

کاهش بعد تصاویر فراطیفی از طریق خوشه بندی فازی باندها

برات مجردی^۱، محمد جواد ولدان زوج^{۲*} و حمید ابریشمی مقدم^۳

^۱ دانشجوی دکتری دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیر

^۲ دانشیار دانشکده مهندسی نقشه برداری - دانشگاه صنعتی خواجه نصیر

^۳ عضو هیئت علمی دانشگاه شهید رجایی

(تاریخ دریافت ۸۶/۹/۲۰، تاریخ دریافت روایت اصلاح شده ۸۷/۸/۱۸، تاریخ تصویب ۸۷/۱۰/۳۰)

چکیده

این مقاله یک روش نوین جهت انتخاب باند از تصاویر فراطیفی از طریق خوشه بندی باندها ارائه می دهد. نوآوری اصلی این تحقیق در دو موضوع قرار می گیرد: الف- ارائه یک فضای محاسباتی جدید با نام فضای پدیده که در آن باندها بر اساس انعکاس طیفی پدیده ها دارای بردار مشخصه می شود. ب- ارائه معیار هایی نظیر عدم قطعیت و زاویه در فضای پدیده برای شناسایی باندهای با وابستگی بالا و باندهای حاوی اطلاعات. پس از آنکه فضای پدیده توسط میانگین کلاسها ایجاد گردید، باندها در این فضا توسط الگوریتم FCM خوشه بندی می شوند. مجموعه باندهای با همبستگی بالا از طریق شاخص عدم قطعیت در یک دسته قرار گرفته و نزدیک ترین باند به مرکز هر دسته به عنوان نماینده باندهای قرار گرفته در آن دسته معرفی می شود. از طرفی باندهای با عدم قطعیت بالا به عنوان باندهای منفرد معرفی شده که از میان آنها باندهای حاوی اطلاعات از طریق زاویه نسبت به قطر فوق مکعب فضای پدیده شناسایی می گردند. از آن جا که دسته بندی باندها مبتنی بر الگوریتم خوشه بندی فازی و نظارت نشده است عمل خوشه بندی چندین بار تکرار شده و باندهایی به عنوان نماینده در فضای پدیده معرفی می شوند که بیشترین صحت طبقه بندی را به ازای داده های اعتبار سنجی حاصل نمایند. نتایج عملی بر روی یک قطع از تصویر فراطیفی که به عنوان یک داده چالش آور و مبنا مطرح است نتایج بهتری را نسبت به الگوریتم های متداول انتخاب باند نظیر پیشرو شناور و پسر و شناور عاید کرد.

واژه های کلیدی: تصاویر فراطیفی، انتخاب باند، طبقه بندی، فضای پدیده

مقدمه

بندی، داده های آموزشی به اندازه کافی زیاد نباشند. نحوه بروز این مشکل به این ترتیب می باشد که با افزایش تعداد باندها صحت طبقه بندی تا یک حدی افزایش یافته سپس شروع به کاهش می کند این پدیده به نام پدیده Hughes معروف می باشد [۳]. این امر در حالت هایی که طبقه بندی کننده غیر پارامتریک باشد وخیم تر نیز می گردد [۴]. از این رو کاهش ابعاد فضا یک امر ضروری می باشد. اگرچه اخیرا الگوریتم های ماشین بردار پشتیبان (SVM)^۲ پایداری نسبتا خوبی را در مواجهه با فضای ویژگی با ابعاد بالا نشان داده اند، لیکن فرضیه تفکیک پذیری بین دو کلاس به صورت خطی یکی از معایب اصلی این الگوریتم است که در داده های سنجش از دور این فرض همیشه صادق نیست. از همین رو الگوریتم انتخاب باند مبتنی بر SVM و پالایش بازگشتی باندها^۳ توسط [۵-۷] ارائه شده اند.

به طور کلی دو روش عمده کاهش ابعاد فضا شامل استخراج ویژگی و انتخاب ویژگی که همان انتخاب باند

فناوری تصویر برداری فراطیفی در مقایسه با تصاویر چند طیفی توانایی بازسازی علائم طیفی پدیده ها را به خوبی کتابخانه طیفی اجسام دارا می باشد. سنجنده های هوا برد AVIRIC و فضا برد HYDICE از این قبیل سنجنده های فراطیفی می باشند. از آن جا که سنجنده فراطیفی در پهنای باند باریک و با فواصل نمونه برداری ثابت از منحنی طیفی اجسام به صورت سیستماتیک نمونه گیری می نماید لزوما تمامی باندها دارای ارزش اطلاعاتی یکسان جهت استخراج اطلاعات از تصویر نمی باشند و چه بسا افزونگی داده نیز بوجود آید. به طور کلی مواجه شدن با داده های فراطیفی در هنگام اخذ داده مشکل نویز، هنگام ذخیره سازی و انتقال مشکل حجم داده، هنگام پردازش پیچیدگی ناشی از ابعاد بالای داده [۱] و به تبع آن مشکل ابعاد بالای فضای فرضیه^۱ به وجود می آید [۲].

از دیدگاه استخراج اطلاعات، ابعاد بالای فضای ویژگی این داده ها یک مشکل عمده است. این مشکل زمانی رخ می دهد که برای تخمین قابل اعتماد پارامترهای طبقه

متقابل ویژگی ها^{۱۱} به ترتیب توسط [۱۴-۱۳] به منظور انتخاب باند ارائه گردیده اند، در واقع چنین روش هایی باندهایی که معیارهایی نظیر بیشینه نمای اطلاعات فشرده شده^{۱۲}، کمترین اطلاعات متقابل، کمترین همبستگی و کمترین شباهت ویژگی ها را ارضا نمی نمایند کنار گذاشته شده و اصطلاحاً فیلتر می کنند همچنین با بکارگیری مفهوم وابستگی توصیفات^{۱۳} که در تئوری مجموعه های تقریب^{۱۴} وجود دارد می توان وابستگی بین باندهای همسایه را بازنمایی نمود و با تعریف بردار خصوصیت برای هر باند امکان کاهش بعد از طریق خوشه بندی فازی باندها به صورت نظارت نشده به وجود می آید [۱۵]. در مقابل، الگوریتم های نظارت شده معمولاً با هدف تفکیک پذیری کلاس ها صورت می پذیرد که به نوبه خود به دو دسته کلی فیلتری^{۱۵} و پوششی^{۱۶} تقسیم بندی می گردند [۱۶]. الگوریتم های فیلتری ویژگی ها را قبل از این که وارد فرآیند طبقه بندی گردند از طریق معیارهای تفکیک پذیری مورد ارزیابی و انتخاب قرار می دهند. روش فیلتر سعی در نگهداشتن ویژگی هایی می نمایند که معیارهای تفکیک پذیری نظیر واگرایی و فواصل نظیر Mohalanobis, Bhattacharyya را بهینه نمایند. برعکس روش پوششی ویژگی های مناسب را بر اساس اثر آنها بر روی بهبود صحت نتایج یک طبقه بندی کننده خاص انتخاب می نماید.

به طور کلی موفقیت روش های انتخاب ویژگی بستگی به سه عامل: الف- روش جستجو جهت ایجاد زیر مجموعه هایی از ویژگی ها ب- تعریف یک معیار برای ارزیابی مناسب بودن زیر مجموعه ویژگی های تولید شده توسط روش های جستجو پ- نحوه ارزیابی زیر مجموعه هایی از ویژگی ها بر اساس معیارهای ب دارد [۱۶]. بنابراین نقاط قوت و ضعف روش های انتخاب ویژگی موجود به این سه عامل بستگی دارد. این مقاله بطور خاص عامل اول را مورد بررسی قرار می دهد. در صورتی که n باند توسط سنجنده فراطیفی اخذ شده باشد تعداد 1-2ⁿ زیرمجموعه از باندها قابل تعریف شدن میباشد. از آنجا که تعداد ویژگی های مناسب از قبل قابل پیشبینی نمی باشد لذا تمامی زیرمجموعه ها می باید مورد ارزیابی قرار گیرند. از طرفی جستجوی سرتاسری همه زیرمجموعه های ممکن از لحاظ زمانی عملی نیست. لذا الگوریتم های جستجو امری ضروری به شمار می آید. در

است وجود دارد. روش های استخراج ویژگی متمرکز بر روی یافتن یک ترکیب خطی یا غیر خطی از ویژگی های اولیه اندازه گیری شده با ابعاد کامل به فضای با ابعاد کم هستند. در واقع ویژگی های با ابعاد کامل توسط یک ترکیب معمولاً خطی به ابعاد کم نگاشته می شوند. یافتن ترکیب خطی معمولاً با هدف بهینه سازی یک معیار تفکیک پذیری کلاس ها نظیر نسبت دترمینان ماتریس پراکندگی بین کلاسی به دترمینان ماتریس پراکندگی داخل کلاسی، یافتن مولفه های مستقل تشکیل دهنده علائم طیفی پیکسل ها و یا برداشتن همبستگی بین ویژگی ها می باشد که به ترتیب منجر به روش های آنالیز تفکیک پذیری خطی^۴ LDA [۴]، آنالیز مولفه های مستقل^۵ ICA [۸] و آنالیز مؤلفه های اصلی^۶ PCA [۴] می گردند. الگوریتم های استخراج ویژگی می توانند به صورت نظارت نشده و نظارت شده مطرح گردند. روش های نظارت نشده روش هایی هستند که نیازی به دانش اولیه و داده های تمرینی ندارند و لزوماً هدف آنها بهینه سازی تفکیک پذیری کلاس ها و بهبود صحت طبقه بندی نمی باشد. روش های PCA, ICA, [9] MNF نمونه هایی از این دست است. در عوض روش های نظارت شده سعی در بهبود صحت طبقه بندی از طریق بهینه سازی یک معیار تفکیک پذیری بین کلاس ها دارند. روش های نظارت شده با توجه به این که ماتریس پراکندگی داخل کلاسی و بین کلاسی به چه نحوی تخمین زده می شوند به دو روش پارامتریک و غیر پارامتریک گونه بندی می شوند. روش های [۱۰] aPAC^۷LDA، نمونه هایی از گونه پارامتریک و روش های آنالیز تفکیک پذیری غیر پارامتریک^۸ NDA^۸، [۱۱] و آنالیز تفکیک پذیری غیر پارامتریک وزن دار^۹ NWFE [۱۲] نمونه هایی از گونه غیر پارامتریک به شمار می آیند. هر یک از روش های استخراج ویژگی مزایا و معایب خاص خود را از نقطه نظر محاسباتی و صحت نتایج دارا هستند.

در مقابل، روش های انتخاب ویژگی به دو دسته عمده نظارت شده و نظارت نشده گونه بندی می شوند. روش های نظارت نشده شامل روش هایی است که در آنها تفکیک پذیری کلاس ها به طور جدی مد نظر نبوده و هدف اصلی ذخیره سازی داده ها با حجم کم می باشد به گونه ای که محتوای اطلاعات تا حد زیادی حفظ گردد. اخیراً روش هایی مبتنی بر شباهت ویژگی ها^{۱۱} و اطلاعات

ویژگی‌ها دارد تا به طور خودکار امکان قرار گرفتن دو ویژگی با همبستگی بالا در یک زیر مجموعه را به حداقل برساند. به علاوه، این روش به صورت پیشرو یا پسرو نبوده که برای رسیدن به یک تعداد ویژگی دلخواه، ویژگی‌ها را از انتها به ترتیب غیر اطلاعات دار بودن حذف یا از ابتدا به ترتیب اطلاعات دار بودن اضافه نماید. روش پیشنهادی از ابتدا سعی در کاوش فضا با تعداد باند مورد نیاز می‌نماید تا مناسب ترین باندهای نماینده را انتخاب نماید. در روش پیشنهادی ابتدا برای باندها بر اساس انعکاسی که در کلاس‌های مختلف داشته‌اند بردار ویژگی تشکیل می‌شود. بدین ترتیب باندها در یک فضا به نام فضای پدیده بازنمایی می‌شوند. در این فضا باندهای مشابه که دارای خصوصیات انعکاسی مشابهی نسبت به کلاس‌ها از خود نشان داده‌اند از طریق الگوریتم خوشه بندی فازی در یک دسته قرار می‌گیرند. سپس در فضای پدیده نزدیک‌ترین باند به مرکز هر دسته به عنوان نماینده باندهای با وابستگی بالا معرفی می‌شود. به علاوه باندهای اطلاعات دار که به صورت منفرد در فضای پدیده قرار گرفته‌اند از طریق دو شاخص آنتروپی و زاویه نسبت به قطر فضای پدیده شناسایی می‌شوند. در ادامه این مقاله ابتدا بازنمایی باندها در فضای پدیده معرفی می‌شود. سپس نحوه انتخاب باند بر اساس خوشه بندی فازی و معیارهای آنتروپی و زاویه بیان می‌شوند. نهایتاً نحوه به کارگیری الگوریتم، آزمایش‌های صورت پذیرفته و نتایج مورد بحث و بررسی قرار می‌گیرد.

بازنمایی باندها

اصولاً یک بردار ویژگی شامل یک مجموعه عنصر است که یک شیء را توصیف می‌کند. از دیدگاه بازشناسی الگو، یک فضا می‌بایست یک مجموعه خصوصیات خاص را به نحوی اختیار نماید تا بتواند با یک بازنمایی محدود که اشیاء را توصیف می‌کند عمل یادگیری را به انجام برساند [۱۷]. بردار ویژگی پیکسل دلخواه x برحسب پاسخ طیفی در باندها مختلف به صورت $x = [x_1, x_2, \dots, x_n]^T$ تعریف می‌شود که n تعداد باندها و x_i میزان انعکاس باند i ام است. از این رو پیکسل‌های تصویر در فضای ویژگی برای پردازش‌هایی نظیر خوشه بندی و طبقه بندی بازنمایی می‌شوند. در نتیجه این فضا صرفاً برای تحلیل پیکسل‌ها و طبقه بندی تصویر مناسب است. از آن جایی که روش

نتیجه موفقیت الگوریتم‌های انتخاب ویژگی در گرو الگوریتم‌های بهینه جستجو جهت بررسی ترکیبات ممکن m تایی (زیر مجموعه ویژگی‌ها) از n تعداد ویژگی (تعداد کل باندها) است. برای این منظور الگوریتم‌های جستجوی بسیاری توسط محققین مختلف ارائه گردیده است از جمله آنها می‌توان به الگوریتم‌های جستجوی سرتاسری^{۱۷}، جستجوی پسروی پی در پی^{۱۸}SBS، جستجوی پسروی پی در پی شناور^{۱۹}SBFS، جستجوی شاخه و حد^{۲۰}، جستجوی پیشرو پی در پی^{۲۱}SFS، جستجوی پی در پی پیشرو شناور^{۲۲}SFFS، جستجوی پرتوی^{۲۳} و جستجوی تصادفی از طریق الگوریتم‌های زیستی^{۲۴} اشاره نمود. از جمله مشکلاتی که می‌توان برای الگوریتم‌های جستجو برشمرده عبارتند از: اساساً از آن جا که جستجو سرتاسری عملی نمی‌باشد الگوریتم‌های موجود معمولاً منتهی به بهینه سازی‌های محلی می‌شوند. در ارتباط با انواع روش‌های پسروی از آن جا که ریزش باندهای غیر اطلاعات دار از مجموعه ویژگی‌های با ابعاد کامل آغاز می‌شود در ابتدای امر نیاز به داده‌های زیاد برای تخمین اعتماد پذیر پارامترهای معیارهای ارزیابی مناسب بودن ویژگی‌ها دارد. از طرف دیگر معمولاً مناسب ترین تعداد باند به مراتب کمتر از تعداد باندهای اخذ شده است بنابراین الگوریتم‌های پسرو بسیار زمانبر بوده و الگوریتم‌های پیشرو از ارجحیت بالایی برخوردار می‌باشند. این مسئله در روش شاخه و حد کاملاً مشهود است [۱]. از طرف دیگر اکثر جستجوگرها از این واقعیت که باندهای مجاور معمولاً دارای همبستگی بالایی می‌باشند و نباید از ابتدا در یک زیر مجموعه قرار گیرند غفلت می‌ورزند. به بیان دیگر کنار گذاشتن باندهای مشابه پس از ارزیابی زیر مجموعه ویژگی‌ها صورت می‌پذیرد. از طرفی قرار گرفتن این ویژگی‌های مشابه در زیر مجموعه‌ها باعث مشکلات محاسباتی نظیر سینگولاریتی ماتریس‌های واریانس کواریانس برای محاسبه معیارهای ارزیابی شده که منجر به ارزیابی نادرست زیر مجموعه‌ها می‌گردد. این موضوع در روش‌های پوششی که نیاز به آموزش طبقه بندی کننده‌های پارامتریک دارد نیز مشهود می‌باشد. لذا گریز از قرار نگرفتن دو باند در یک زیر مجموعه که دارای وابستگی زیاد نسبت به یک دیگر می‌باشند امری مهم در الگوریتم‌های جستجو است.

الگوریتم پیشنهادی سعی در یافتن زیرمجموعه‌ای از

طیفی مجاور هم نیستند. با این توصیفات یک بلوک یا مجموعه بلوک های مشابه از باندها از لحاظ انعکاسی در فضای طیفی به صورت یک دسته در فضای پدیده نمایان می شوند. علاوه بر این همان طور که می توان در شکل (۱) ملاحظه کرد باند t در مواجهه با سه کلاس، انعکاس یکسانی را از خود نشان داده است. بنابراین از دیدگاه هندسی این باند بر روی قطر فضای پدیده قرار گرفته است. بنابراین اینطور می توان استدلال نمود که باندهایی که نزدیک قطر در فضای پدیده قرار گرفته اند حاوی ارزش اطلاعاتی جهت جداسازی پدیده ها نیستند. در عوض باندهایی وجود دارند نظیر s که نتنها در مقایسه با باندهای مجاور میزان انعکاس در آنها متفاوت است بلکه میزان انعکاس در کلاس های مختلف نیز متفاوت است. در نتیجه چنین کلاس هایی در فضای پدیده جدا از باندهای با وابستگی بالا که به شکل خوشه ظاهر شده اند قرار می گیرند و از فاصله نسبتاً زیادی نسبت به قطر نیز برخوردار می باشند. بنابراین می توان گفت چنین باندهای مجزایی دارای پتانسیل تفکیک پذیری مناسبی و به الطبع ارزش اطلاعاتی دارند. با این توصیفات ۳ نوع باند در فضای پدیده قابل شناسایی هستند: الف- باندهایی که وابستگی و شباهت زیادی نسبت به یکدیگر دارند که در غالب یک دسته در فضای پدیده ظاهر شده اند و از میان آنها باید یک نماینده انتخاب گردد. ب- باندهایی که حاوی اطلاعات نیستند و در مجاورت قطر فضای پدیده پهلو گرفته اند. پ- باندهایی که در درون یک دسته قرار نگرفته و تا حد امکان از قطر نیز فاصله گرفته اند. چنین باندهایی، باندهای اطلاعات دار قلمداد می شوند.

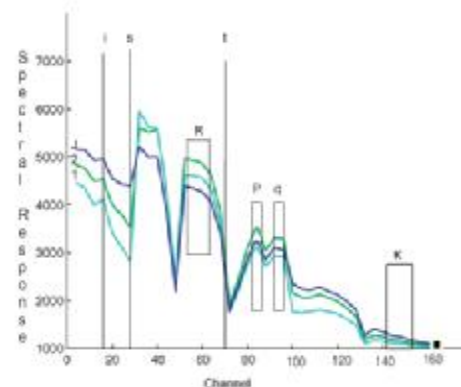
انتخاب باند مبتنی بر خوشه بندی فازی

پس از معرفی انواع باندها در فضای پدیده، در این مرحله نحوه شناسایی باندهای با وابستگی بالا و باندهای اطلاعات دار تشریح می شود.

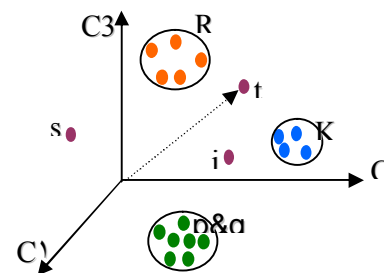
مرحله اول: شناسایی باندهای با وابستگی بالا

در این مرحله ابتدا باندهایی که وابستگی بالایی نسبت به یکدیگر دارند به وسیله الگوریتم خوشه بندی FCM در یک دسته قرار می گیرند. اگر فرض کنیم h_i باند i ام با L بعد (تعداد کلاس ها) و دارای m_{ki} عنصر باشد که در فضای پدیده قرار گرفته است. هدف خوشه بندی n باند

های انتخاب باند سعی در مطالعه و تحلیل خصوصیات باندها دارند منطقی آن است که باندها بر حسب خصوصیاتشان بیان گردند. بر این اساس پیشنهاد ما بازنمایی باندها در فضای جدید با نام فضای پدیده است. در این فضا یک باند بر حسب پاسخ طیفی که از کلاس های مختلف گرفته است توصیف می شود. در صورتی که فرض کنیم یک تصویر شامل L کلاس و n باند باشد منحنی طیفی کلاس ها را می توان با شکل (۱) نشان داد. بنابراین بردار ویژگی باند h_i در فضای پدیده (شکل (۲)) به صورت قابل بیان است.



شکل ۱: منحنی طیفی.



شکل ۲: فضای پدیده.

که h_i بردار ویژگی باند i ام، m_{ji} میانگین کلاس j ام در باند i ام است. از این رو فضای پدیده دارای L بعد است و مطالعه و بررسی باندها را در این فضا تسهیل می کند. از دیدگاه فیزیکی انعکاس پدیده هایی که در بلوک R قرار گرفته اند مانند هم هستند. بنابراین این طور می توان استدلال کرد که باندهای واقع شده در بلوک R وابستگی بالایی نسبت به هم دارند. در نتیجه انتظار می رود که این باندهای با وابستگی بالا در فضای پدیده در کنار یک دیگر قرار گیرند و یک دسته را تشکیل دهند. این وضعیت می تواند برای بلوک های p, q نیز رخ دهد که از لحاظ

میان باندهای مجزا چندین باند اولی که دارای بیشترین زاویه نسبت به قطر هستند انتخاب شده و به مجموعه باندهای انتخاب شده از مرحله اول اضافه گردیدند. بدین ترتیب تعداد باندهای انتخاب شده برابر $(c+m)$ می گردد که c تعداد دسته ها و m باندهای منفردی می باشند که زاویه بزرگی نسبت به قطر دارند.

نحوه به کارگیری الگوریتم

الگوریتم پیشنهادی انتخاب ویژگی به صورت پوششی مورد استفاده قرار می گیرد. بدین ترتیب که داده های آموزشی به دو مجموعه تقسیم گردیدند. مجموعه اول شامل ۴۰٪ داده های آموزشی است که برای محاسبه منحنی طیفی کلاس ها و به تبع آن فضای پدیده مورد استفاده قرار می گیرند. همچنین از این داده ها برای تخمین پارامترهای آماری و یادگیری طبقه بندی کننده بیشترین شباهت استفاده گردید. مجموعه دوم شامل ۶۰٪ مابقی داده های آموزشی است که به عنوان مجموعه داده های اعتبارسنجی برای ارزیابی مجموعه باندهای انتخاب شده توسط الگوریتم خوشه بندی و معیار زاویه مورد استفاده قرار گرفتند. از آنجا که روش پیشنهادی وابسته به الگوریتم خوشه بندی فازی می باشد در هر اجرای برنامه، برای شروع خوشه بندی نیاز به مقدار دهی اولیه است که منجر به تغییر نتایج حاصله می شود. بنابراین برای فائق آمدن به مشکل مقدار دهی اولیه، الگوریتم خوشه بندی ده بار تکرار شده و بهترین جواب بر حسب بیشترین صحت کلی طبقه بندی داده های اعتبارسنجی از میان فضای جواب ها انتخاب می شود.

آزمایش ها و ارزیابی

برای ارزیابی روش پیشنهادی یک قطعه تصویر که از منطقه شمال غربی ایندیانا توسط سنجنده AVIRIS با قدرت تفکیک ۲۰ متر در سال ۱۹۹۲ اخذ شده مورد استفاده قرار گرفت [۱۹]. این داده دارای ۱۶ کلاس بوده و از آنجا که کلاس های واقع شده در این قطعه تصویر شباهت طیفی بسیاری نسبت به هم دارند این داده به عنوان یک داده چالش آور و مینا برای محک زنی و ارزیابی الگوریتم های پیشنهادی محققان مطرح است. این داده دارای ۲۲۰ باند طیفی است که پهنای هر باند ۱۰ نانومتر می باشد و در محدوده طیفی ۰.۴ تا ۲.۴۵ تصویر

به دسته در فضای پدیده با استفاده از الگوریتم FCM است. FCM یک الگوریتم خوشه بندی تکراری است که مقادیر درجه عضویت کلاس ها به وسیله کمینه کردن خطای تابع ذیل به دست می آید [۱۸].

$$J(U, V) = \sum_{i=1}^n \sum_{j=1}^c (u_{ji})^p d_{ij}^2 \quad (1)$$

$$d_{ij} = |h_i - v_j| = \sqrt{\sum_{k=1}^L (m_{ki} - v_{kj})^2} \quad (2)$$

در این رابطه v_j بردار میانگین دسته z ام، U_{ji} مقادیر درجه عضویت باند z ام به دسته z ام، c تعداد دسته ها، n تعداد باندها، d_{ij} فاصله بین پاسخ طیفی کلاس ها در باند z ام و میانگین دسته z ام و p مقداری است که درجه فازی بودن را کنترل می کند. مقادیر p می تواند از ۱ (غیر فازی) تا بی نهایت کاملاً فازی تغییر کند. به منظور جداسازی باندهای با وابستگی زیاد و باندهای مجزا در فضای پدیده، آنتروپی باندها بر حسب مقادیر درجه عضویت کلاس ها به صورت ذیل محاسبه شد.

$$E_i = -\sum_{j=1}^c U_{ji} \log_2 U_{ji} \quad \forall i=1,2,\dots,n \quad (3)$$

که در رابطه (۳) E_i آنتروپی باند z ام می باشد. بدین ترتیب باندهای با آنتروپی و عدم قطعیت بالا به عنوان باندهای مجزا کنار گذاشته شدند. و بقیه باندها با قطعیت بالا بر اساس بزرگترین درجه عضویت تنها به یک دسته تعلق یافتند. در هر دسته فاصله بین باندهای هر دسته و مرکز دسته محاسبه شده و باندی که نزدیک ترین فاصله را به مرکز دسته داشت به عنوان نماینده مجموعه باندهای قرار گرفته در آن دسته معرفی شد.

مرحله دوم: شناسایی باندهای اطلاعات دار

هدف از این مرحله شناسایی و حذف باندهای غیر اطلاعات دار می باشد. در این مرحله ابتدا زاویه بین باندهای انتخاب شده (نزدیکترین باند به مرکز هر دسته) از مرحله قبل با قطر مطابق رابطه ذیل محاسبه شد.

$$A_l = \cos^{-1} \left(\frac{\sum_{k=1}^L m_{kl}}{\|h_l\| \sqrt{L}} \right) \quad l=1,2,3,\dots,q \quad (4)$$

بدین ترتیب باندی که زاویه کمی با قطر می سازد حذف می گردد. به عبارت دیگر دسته ای از باندها که در اطراف قطر گرد هم آمده اند کنار گذاشته می شوند. در مرحله بعدی زاویه بین باندهای منفرد شده از مرحله اول که دارای آنتروپی بالا بودند با قطر محاسبه می شوند. از

توسط روش‌های مختلف جستجوی پیشرو شناور، پسر و شناور و روش پیشنهادی مبتنی بر خوشه بندی می باشد. از میان سه روش جستجوی فوق، روش های پیشرو شناور و روش پیشنهادی مبتنی بر خوشه بندی از طبقه بندی کننده بیشترین شباهت برای ارزیابی و مقایسه مجموعه ویژگی‌های ایجاد شده (توسط الگوریتم‌های جستجو) استفاده شد. در واقع در هر یک از دو روش فوق الذکر تصمیم گیری بر روی اضافه یا حذف نمودن یک ویژگی به عملکرد آن ویژگی در بهبود نتایج طبقه بندی داده های اعتبار سنجی بستگی دارد.

در مقابل از آنجا که قسمت اول داده های آموزشی (۴۰٪ داده‌های آموزشی) به اندازه کافی برای آموزش طبقه بندی کننده بیشترین شباهت (یا طبقه بندی کننده تفکیک کننده درجه دو QDC) در مراحل ابتدایی روش جستجوی پسر و شناور (فضای با ابعاد کامل) نبودند، لذا معیاری که برای ارزیابی مناسب بودن زیر مجموعه باندهای تولید شده توسط این الگوریتم مورد استفاده قرار گرفت رد پای نسبت ماتریس پراکندگی بین کلاسی به داخل کلاسی بود که مطابق رابطه ذیل محاسبه شد.

$$J = \text{trace} ((S_w)^{-1} * S_b) \quad (5)$$

که در رابطه فوق S_w و S_b از رابطه های ذیل به دست می آیند.

$$S_w = \sum_{i=1}^L P_i E \{ (X - m_i)(X - m_i)^T | \omega_i \} = \sum_{i=1}^L P_i \sum_i \quad (6)$$

$$S_b = \sum_{i=1}^L P_i (m_i - m_0)(m_i - m_0)^T = \sum_{i=1}^L \sum_{j=i+1}^L P_i P_j (m_i - m_j)(m_i - m_j)^T \quad (7)$$

که در روابط فوق \sum_i ماتریس واریانس کواریانس کلاس نام P_i ، احتمال اولیه کلاس نام، L تعداد کلاس‌ها، m_i و m_j نیز به ترتیب میانگین کلاس نام و نام هستند.

پس از اجرای دو الگوریتم اول، صحت کلی داده های اعتبارسنجی به ازای هر زیرمجموعه ویژگی محاسبه گردید. همان‌طور که در شکل (۵) می توان ملاحظه کرد دو روش اول توانسته اند تا حد امکان ویژگی هایی را انتخاب نمایند که داده های اعتبار سنجی را به خوبی مدل نمایند و صحت طبقه بندی آنها را به ۱۰۰٪ برسانند. از سوی دیگر مقدار معیار J برای الگوریتم پسر و شناور نیز به ازای هر زیر مجموعه ویژگی در شکل (۶) نشان داده شده است. همان‌طور که در شکل (۶) ملاحظه می گردد با افزایش تعداد ویژگی ها میزان J افزایش یافته که این امر

برداری صورت پذیرفته است. ابتدا ۲۰ باند جذبی آب در محدوده‌های باندهای (۱۰۴ - ۱۰۸ و ۱۵۰-۱۶۳ و ۲۲۰) حذف گردیدند و ۱۵ باند نویزدار از مجموع باندهای باقی مانده حذف شدند. بدین ترتیب ۱۸۵ باند به منظور گزینش باند باقی ماندند. ابعاد تصویر ۱۴۵ در ۱۴۵ است که یک ترکیب رنگی از تصویر و واقعیت زمینی مربوطه به ترتیب در شکل های (۳) و (۴) نشان داده شده است. به منظور بررسی مناسب تر، ۷ کلاس که تقریباً داده‌های تمرینی کمی داشتند حذف شده و صرفاً ۹ کلاس باقیمانده مورد ارزیابی قرار گرفتند. علاوه بر این داده‌های آموزشی و آزمایشی مطابق با جدول (۱) انتخاب گردیدند. همان‌طور که ملاحظه می‌شود محدوده مناطق آموزشی و آزمایشی از لحاظ مکانی مجزا هستند و همبستگی مکانی بین پیکسل‌ها وجود ندارد.



شکل ۳: یک ترکیب رنگی از تصویر فراطیفی.

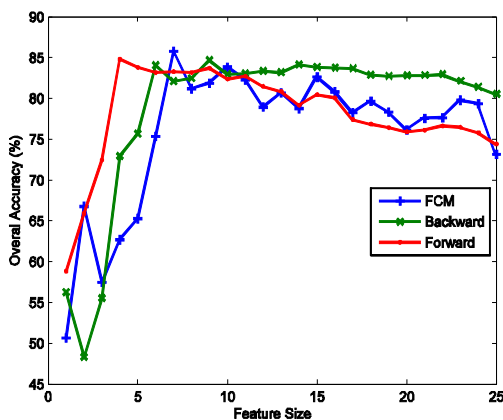


شکل ۴: واقعیت زمینی تصویر شکل ۳.

آزمایش ۱

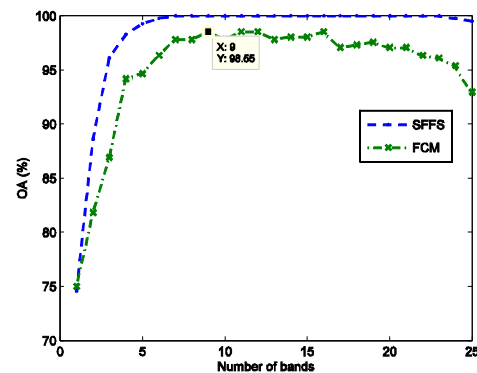
هدف از این تجربه بررسی چگونگی انتخاب باند

طبقه‌بندی کننده بیشترین شباهت صحت کلی و متوسط بالایی نسبت به دو طبقه بندی کننده دیگر دارد. برای همین منظور در جدول (۲) نتایج مربوط به صحت هر کلاس بر حسب طبقه بندی کننده بیشترین شباهت ارائه گردیده است. با مقایسه صحت کلاس به کلاس الگوریتم‌های مختلف می‌توان دریافت که تقریباً کلاس‌های علفزار مرتعی، لوبیای کاشته شده در زمین با عملیات خاک ورزی تمیز و جنگل دارای مشکل یکسان از نقطه نظر تفکیک پذیری با سایر کلاس‌ها هستند. جدول (۳) تعداد باندهای بهینه و شماره باندهای انتخاب شده را مبتنی بر طبقه بندی کننده بیشترین شباهت نشان می‌دهند. صحت کلی داده‌های آزمایشی بر اساس طبقه‌بندی کننده بیشترین شباهت برای روش‌های جستجوی پیشرو و پسرو شناور و FCM به ازای مجموعه باندهای مختلف در شکل (۷) نشان داده شده است. همان‌طور که در شکل ملاحظه می‌شود منحنی روش پیشنهادی به ازای ۹ باند از صحت نسبی بالایی نسبت به دو الگوریتم برخوردار است. به علاوه منحنی طیفی کلاس‌ها بر حسب باندهای انتخابی سه روش که در جدول (۳) آمده است در شکل‌های (۸)، (۹) و (۱۰) نشان داده شده است. با مقایسه منحنی طیفی اصلی کلاس‌ها (شکل (۱۱)) و منحنی طیفی کلاس‌ها که حاصل باندهای انتخابی سه روش می‌باشد می‌توان نتیجه گرفت که علاوه بر حفظ ساختار اصلی منحنی طیفی مینا در تعداد باندهای نسبتاً کم همانند سنجنده‌های چند طیفی این روش توانسته است به نتایج قابل توجهی دست یابد.

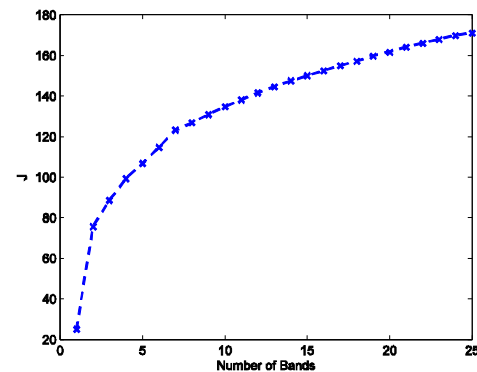


شکل ۷: صحت کلی داده‌های آزمایشی برای روش‌های جستجوی پیشرو و پسرو شناور و FCM به ازای مجموعه باندهای مختلف.

ناشی از افزایش ابعاد ویژگی است. باید توجه داشت که افزایش این نسبت لزوماً صحت طبقه بندی را در ابعاد بالا برای داده‌های آزمایشی توجیه نمی‌نماید.



شکل ۵: صحت کلی داده‌های اعتبار سنجی برای روش‌های پیشرو شناور و FCM به ازای مجموعه باندهای مختلف.



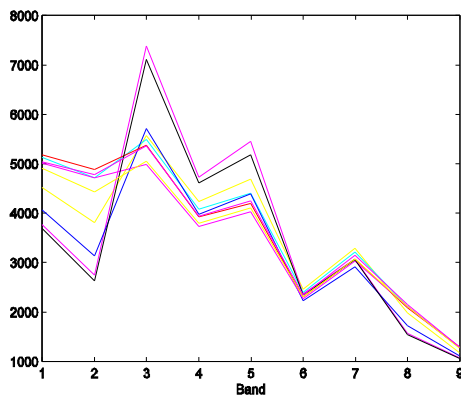
شکل ۶: معیار J برای کل داده‌های آموزشی به ازای مجموعه باندهای مختلف.

آزمایش ۲: انتخاب تعداد باندهای بهینه

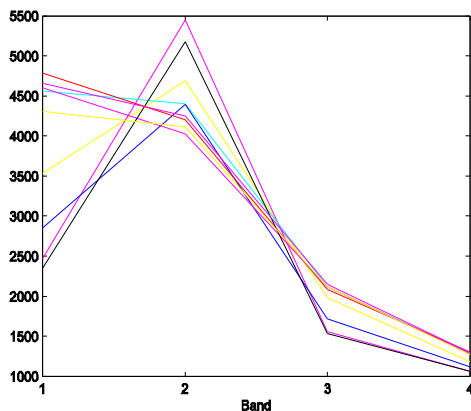
در این قسمت باندهایی که در مرحله قبل تعیین گردیده اند در مقابل داده‌های آزمایشی مورد آزمون قرار می‌گیرند. در این تجربه تمامی داده‌های آموزشی برای محاسبه پارامترهای طبقه بندی کننده بیشترین شباهت مورد استفاده قرار گرفتند. البته برای بررسی این که این روش واقعا از طبقه بندی کننده بیشترین شباهت جانبداری نمی‌کند، روش‌های دیگر طبقه بندی کننده نظیر k نزدیک ترین همسایگی (KNNC) و طبقه بندی کننده تفکیک کننده خطی (LDC) نیز برای ارزیابی مورد استفاده قرار گرفتند. نتایج حاصل از سه طبقه بندی کننده برای سه روش انتخاب باند بر حسب صحت کلی OA^{۲۵} و متوسط صحت AA^{۲۶} کلاس‌ها در جدول (۲) نشان داده شده است. همان‌طور که ملاحظه می‌شود

جدول ۱: تعداد و موقعیت محدوده های آموزشی و آزمایشی.

Class Name	Training Data			Test Data			نام کلاس
	Row	Column	N.train	Row	Column	N.test	
Corn-no till	64-71	38-47	80	31-39	27-59	297	ذرت کاشته شده در زمین بدون شخم
Corn-min	61-71	13-20	88	60-70	4-12	99	ذرت کاشته شده در زمین با شخم اندک
Grass/Pasture	74-78	1-12	60	75-83	13-23	99	علفزار مرتعی
Grass/Trees	100-114	59-70	180	49-71	27-36	230	زنبق استرالیایی
Hay-windrowed	37-57	132-140	189	39-60	122-132	242	یونجه درو شده و خشک شده
Soybeans-no till	35-38	82-93	48	40-63	76-93	432	لوبیای کاشته شده در زمین بدون شخم
Soybeans-min	64-69	50-70	126	75-115	33-47	615	لوبیای کاشته شده در زمین با شخم اندک
Soybeans-clean	7-19	37-43	91	17-25	54-68	135	لوبیای کاشته شده در زمین با عملیات خاک ورزی تمیز
Woods	31-38	98-119	176	120-138	87-122	684	جنگل



شکل ۸: منحنی طیفی کلاس ها به ازای ۹ باند انتخابی روش پسرو.



شکل ۹: منحنی طیفی کلاسها به ازای ۴ باند انتخابی روش پیشرو.

جدول ۲: صحت کلی و متوسط طبقه بندی کننده های مختلف به ازای روش های مختلف انتخاب باند به همراه صحت طبقه بندی کننده بیشترین شباهت برای هر یک از کلاس ها.

	SBFS	SFFS	FCM
Corn-no till	100	99.59	100
Corn-min	98.99	80.81	82.83
Grass/Pasture	37.04	45.18	46.67
Grass/Trees	100	100	100
Hay-windrowed	94.95	95.96	95.96
Soybeans-no till	91.43	96.06	94.21
Soybeans-min	99.56	97.66	99.85
Soybeans-clean	50.50	59.60	68.69
Woods	74.80	70.73	68.94
AA-QDC	83.03	82.84	84.13
OA	84.72	84.79	85.77
AA-LDC	71.07	66.63	71.83
OA	67.70	59.94	68.94
AA-KNNC	68.28	66.66	68.61
OA	66.64	68.83	69.22

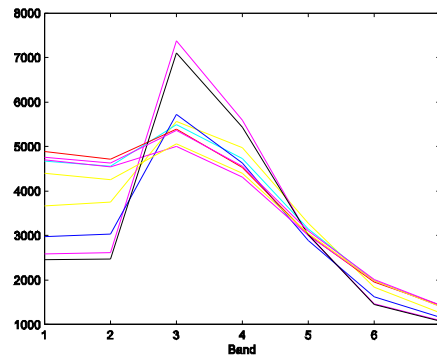
جدول ۳: باندهای انتخاب شده در سه روش جستجوی باندها.

SBFS	SFFS	FCM
11	24	4
21	68	7
45	123	9
59	165	11
68		14
78		17
83		20
123		
165		

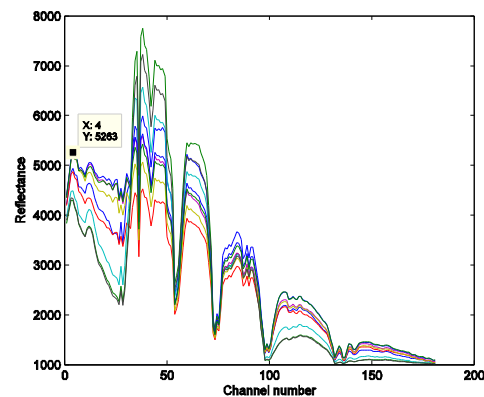
می‌دهد. باندهایی را نیز که از لحاظ انعکاسی با سایر باندها متفاوت باشد با استفاده از معیارهای آنتروپی و زاویه بعنوان باند منفرد و حاوی اطلاعات در نظر می‌گیرد. از آنجا که این روش از همان ابتدا m باند را در فضای پدیده جستجو می‌نماید و نیازی به اضافه یا حذف نمودن باندها تا رسیدن به m باند را ندارد لذا تعداد بررسی های انجام شده (از لحاظ مناسب بودن فضای پیشنهادی و مقایسه بین زیر مجموعه ها) کاهش یافته و باعث افزایش سرعت جستجو می‌شود. اگرچه در این روش عملیات خوشه بندی چندین بار تکرار می‌گردد تا به مشکل تصادفی بودن نقاط شروع برای خوشه بندی فازی فائق آید با این حال زمان انجام الگوریتم از روش های جستجوی متداول کمتر است (زمان محاسبه الگوریتم پیشرو شناور و پسرو شناور به ترتیب $\frac{2}{3}$ و $\frac{2}{1}$ برابر زمان جستجوی مبتنی بر خوشه بندی بود). از سوی دیگر در روش های متداول هر مجموعه باند با مجموعه باندهای بعد و قبل خود در یک باند تفاوت دارد و این امر وابستگی نسبی بین دو مجموعه باند از لحاظ ابعاد فضا ایجاد می‌کند. در حالی که در روش پیشنهادی هر مجموعه باند به طور مستقل به دست می‌آید و وابسته به مجموعه های قبل و بعد از خود نمی‌باشد. نتایج به دست آمده بهبود درستی متوسط به میزان $\frac{1}{10}$ را نسبت به روش پسرو شناور و بهبود درستی کلی به میزان $\frac{0.98}{100}$ نسبت به روش پیشرو شناور را نشان دادند. این بهبود درستی در طبقه بندی کننده های کمترین فاصله و KNN نیز مشهود است. نکته قابل توجه اینکه در سه روش کلاس‌های مشکل دار از لحاظ صحت همچنان باقی مانده اند.

تقدیر و تشکر

مولفین مراتب تشکر و قدر دانی خود را از آقای دکتر Landgrebe از دانشگاه پردو جهت در اختیار قرار دادن تصویر و واقعیت زمینی دارند.



شکل ۱۰: منحنی طیفی کلاس‌ها به ازای ۷ باند انتخابی روش FCM.



شکل ۱۱: منحنی طیفی اصلی کلاس‌ها.

نتیجه گیری

در این مقاله یک روش نوین انتخاب باند مبتنی بر خوشه بندی فازی باندها ارائه گردید. بطور خاص روش پیشنهادی متمرکز بر جستجو و ارائه مجموعه باندهای مناسب جهت ارزیابی به روش پوششی دارد. روش پیشنهادی نیازی به جستجوی دنباله دار و پیوسته برای اضافه یا حذف نمودن تدریجی باندها ندارد. در واقع این روش الگوریتم های جستجوی متداول را کنار گذاشته و بجای آن باندهایی که در مواجهه با پدیده های مختلف دارای خصوصیات انعکاسی یکسانی و در نتیجه وابستگی بالا می باشند را در فضای پدیده در یک دسته قرار

مراجع

- 1 - Serpico, S. B, J. and Moserd, G. (2007). "Extraction of spectral channels from hyperspectral images for classification purposes." *IEEE Trans. Geosci. Remote Sens.*, Vol. 45, No. 2, PP. 484-494, February.
- 2 - Kumar, S. (2000). *Modular Learning Through Output Space Decomposition*. Ph.D. Dissertation, Department of Electrical and Computer Engineering, The University of Texas at Austin.

- 3 – Gordon, P. (1968). “Hughes On the Mean Accuracy of Statistical Pattern Recognizers.” *IEEE Transactions on Information Theory*, PP. 55-63.
- 4 - Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic.
- 5 - Mao, K. Z. (2004). “Feature subset selection for support vector machines through discriminative function pruning analysis.” *IEEE Transactions on Systems, man, and Cybernetics—Part B: Cybernetics*, Vol. 34, No. 1, PP. 60-67.
- 6 - Pal, M. (2006). “Support vector machine based feature selection for land cover classification: a case study with DAIS hyperspectral data.” *International Journal of Remote Sensing*, Vol. 27, PP. 2877-2894 Number 14, 20 July.
- 7 - Skurichina, M., Paclik, P., Duin, R. P. W., de Veld, D., Sterenborg, H. J. C. M. and Witjes, M. J. H. (2004). “Selection/Extraction of Spectral Regions for Autofluorescence Spectra Measured in the Oral Cavity.” *Structural, Syntactic, and Statistical Pattern Recognition*, ISBN 3-540-22570-6, no. 3139, PP. 1095-1103, Springer Verlag, Berlin.
- 8 – Hyvarinen, A. and Oja, E. (2000). “Independent component analysis: Algorithms and applications.” *Neural Netw.*, Vol. 13, No. 4/5, PP. 411–430, May/June.
- 9 - Green, A. A., Berman, M., Switzer, P. and Craig, M. D. (1988). “A transformation for ordering multispectral data in terms of image quality with implications for noise removal,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 26, No. 1, Pp. 65–74.
- 10 - Loog, M., Duin, R. P. W. and Haeb-Umbach, R. (2001). “Multiclass linear dimension reduction by weighted pairwise Fisher criteria.” *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 23, No. 7, PP. 762–766.
- 11 - Bressan, M. and Vitri, J. (2003). “Nonparametric discriminant analysis and nearest neighbor classification.” *Pattern Recognit. Lett.*, Vol. 24, No. 15, PP. 2743–2749.
- 12 - Kuo, B. C. and Landgrebe, D. A. (2004). “Nonparametric weighted feature extraction for classification.” *IEEE Trans. Geosci. Remote Sens.*, Vol. 42, No. 5, PP. 1096–1105.
- 13 - Mitra, P., Murthy, C. A. and Pal, S. K. (2002). “Unsupervised Feature Selection Using Feature Similarity.” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 3, PP. 301-312.
- 14 - Martínez-Usó, A., Pla, F., Sotoca, J. M. and García-Sevilla, P. (2006). “Clustering-based multispectral band selection using mutual information.” *The 18th International Conference on Pattern Recognition (ICPR'06) IEEE computer society*.
- 15 - Shi, H., Shen, Y. and Liu, Z. (2003). “Hyperspectral Bands Reduction Based on Rough Sets and Fuzzy C-Means Clustering.” *Instrumentation and Measurement Technology Conference, Proceedings of the 20th IEEE*, Vol. 2, PP. 1053 – 1056.
- 16 - Guyon, I., Gunn, S., Nikravesh, M. and Zadeh, A. (2006). *Feature Extraction Foundations and Applications*, ISBN 3540354875, Springer; 1 edition.
- 17 - Pekalska, E. (2005). *dissimilarity representations in pattern recognition* Phd thesis, Delft University.
- 18 - Bezdek, J. C., Ehrlich, R. and Fcm, F. W. (1984). “The fuzzy c-means clustering algorithm.” *Computers and Geosciences*, Vol. 10, PP. 191-203.
- 19 - <ftp://ftp.ecn.purdue.edu/biehl/MultiSpec/>

واژه های انگلیسی به ترتیب استفاده در متن

- 1- Large hypothesis space
- 2 - Support Vector Machine
- 3 - Recursive Band Elimination
- 4 - Linear Discriminant Analysis
- 5 - Independent Component Analysis
- 6 - Principle Component Analysis
- 7 - Approximate pair-wise Accuracy Criterion
- 8 - Nonparametric Feature Extraction
- 9 - Nonparametric Weighted Feature Extraction
- 10 - Feature Similarity
- 11 - Mutual Information
- 12 - Maximum information compression index
- 13 - Attribute Dependency
- 14 - Rough Set
- 15 - Filter
- 16 - Wrapper
- 17 - Exhaustive Search
- 18 - Sequential Backward Search
- 19 - Sequential Backward Floating Search
- 20 - Branch and Bound
- 21 - Sequential Forward Search
- 22 - Sequential Forward Floating Search
- 23 - Beam Search
- 24 - Genetic Algorithm
- 25 - Overall Accuracy
- 26 - Average Accuracy