین تبار - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۸۶ | <ul style="list-style-type: none"> بهروز باباجانی شانی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۸۵ |
| <ul style="list-style-type: none"> باباعلی جعفرپور - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۸۸ | <ul style="list-style-type: none"> علی نوروزی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۸۷ |
| <ul style="list-style-type: none"> میترا شاهرخی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۰ | <ul style="list-style-type: none"> فرشته صالحیان - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۸۹ |
| <ul style="list-style-type: none"> حرمت ا... فرهنگ - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۲ | <ul style="list-style-type: none"> رضا شعبانی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۱ |
| <ul style="list-style-type: none"> محمد اسماعیل زاهدی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۴ | <ul style="list-style-type: none"> سیدعلی مظلومی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۳ |
| <ul style="list-style-type: none"> محسن سفیدگر نیا - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۶ | <ul style="list-style-type: none"> فرهاد اسحاقی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۵ |
| <ul style="list-style-type: none"> محمد مهدی صنعتیان - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۸ | <ul style="list-style-type: none"> مهدی اسماعیل زاده - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۷ |
| <ul style="list-style-type: none"> سید مجتبی سیف اله پور - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۷۰۰ | <ul style="list-style-type: none"> علی اخضری - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۶۹۹ |

| | |
|---|--|
| علیرضا رحیمی گرجی - شرکت مخابرات مازندران شماره عضویت: ۹۱۱۱۵۱۷۰۱ | روشن احمدی - دانشگاه تربیت دبیر شهید رجایی تهران شماره عضویت: ۹۱۱۲۱۳۷۰۲ |
| مهدی منصوری زاده شماره عضویت: ۹۱۱۷۱۳۷۰۳ | |
| اعضای وابسته انجمن در سال ۹۱ | |
| حسین گودرزی - شرکت رایتل شماره عضویت: ۹۱۵۲۱۵۳۴۷ | سعید صفوی شماره عضویت: ۹۱۵۲۱۵۳۴۸ |
| محمد عابدیان جلودار - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۴۹ | حسن حسینی - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۰ |
| داود کاکه زاده - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۱ | کمال نیک صیرت پهمدانی - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۲ |
| سیده فائزه حسینی - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۳ | احسان وارث اصفهانی - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۴ |
| محمد لریان فرد - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۵ | حمیدرضا حسینی - بانک دی شماره عضویت: ۹۱۶۲۱۶۳۵۶ |
| پیام شعبانیان - شرکت صنایع ارتباطی آوا شماره عضویت: ۹۱۵۷۱۶۳۵۷ | محسن ناصری - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۵۸ |
| حسن قزبانگلی گنجی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۵۹ | محسن قنبر زاده - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۰ |
| عباس عروجی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۱ | اسماعیل زرین فر - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۲ |
| سیدعلی اکبر جلیلی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۳ | ایرج آقاچان نژاد - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۴ |
| هادی بریمانی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۵ | حسن احمد زاده - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۶ |
| عباس نصیری - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۷ | آرزم یگانه فلاح - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۸ |
| امید آبچر - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۶۹ | سید محمد قاضی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۰ |
| سید جواد رسولی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۱ | حسن شیرزاد - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۲ |
| ذبیح ا... سروش - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۳ | مصطفی فلاحتیان - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۴ |
| شعبانعلی کاظمی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۵ | یعقوب سگرددون - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۶ |
| محسن توسلی - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۷ | سید مقصود صالحی نژاد - شرکت مخابرات مازندران شماره عضویت: ۹۱۲۱۵۱۳۷۸ |

| | |
|---|---|
| مهدي خاكپور - شركت مخابرات مازندران شماره عضويت: ۹۱۲۱۵۱۳۷۹ | مصطفی مفیدیان - شرکت مخابرات مازندران شماره عضويت: ۹۱۲۱۵۱۳۸۰ |
| جلال شجاعی - شرکت مخابرات مازندران شماره عضويت: ۹۱۲۱۵۱۳۸۱ | ايوب عباسی فیروزجاه - شرکت مخابرات مازندران شماره عضويت: ۹۱۲۱۵۱۳۸۲ |
| مريم شفيعی - دانشگاه تهران شماره عضويت: ۹۱۶۲۱۶۳۸۳ | |
| اعضای دانشجویی انجمن در سال ۹۱ | |
| علیرضا مجید - دانشگاه آزاد اسلامی واحد تهران جنوب شماره عضويت: ۹۱۳۲۱۳۶۲۴ | حامد عربزاده - دانشگاه امیرکبیر شماره عضويت: ۹۱۳۲۱۳۶۲۵ |
| مهديه پاک طينت - دانشگاه علم و صنعت شماره عضويت: ۹۱۳۲۱۳۶۲۶ | فرزانه خسروی - دانشگاه صنعتی امیرکبیر شماره عضويت: ۹۱۳۲۱۳۶۲۷ |
| آيدا موحدنيا - دانشگاه صنعتی امیرکبیر شماره عضويت: ۹۱۳۲۱۳۶۲۸ | رضا شفيعی - دانشگاه صنعتی امیرکبیر شماره عضويت: ۹۱۳۲۱۳۶۲۹ |
| ليلا قنبري عربلو - دانشگاه صنعتی امیرکبیر شماره عضويت: ۹۱۳۲۱۳۶۳۰ | محمدحسن محمدپور - دانشگاه فرهنگیان پردیس شماره عضويت: ۹۱۳۷۶۳۶۳۱ |

Image Processing of steel sheets for Defect Detection by using Gabor Wavelet

M. Sadeghi^{*1} M. Shafiee^{}**

^{*} MSc Student, Young Researchers and Elite Clube, Islamic Azad University Zavareh Branch, Isfahan

^{**} Professor, Department of Electrical Engineering, Amirkabir University of Technology, Tehran

Abstract

In different steps of steel production, various defects appear on the surface of the sheet. Putting aside the causes of defects, precise identification of their kinds helps classify steel sheet correctly, thereby it allocates a high percentage of quality control process. QC of steel sheet for promotion of product quality and maintaining the competitive market is of great importance. In this paper, in addition to quick review of image process techniques used, using image process by means of Gabor wavelet, a fast and precise solution for detection of texture defects in steel sheet is presented. In first step, the approach extracts considerable texture specification from image by using Gabor wavelet. The specification includes both different directions and different frequencies. Then using statistical methods, images are selected that have more obvious defects, and location of defects is determined. Offering the experimental samples, the accuracy and speed of the method is indicated.

Keywords

Cell counting, Color segmentation, CIElab color space, Fuzzy c-means, Immunohistochemistry, K-means, Meningioma, Mitosis index, Ultra erosion

1. Corresponding author (msadeghi@khuisf.ac.ir)

Medial-axis Enhancement of Tubular Structures and its Application in the Extraction of Portal Veins

A.H. Foruzan^{*1} R. A. Zoroofi^{} Y. Satob^{***} M. Horic^{***}**

^{*} Assistant Professor, Department of Biomedical Engineering, Shahed University, Tehran

^{**} Professor, Department of Electrical and Computer Engineering, University of Tehran,
Tehran

^{***} Associate Professor, Department of Medicine and Radiology, Osaka University,
Japan

Abstract

In this paper, a new filter is designed to enhance medial-axis of tubular structures. Based on a multi-scale method and using eigenvectors of Hessian matrix, the distance of a point to the edges of the tube is found. To do this, a hypothetical line with a deliberate direction is passed through the point which cuts the tube at its edges. For points which are located on the medial-axis, this distance is symmetric with respect to any deliberate direction. We find samples of the distances in different directions and assign a measure to the points based on this symmetry property. The output of this step is an enhanced image in which noise is removed and tubes can be seen more clearly. Then, we employ the filter developed by Pock et al. to enhance medial axis. Evaluation of the proposed method is performed using 2D/3D synthetic/clinical datasets both quantitatively and qualitatively.

Keywords

Extraction of the tubular structures, Extracted from the central axis of the portal vein, Hepatic CT image analysis, Medical image processing.

1. Corresponding author (a.foruzan@shahed.ac.ir)

Outdoor Color Scene Segmentation towards Object Detection using Dual-Resolution Histograms

J. Rasti^{*1}

A. H. Monadjemi^{**}

A. Vafaei^{***}

^{*}Assistant Professor, Department of Biomedical Engineering, University of Isfahan,
Isfahan

^{**} Associate Professor, Department of Computer Engineering, University of Isfahan,
Isfahan

^{***}Assistant Professor, Department of Computer Engineering, University of Isfahan,
Isfahan

Abstract

One of the most important problems in automatic outdoor scene analysis is the approach of segmentation towards object detection. The special characteristics of such images -like color variety, different luminance effects and color shades, abundant texture details, and diversity of objects- lead to major challenges in the segmentation process. In the previous research, we proposed a k-means clustering algorithm in a multi-resolution platform for preliminary color segmentation. In this method, the texture details are deliberately expunged and apparent clusters are gradually removed in the blurred versions of the image to let more detailed classes expose in the more clarified versions. The performance of this step-by-step approach is relatively higher than the traditional k-means in color clustering for outdoor scene segmentation. In this paper, an adaptive method based on the circular hue histogram in a dual-resolution platform is suggested to detect the apparent clusters in the blurred images. Experimental results on two outdoor datasets show about 20% decrease in the pixel segmentation error as well as around 30% increase in both precision and speed in the convergence of the clustering algorithm.

Keywords

External image, Clustering, Segmentation, Color resolution.

1. Corresponding author (rasti@eng.ui.ac.ir)

SPIHT-DCT hybrid method using Placement spatial scalability and Place - when encoding a video HDTV

V. Seirafian*¹

S. Talebi **

* MSc Student, Department of Electrical Engineering, Islamic Azad University Maimeh Branch, Isfahan

** MSc Student, Electrical Engineering Department, University of Bahonar, Kerman

Abstract

This paper provided a hybrid coding using spatial and spatial - temporal scalability terms, for coding video with high resolution. In the hybrid method, Intra and Inter video frames are encoded in two different ways. Intra frames are coded using SPIHT algorithm based on wavelet transform and Inter frames are coded using conventional MPEG-2 standard, based on DCT transform. With using this procedure, video is sent through two or three layers. The transmitted data from layers provided the video with different resolution and quality. There for, the user can choose the appropriate services based on their needs. In the spatial scalability, the base layer and enhanced layer have the same coding structure. But in the spatial - temporal scale when the second-upgraded-layer, because it contains only Inter frame, just coded based on standard methods. Simulation results of various videos with high resolution show the image quality of the proposed method in comparison of scalability in multiple layers in MPEG-2 standards.

Keywords

MPEG-2 standard, SPIHT, DCT, Scalability, DHTV

1. Corresponding author (vahid.seirafian@iaumeymeh.ac.ir)

Determination of Optimum SVMs Based on Genetic Algorithm in Classification of Hyper spectral Imagery

F. Samadzadegan *

H. Hasani¹**

* Professor, Department of Geomatic, Faculty of Engineering, University of Tehran, Tehran

** PhD Student at Department of Geomatic, Faculty of Engineering, University of Tehran, Tehran

Abstract

Hyper spectral remote sensing imagery, due to its rich source of spectral information provides an efficient tool for ground classifications in complex geographical areas with similar classes. Referring to robustness of Support Vector Machines (SVMs) in high dimensional space, they are efficient tool for classification of hyper spectral imagery. However, there are two optimization issues which strongly effect on the SVMs performance: Optimum SVMs parameters determination and optimum feature subset selection. Traditional optimization algorithms are appropriate in limited search space but they usually trap in local optimum in high dimensional space, therefore it is inevitable to apply meta-heuristic optimization algorithms such as Genetic Algorithm to obtain global optimum solution. This paper evaluates the potential of different proposed optimization scenarios in determining of SVMs parameters and feature subset selection based on Genetic Algorithm (GA). Obtained results on AVIRIS Hyper spectral imagery demonstrate superior performance of SVMs achieved by simultaneously optimization of SVMs parameters and input feature subset. In Gaussian and Polynomial kernels, the classification accuracy improves by about 5% and 15% respectively and more than 90 redundant bands are eliminated. For comparison, the evaluation is also performed by applying it to Simulated Annealing (SA) that shows a better performance of Genetic Algorithm especially in complex search space where parameter determination and feature selection are solve simultaneously.

Keywords

Support vector machines, Hyper spectral images, Class classification, Model selection, Feature selection, Genetic algorithms.

1. Corresponding author (hasani@ut.ac.ir)

Improvement of mean shift tracker for tracking of target with variable photometric pattern

P. Moallem^{*1} J. Abbaspour^{**} Ar. Memarmoghaddam^{***} M. KavoshTehrani^{****}

^{*}Associate Professor, Department of Electrical Engineering, University of Isfahan,
Isfahan

^{**}M.Sc in Electro-Optical Engineering, Research Institute of Optoelectronics,
Malek-ashtar University of Technology, Tehran

^{***}PhD Candidate in Electronics Engineering, Department of Electrical Engineering,
University of Isfahan, Isfahan

^{****}Associate Professor, Department of Physics, Malek-ashtar University of
Technology, Shahin-shahr

Abstract

The mean shift algorithm is one of the popular methods in visual tracking for non-rigid moving targets. Basically, it is able to locate repeatedly the central mode of a desirable target. Object representation in mean shift algorithm is based on its feature histogram within a non-oriented individual kernel mask. Truly, adjusting of the kernel scale is the most critical challenge in this method. Up to now, no methods are presented that can perfectly as well as efficiently adjust and adapt the kernel scale during track when a target is resized. Another problem of mean shift tracking algorithm will be encountered whenever photometric properties of target texture changes. In order to solve these problems, this paper presents a modified mean shift tracking algorithm that is used a robust adaptive sizing technique. It can also cope with photometric changes of target template by adapting of its model in every frame of image sequence. In our proposed method, at first, the target window is adaptively resized with respect to spatio-temporal gradient powers of its pixel intensities in current frame and then mean shift algorithm is consequently applied to the resulted sizing window. Compared to standard mean shift algorithm, experimental results show that our proposed method, not only reduces center location errors of target, but also efficiently track it in the presence of changing illumination.

Keywords

Mean shift tracker, Histogram model, Kernel, Tracking window, Adaptive sizing, Contrast.

1. Corresponding author (p_moallem@eng.ui.ac.i)

Contents

**Iranian Journal of
Information Technology & Communication**
No. 13-14, Vol.4 , October 2012 - March 2013

| | |
|--|----|
| ▪ Improvement of Mean Shift Tracker for Tracking of Target with Variable Photometric Pattern | 1 |
| P. Moallem, J. Abbaspour, A.R. Memarmoghaddam, M. KavoshTehrani | |
| ▪ Determination of Optimum SVMs Based on Genetic Algorithm in Classification of Hyperspectral Imagery | 9 |
| F. Samadzadegan, H. Hasani | |
| ▪ SPIHT-DCT hybrid method using Placement spatial scalability andPlace - when encoding a video HDTV | 25 |
| V.R. Seirafian, S. Talebi | |
| ▪ Outdoor Color Scene Segmentation towards Object Detection using Dual-Resolution Histograms | 37 |
| J. Rasti, A.H. Monadjemi, A. Vafaei | |
| ▪ Medial-axis Enhancement of Tubular Structures and its Application in the Extraction of Portal Veins | 57 |
| A. H. Foruzan, R. A. Zoroofi, Y. Satob, M. Horic | |
| ▪ Image Processing of steel sheets for Defect Detection by using Gabor Wavelet | 67 |
| M. Sadeghi, M. Shafiee | |
| Instruction for Authors | 75 |
| Abstracts in English | |
