

نگاشت مناطق آلتراسیون با استفاده از مدل آمیخته‌ی گاوسی و نقشه‌بردار زاویه طیفی

مریم لطفی^۱، حمید قنبری^{*}، حسین عارفی^۲، عباس بحرودی^۳

^۱ دانشجوی کارشناسی ارشد - دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی -

دانشگاه تهران

{*hamid.ghanbari, lotfi.maryam}@ut.ac.ir

^۲ استادیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی - پردیس دانشکده‌های فنی - دانشگاه تهران

hossein.arefi@ut.ac.ir

^۳ استادیار دانشکده مهندسی معدن - پردیس دانشکده‌های فنی - دانشگاه تهران

bahroudi@ut.ac.ir

(تاریخ دریافت مرداد ۱۳۹۵، تاریخ تصویب اسفند ۱۳۹۵)

چکیده

در ایران باتوجه به وسعت زیاد کشور و گستردگی مناطق پتانسیل‌دار ذخایر معدنی، شناسایی و مدیریت صحیح ذخایر اهمیت بسیار بالایی دارد. از طرف دیگر به دلیل پیشرفت‌های فن‌آوری سنجش از دور و تولید داده‌های فراطیفی با اطلاعات طیفی فراوان، استفاده از این داده‌ها جهت مطالعه‌ی دقیق پدیده‌ها به سرعت در حال گسترش است. تصاویر فراطیفی به دلیل نمایش گسترده خصوصیات طیفی عوارض و پدیده‌های سطح زمین در بسیاری از علوم مرتبط با زمین مورد توجه قرار گرفته‌اند. از آنجایی که بسیاری از مواد معدنی در محدوده مادون قرمز کوتاه و ویژگی‌های طیفی منحصر به فردی دارند، تشخیص و تفکیک آن‌ها به کمک تصاویر فراطیفی به جهت پهنای بانندی بسیار کمی که دارند، بسیار حائز اهمیت است. در این پژوهش قصد داریم به تفکیک مناطق آلتراسیون مرتبط با مس پورفیری به کمک تصاویر فراطیفی هایپریون و با استفاده از دو الگوریتم نقشه بردار زاویه طیفی (SAM) و طبقه‌بندی کننده مدل آمیخته گاوسی پردازیم. شایان ذکر است که انجام پیش‌پردازش‌های لازم بر روی تصاویر فراطیفی هایپریون به دلیل داشتن نسبت سیگنال به نویز پایین، یکی از گام‌های بسیار ضروری به جهت رسیدن به نتایج مطلوب می‌باشد. به منظور مقایسه و ارزیابی صحت روش‌های پیشنهادی در این تحقیق، ابتدا داده تصویری از سنجنده هایپریون مورد استفاده قرار گرفت. همچنین یک داده شبیه‌سازی شده مورد استفاده قرار گرفت، تا کارایی الگوریتم‌های مورد استفاده در این تحقیق را نشان دهند. برای ارزیابی نتایج نیز از ماتریس خطا به کمک داده‌های زمینی استفاده شد و مولفه‌های دقت کلی و ضریب کاپا به عنوان معیار ارزیابی مورد استفاده قرار گرفتند. دقت کلی و ضریب کاپا برای الگوریتم SAM و آمیخته گاوسی به ترتیب معادل با ۰/۸۲، ۰/۷۵ و ۰/۸۰، ۰/۷۱ حاصل شد.

واژگان کلیدی: نقشه‌بردار زاویه طیفی، تصاویر فراطیفی، مدل آمیخته گاوسی، پتانسیل‌یابی مواد معدنی

۱- مقدمه

با ظهور علم سنجش از دور و دسترسی به تصاویر ماهواره‌ای، امکان بهره‌گیری از این علم در زمینه پتانسیل‌یابی مواد معدنی به‌وجود آمده است. از آنجایی که هر یک از کانی‌ها دارای منحنی‌های طیفی منحصر به فردی می‌باشند، می‌توان از این موضوع به‌عنوان وجه تمایزی برای تشخیص این کانی‌ها از یکدیگر استفاده نمود. با بررسی‌های میدانی انجام شده، مشخص شده است که نتایج حاصل از تحلیل‌های طیفی که با استفاده از تصاویر سنجش از دور چندطیفی بدست آمده‌اند، معمولاً به دلیل استفاده از باندهای طیفی کم و نسبتاً عریض، غیر قابل اعتماد هستند و تنها به منظور شناسایی اولیه می‌توان از آن‌ها بهره برد [۱]. در این راستا ژانگ و همکاران در سال ۲۰۰۷، از تصاویر ASTER و روش طبقه‌بندی نظارت شده‌ی بیشترین شباهت کمک گرفته‌اند و پس از تولید اندکس‌های معدنی و کمک‌گیری از آنالیز مؤلفه‌های اصلی^۱ و همچنین با روش کمینه‌ساز انرژی مقید به تولید نقشه‌ی مواد معدنی پرداختند [۲]. به منظور بررسی توانایی روش نقشه‌بردار زاویه طیفی در سال ۲۰۱۰ گابر و همکاران الگوریتم SAM را روی تصاویر ASTER اجرا نمودند و پس از ارزیابی نتایج مشخص شد که روش نقشه‌بردار زاویه طیفی برای تصاویر چندطیفی مناسب نیست، زیرا با در نظر گرفتن حد‌آستانه معادل با ۰/۱ رادیان تقریباً تمام پیکسل‌ها در کلاس صفر قرار می‌گرفتند [۳]. بدینی در سال ۲۰۱۱ طی تحقیقی طیف مربوط به ۹۰ نمونه کانی مختلف را به کمک طیف سنج زمینی برداشت کرد و روش فیلتر تطبیقی را روی تصاویر ASTER و HyMap اعمال نمود. ضریب کاپا معادل با ۸۳٫۵٪ برای طبقه‌بندی حاصل از تصاویر HyMap به‌دست آمد البته کاستی‌هایی به دلیل عدم وجود داده‌های ژئوفیزیکی وجود داشت و به دلیل قدرت تفکیک مکانی و طیفی پایین سنجنده ASTER خروجی‌ها دقت یالایی نداشتند [۴].

در سال ۲۰۱۴ بورناس و همکاران به کمک تصاویر حاصل از سنجنده‌ی Landsat TM و همچنین با بهره‌گیری از داده‌های آموزشی و ارائه‌ی آن‌ها به الگوریتم‌های شبکه عصبی و طبقه‌بندی‌کننده‌ی بیشترین شباهت به حضور مواد معدنی مختلفی از قبیل طلا، مس، قلع و همچنین مواد رادیواکتیو مثل اورانیوم و توریم پی بردند [۵].

مطالعات نشان داده‌اند که سنجنده‌های فراطیفی ماهواره‌ای قادر به تولید اطلاعات مواد معدنی هستند، ولی نیاز به بهبود نسبت سیگنال به نویز در این گونه داده‌ها وجود دارد تا بتوان نقشه‌هایی تولید نمود که از نظر دقت با نقشه‌های حاصل از سنجنده‌های فراطیفی هوایی در یک سطح باشند. در این راستا حسین‌جانی‌زاده و همکاران در سال ۲۰۱۴ از تصاویر فراطیفی هایپریون استفاده کردند و الگوریتم فیلتر ترکیبی تطبیقی مختلط^۲ را اعمال کردند که پس از ارزیابی نتایج طبقه‌بندی به کمک نمونه‌های برداشت شده زمینی و انجام آنالیزهای طیفی روی آن‌ها، صحت بالای ۷۵٪ در طبقه‌بندی بیشتر کلاسهای مواد معدنی ثابت شد. از طرف دیگر نتایج حاکی از برتری سنجنده هایپریون نسبت به سنجنده‌های چندطیفی مثل ETM و ASTER بود، البته به شرط آن‌که پیش پردازش‌های مناسب روی تصاویر فراطیفی هایپریون اعمال شده باشد [۶]. یکی دیگر از موضوعات تحقیقاتی که سالهاست ذهن محققین رشته‌های مختلف را به خود معطوف نموده، طبقه‌بندی تصاویر فراطیفی به منظور استخراج اطلاعات مواد معدنی است. به این منظور، نیاز به الگوریتم‌هایی که علاوه بر دقت بالا، قادر به حل مسائل با سرعت بیشتری باشند، احساس می‌شود. تاکنون الگوریتم‌های متنوعی به منظور طبقه‌بندی تصاویر فراطیفی ارائه شده است. یکی از این الگوریتم‌ها که به تازگی در زمینه طبقه‌بندی تصاویر فراطیفی به کار گرفته شده است مدل آمیخته‌ی گاوسی^۳ است [۷]. پس از نتایج خوبی که این روش در مورد داده‌های با حجم زیاد ارائه نمود، محققین سنجش از دوری در پی بررسی این روش در مورد تصاویر فراطیفی برآمدند. مدل آمیخته گاوسی زیرمجموعه و یکی از پرکاربردترین مدل‌های آمیخته متناهی^۴ است. مدل آمیخته‌ی متناهی اولین بار در قرن نوزدهم وارد ادبیات آمار شد، [۸] و [۹]. از جمله مدل‌های آمیخته متناهی میتوان به مدل آمیخته پواسن برای گروه بندی اسناد در امر بازیابی اطلاعات و مدل آمیخته فیشر برای برای تحلیل متون و آزمایشهای ژنی اشاره کرد. مشهورترین مدل آمیخته متناهی، مدل آمیخته گاوسی می‌باشد، [۱۰] و [۱۱]. از آنجا که مهمترین بخش یک مدل برآورد پارامترهای آن می‌باشد، روش‌های مختلفی

^۱ Mixture Tuned Mached Filter (MTMF)

^۲ Gaussian Mixture Model (GMM)

^۳ Finite Mixture Model

^۴ Principal Component Analysis (PCA)

آمیخته‌ی گاوسی متناهی تقریب زده شوند، این مدل در مور مالی مورد توجه بسیاری قرار گرفته است.

همچنین مدل آمیخته‌ی گاوسی در نجوم، زیست شناسی، پزشکی و مهندسی نیز کاربرد بسیار ی دارد که برای جزییات بیشتر می‌توان به [۱۵]، [۱۶]، [۱۷] و [۱۸] مراجعه کرد.

در مجموع می‌توان گفت که بسیاری از مقالات مرور شده در مورد اکتشاف مواد معدنی در حوزه‌ی سنجش از دور بر مبنای استفاده از تصاویر چندطیفی بوده و از تصاویر فراطیفی در این حیطه کمتر بهره گرفته شده است. سنجنده‌های چندطیفی باند های محدود و جدا از یکدیگر دارند؛ در حالی که سنجنده‌های فراطیفی علاوه بر داشتن باندهای زیاد، پهنای باند بسیار باریکتری (حدود ۱۰ تا ۲۰ نانومتر) نسبت به سنجنده‌های چندطیفی دارند. از آنجایی که بسیاری از مواد معدنی برحسب ویژگیهای طیفی آنها که معمولا در حدود ۲۰ تا ۴۰ نانومتر پهنای دارند؛ قابل تشخیص از یکدیگر هستند، به سادگی میتوان به اهمیت این نوع از سنجنده‌ها و مزیت آنها نسبت به سنجنده‌های چندطیفی پی برد. از این رو در این تحقیق از تصاویر فراطیفی Hyperion به منظور اکتشاف مواد معدنی و همچنین از روش طبقه‌بندی بدون نظارت مدل آمیخته‌ی گاوسی به منظور تشخیص کانی‌ها از یکدیگر استفاده شده است. در ادامه پس از تشریح الگوریتم مدل آمیخته‌ی گاوسی و روش‌های تخمین پارامتر الگوریتم EM، روند پیش پردازش داده‌ها و داده‌های مورد استفاده جهت پیاده‌سازی الگوریتم معرفی شده است. در ادامه از الگوریتم پیشنهادی برای خوشه‌بندی داده‌های فراطیفی استفاده شده است. در نهایت در بخش نتایج و بحث معیار ارزیابی، معرفی و نتایج حاصل از اجرای الگوریتم‌ها بررسی می‌گردد. در بخش ۶ نیز نتیجه‌گیری تحقیق ارائه می‌شود.

۲- مدل آمیخته‌ی گاوسی

مدل آمیخته‌ی گاوسی یکی از روش‌های نوینی است که به منظور طبقه‌بندی تصاویر سنجش از دور فراطیفی ارائه شده است. استفاده از این الگوریتم به دلیل نتایج رضایت بخش و سرعت نسبتا بالا در حال افزایش است. مدل آمیخته‌ی گاوسی به میزان قابل توجهی در مدل‌سازی تراکم و خوشه‌بندی مورد استفاده قرار می‌گیرد. این مدل‌ها دارای قابلیت تخمین کاملی می‌باشند، چرا که آنها می‌توانند هر نوع تابع تراکم را به صورت دقیقی با تعداد

برای برآورد پارامترهای مدل آمیخته ارائه شده است. دمپستر و همکاران در سال ۱۹۹۷ الگوریتم Expectation Maximization (EM) را ارائه دادند. این الگوریتم روشی برای محاسبه‌ی برآوردگر ماکسیمم درست‌نمایی است، هنگامی که داده گمشده وجود داشته باشد یا روش‌های ساده‌ی بهینه سازی با شکست مواجه شوند [۱۲]. در سال‌های اخیر برای شناسایی گوینده (تشخیص صدا)، در متون مستقل، از مدل‌های آمیخته‌ی گاوسی استفاده می‌کنند. شناسایی گوینده، هنگامی که هیچ پیش فرضی از آنچه گوینده به زبان می‌آورد وجود ندارد، اصطلاحا شناسایی گوینده با متون مستقل گفته می‌شود. برای هر گوینده مدل آمیخته‌ی گاوسی به گونه‌ای در نظر گرفته می‌شود که تابع احتمال پسین‌اش ماکسیمم مقدار شود. رینولد و همکارانش در سال ۱۹۹۵ در مقاله‌ای [۱۳] نشان دادند که مدل آمیخته‌ی گاوسی برای شناسایی گوینده در متون مستقل، مدلی استوار می‌باشد.

علاوه بر تشخیص گوینده، از مدل آمیخته‌ی گاوسی برای شناسایی چهره‌ی افراد نیز استفاده می‌شود [۱۴]. مهم‌ترین مشکل در تشخیص چهره افراد و بازیابی آن، سایه روشن‌ها، تغییرات نور و پس‌زمینه‌های هم‌رنگ است. تشخیص چهره در مسائل امنیتی، تشخیص تغییرات در افراد و فهرست‌گذاری در تصاویر ویدئویی، کاربرد دارد. برای مدل‌بندی رنگ چهره‌ی (پوست) افراد نیز، می‌توان از مدل آمیخته‌ی گاوسی استفاده کرد.

در روش خوشه‌بندی مبتنی بر مدل، که برای مشاهدات، مدلی احتمالاتی در نظر گرفته می‌شود، از مدل آمیخته‌ی گاوسی استفاده می‌شود. بدین صورت که در این روش، هر خوشه به وسیله یک توزیع پارامتری نشان داده می‌شود. آن‌گاه مدلی که برای کل داده‌ها ارائه می‌شود ترکیب آمیخته‌ی متناهی از این توزیع‌ها، می‌باشد. با استفاده از مدل آمیخته‌ی گاوسی، اطلاعات کاملتری درباره‌ی خوشه‌ها به دست می‌آوریم. به دلیل انعطاف‌پذیری مدل آمیخته‌ی گاوسی برای انواع مختلفی از توزیع‌ها، در یافتن الگوهایی برای امور مالی تجربی نیز، از مدل آمیخته‌ی گاوسی استفاده می‌شود. در مدل‌سازی مالی و کاربردهای آن، توزیع نرخ سود (بازده) در دارایی‌های مالی نقش مهمی دارد. متداول‌ترین فرض این است که نرخ سود دارایی‌ها، توزیع گاوسی دارد و از آنجا که دیگر توزیع‌ها نیز می‌توانند به خوبی با یک مدل

کافی تابع گاوسی مدل‌سازی نمایند [۱۹]. این مدل تابع چگالی احتمال متغیرهای مشاهده شده را با استفاده از چگالی‌های آمیخته گاوسی مدل می‌کند.

مدل آمیخته‌ی گاوسی زیرمجموعه مدل‌های آمیخته‌ی متناهی است. یک مدل آمیخته متناهی به صورت زیر تعریف می‌شود [۷]: فرض کنید x_1, \dots, x_n مشاهداتی از نمونه تصادفی مستقل و هم توزیع X_1, \dots, X_n با بردار پارامتر $\theta = (\theta_1, \dots, \theta_k)$ باشند، آن‌گاه تابع چگالی آمیخته مشاهدات به صورت زیر خواهد بود:

$$f(x_i; \theta) = \sum_{k=1}^K \alpha_k f_k(x_i; \theta_k); i=1, \dots, n$$

را تابع چگالی آمیخته‌ی k مولفه‌ای نیز گویند. منظور از مولفه همان زیر جامعه تشکیل دهنده‌ی جامعه می‌باشد که تعدادشان را با k نشان می‌دهند. θ_k پارامتر مربوط به زیر جامعه‌ی k می‌باشد و α_k ضریب وزنی یا ضریب آمیخته‌ی k ام است که در شرایط زیر صدق می‌کند:

$$\sum_{k=1}^K \alpha_k = 1; \quad 0 \leq \alpha_k \leq 1$$

چگالی مولفه‌ی k ام با پارامتر θ_k می‌باشد. مدل آمیخته گاوسی برای مشاهدات مستقل و هم توزیع x_1, \dots, x_n ، مجموع وزندار k مولفه، با تابع چگالی گاوسی است، که با معادله زیر نشان داده می‌شود:

$$f(x_i; \mu, \sigma^2) = \sum_{k=1}^K \alpha_k \phi(x_i; \mu_k, \sigma_k^2)$$

در اینجا $\mu = (\mu_1, \dots, \mu_k)$ و $\sigma^2 = (\sigma_1^2, \dots, \sigma_k^2)$ بردار پارامترهای مدل آمیخته می‌باشند. $\phi(x_i; \mu_k, \sigma_k^2)$ تابع چگالی گاوسی با پارامترهای μ_k و σ_k^2 مربوط به مولفه k ام می‌باشد که از رابطه زیر به دست می‌آید:

$$\phi(x_i; \mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}\right]$$

این مدل تابع چگالی احتمال متغیرهای مشاهده شده را با استفاده از چگالی‌های آمیخته گاوسی مدل می‌کند.

پارامترها و وزن‌های هر توزیع را با دادن یک سری مقادیر ورودی می‌توان با استفاده از الگوریتم‌هایی از جمله الگوریتم EM بدست آورد.

۲-۲- الگوریتم EM

مسئله تخمین پارامترهای مدل آمیخته‌ی گاوسی به طور گسترده‌ای مورد مطالعه قرار گرفته است. اوریت و هند در سال ۱۹۸۱ یک مرور جامع از این موضوع ارائه دادند و چندین روش برای تخمین این پارامترها پیشنهاد کردند [۲۰]. یکی از این روش‌ها الگوریتم EM است. این روش یک روش کلی برای بهینه‌سازی توابع درست‌نمایی است، هنگامی که داده‌ی گمشده وجود داشته باشد یا روش‌های ساده بهینه‌سازی با شکست مواجه شوند [۲۱] و [۲۲]. این روش در سال ۱۹۷۷ توسط دمپستر و همکارانش ارائه شد [۲۳] و در سال ۱۹۸۴ توسط رندر و واکر برای مدل‌های آمیخته به کار گرفته شد [۲۴]. برای استفاده از این الگوریتم باید تعداد مولفه‌های آمیخته C مشخص باشد. تعداد مولفه‌های آمیخته نشان‌دهنده تعداد کلاس‌های طبقه‌بندی است که می‌تون آن را با دانش قبلی از مسئله، روش‌های گرافیکی آنالیز دیتا و یا سایر روش‌های تخمین کلاس از جمله آمیخته سازگار^۲ بدست آورد [۲۵]. اگرچه روش‌هایی وجود دارند که تعداد خوشه‌ها را قبل از انجام خوشه‌بندی محاسبه می‌کنند، اما در این پژوهش برای بررسی هرچه بهتر الگوریتم‌های مورد استفاده، تعداد خوشه‌ها را معلوم در نظر می‌گیریم. علاوه بر تعداد مولفه‌های آمیخته، می‌بایست یک حدس اولیه برای برای مقادیر پارامترهای بردار میانگین، ماتریس کواریانس و ضرایب اختلاط تولید کرد. الگوریتم EM با استفاده از تخمین‌های اولیه و با در نظر گرفتن متغیرهای پنهان از چرخه مکرر برای برآورد پارامترها استفاده می‌کند. این الگوریتم با در نظر گرفتن مقدار اولیه برای پارامترهای مدل شروع می‌شود و در گام بعد که مرحله‌ی تکرار نامیده می‌شود، این پارامترها به روز می‌شوند و چرخه تا جایی تکرار می‌شود که الگوریتم همگرا شود. مرحله‌ی تکرار از دو گام محاسبه امید ریاضی و ماکسیمم‌سازی تشکیل می‌شود. در گام اول احتمالات پسین توسط معادله ۱ محاسبه می‌گردد.

$$\hat{\tau}_{ij} = \frac{\hat{p}_i \Phi(x_j; \hat{\mu}_i, \hat{\Sigma}_i)}{\hat{f}(x_j)} \quad i=1, \dots, c; \quad j=1, \dots, n. \quad (1)$$

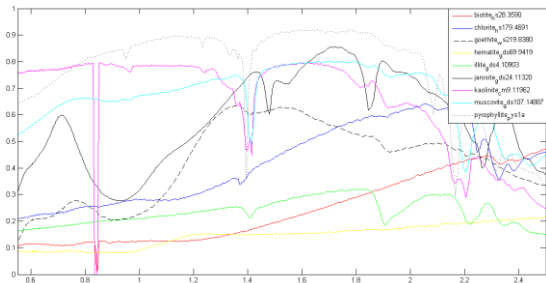
^۲ Adaptive Mixture

^۱ Mixing coefficient

همچنین داده فراطیفی واقعی hyperion از سنجنده EO-1 مورد استفاده قرار گرفت تا کارایی الگوریتم مورد استفاده در این تحقیق را نشان دهند.

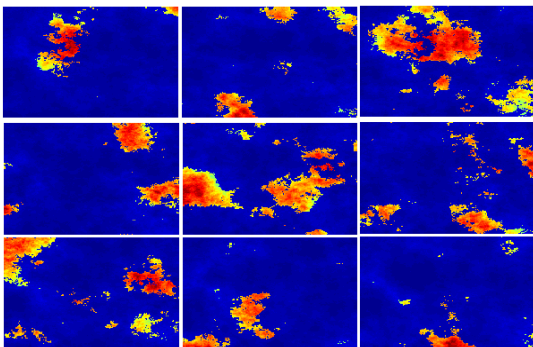
۳-۱- داده‌ی شبیه‌سازی

داده‌های شبیه‌سازی شده از عناصر خالص با مشخصه طیفی و نقشه‌های فراوانی معلوم ساخته شده‌اند. بنابراین از این داده‌ها می‌توان به عنوان ابزاری برای ارزیابی نتیجه الگوریتم‌ها استفاده نمود. داده‌ی شبیه‌سازی شده‌ی مورد استفاده در این پژوهش از ۹ کلاس طیفی با مشخصه‌های طیفی معلوم تشکیل شده است. امضای طیفی این عناصر که از کتابخانه طیفی ایالات متحده انتخاب شده‌اند، در شکل ۱ آورده شده است.



شکل ۱- مشخصه طیفی ۹ عنصر خالص

با اعمال وزن‌های مختلف برای عناصر شکل ۱، پیکسل‌های مختلط ایجاد شده‌اند. وزن‌های اعمال شده برای هر عنصر خالص در هر پیکسل، نقشه فراوانی عنصر مربوط می‌باشد. نقشه‌های فراوانی توسط جعبه ابزار شبیه‌سازی که تحت عنوان پروژه HYDRA انجام شده است، تولید شده‌اند. این جعبه ابزار که قابل دسترس در [۲۶] می‌باشد، تمامی ویژگی‌های فراوانی عناصر خالص، اعم از نامنفی بودن فراوانی‌ها و نیز برابر یک بودن مجموع فراوانی‌ها در هر پیکسل را دارا می‌باشد. در شکل ۲ نقشه‌های فراوانی ۹ عنصر به کار گرفته شده برای تولید داده شبیه‌سازی شده آورده شده است.



شکل ۲- نقشه‌های فراوانی ۹ عنصر خالص

در رابطه بالا $\hat{\tau}_{ij}$ احتمال پسین برای پیکسل j ام می‌باشد و نشان‌دهنده میزان تعلق پیکسل j به مولفه‌ی i ام است. $\Phi(x_j; \hat{\mu}_i, \hat{\Sigma}_i)$ چگالی نرمال مولفه‌ی i ام و $\hat{f}(x_j)$ نیز تخمین آمیخته‌ی گاوسی برای پیکسل j ام است که رابطه آن در زیر آمده است.

$$\hat{f}(x_j) = \sum_{k=1}^c p_k \Phi(x_j; \hat{\mu}_k, \hat{\Sigma}_k) \quad (2)$$

در گام بعد پارامترهایی انتخاب می‌شوند که بر اساس آن‌ها امید ریاضی بدست آمده از مرحله قبل ماکسیمم مقدار شود. در این مرحله پارامترهای مدل بر حسب $\hat{\tau}_{ij}$ که از گام قبل بدست آمده، محاسبه می‌شود.

$$\hat{p}_i = \frac{1}{n} \sum_{j=1}^n \hat{\tau}_{ij} \quad (3)$$

$$\hat{\mu}_i = \frac{1}{n} \sum_{j=1}^n \frac{\hat{\tau}_{ij} x_{ij}}{\hat{p}_i} \quad (4)$$

$$\hat{\Sigma}_i = \frac{1}{n} \sum_{j=1}^n \frac{\hat{\tau}_{ij} (x_j - \hat{\mu}_i)(x_j - \hat{\mu}_i)^T}{\hat{p}_i} \quad (5)$$

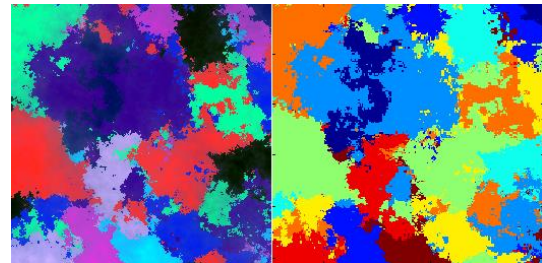
بطور خلاصه روند اجرایی الگوریتم EM را به صورت زیر بیان کرد:

- ۱- تعیین تعداد چگالی‌های گوسین c در مدل آمیخته.
- ۲- تولید مقادیر اولیه‌ی پارامترهای مولفه‌ی آمیخته. این پارامترها شامل ضریب اختلاط، بردار میانگین و ماتریس کواریانس برای هر مولفه‌ی آمیخته می‌باشد.
- ۳- محاسبه امید ریاضی. در این مرحله احتمال متعلق بودن مشاهده j ام به مولفه i ام محاسبه می‌شود.
- ۴- ماکسیمم‌سازی. در این مرحله پارامترهای مدل بر حسب $\hat{\tau}_{ij}$ که از گام سوم بدست آمده، محاسبه می‌شود.
- ۵- تکرار گام ۳ و ۴ تا رسیدن به همگرایی در برآورد پارامترهای مدل.

۳- داده‌های مورد استفاده

جهت ارزیابی دقت الگوریتم‌های پیشنهادی در این تحقیق، ابتدا از یک داده شبیه‌سازی شده استفاده شد.

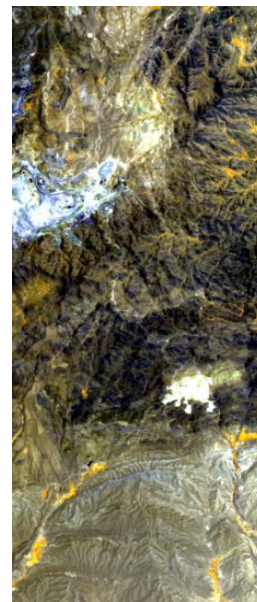
همچنین ترکیب رنگی داده مورد استفاده و نقشه واقعیت زمینی آن در شکل ۳ نمایش داده شده است. این داده شامل یک مکعب داده با ابعاد 256×256 در 240 باند طیفی است.



شکل ۳- الف) نقشه واقعیت زمینی ب) نقشه رنگی کاذب داده شبیه سازی

۳-۲- داده‌ی واقعی

این داده توسط ماهواره EO-1 از منطقه‌ی جنوب غربی شهر کرمان اخذ شده است. توان تفکیک مکانی این تصویر 30 متر و ابعاد تصویر 700 سطر در 254 ستون است. داده LIR هایپریون شامل 242 باند طیفی می‌باشد که فقط 198 باند آن دارای مقادیر غیر صفر است. در شکل ۴ ترکیب رنگی کاذب تصویر مورد مطالعه آورد شده است.



شکل ۴- تصویر رنگی کاذب HYPERION

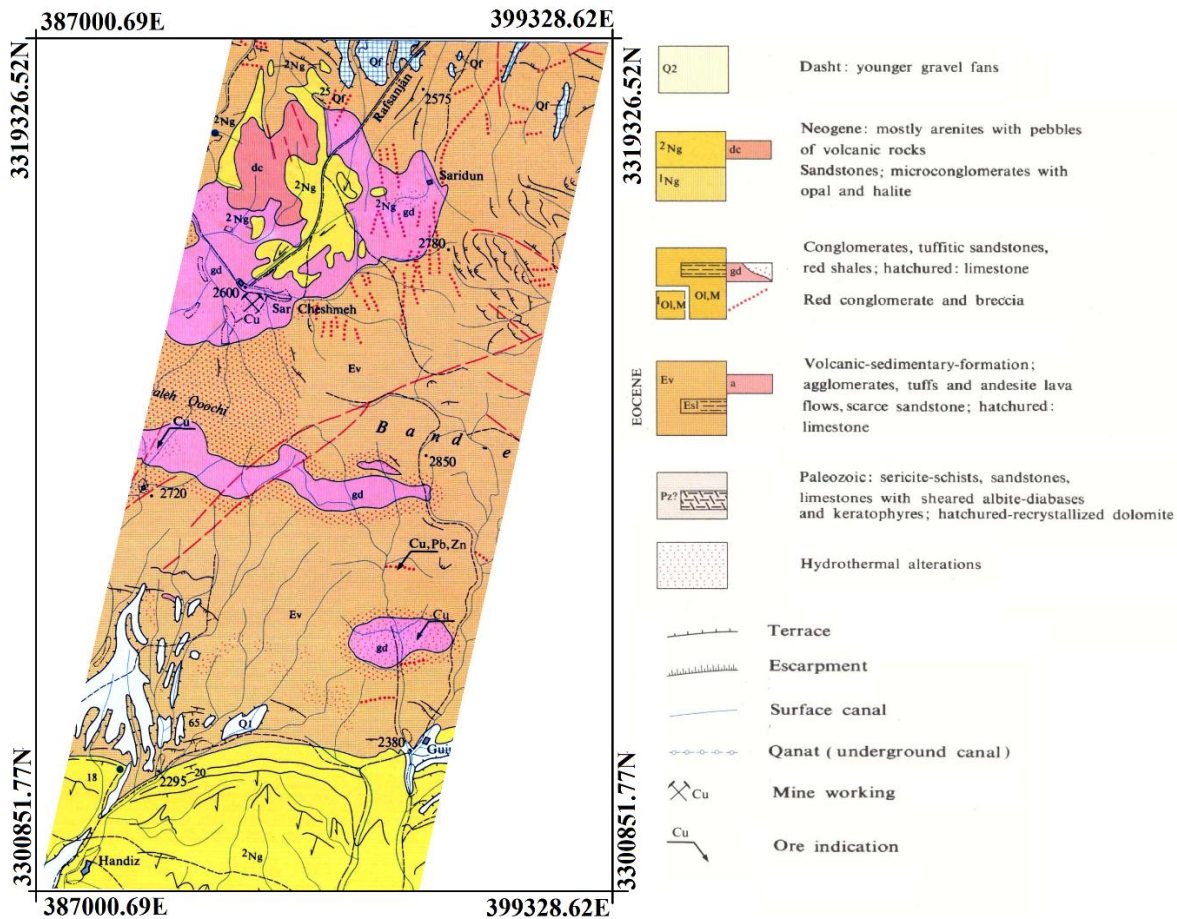
۳-۲-۱- زمین‌شناسی منطقه مورد بررسی

منطقه مورد مطالعه این تحقیق بخشی از کمربند ماگمایی اورمیة-دختر با سن ترشیری است. براساس

مطالعاتی که تاکنون انجام شده این بخش یکی از غنی‌ترین نواحی از نظر کمربند کانی سازی مس نوع پورفیری در ایران محسوب می‌شود. این منطقه در بخش جنوبی زون اورمیة-دختر، جنوب غربی شهر کرمان در مرکز ایران قرار دارد و ارتفاع متوسط منطقه از سطح دریا 2500 متر می‌باشد [۲۶]. در این منطقه ذخایر مس پورفیری برجسته‌ای همانند معادن مس سرچشمه، دره زار و سریدون، میدوک، محسوب می‌شود. این منطقه در بخش جنوبی زون اورمیة-دختر، جنوب غربی شهر کرمان در مرکز ایران قرار دارد و ارتفاع متوسط منطقه از سطح دریا 2500 متر می‌باشد [۲۷]. در این منطقه ذخایر مس پورفیری برجسته‌ای همانند معادن مس سرچشمه، دره زار و سریدون، میدوک، سونگون و غیره حضور دارند. منطقه مورد مطالعه در بین طول‌های جغرافیایی 55 درجه و 51 دقیقه تا 55 درجه و 57 دقیقه و عرض‌های جغرافیایی 29 درجه و 53 دقیقه تا 30 درجه و 50 دقیقه محصور شده است. طبق مطالعات انجام شده و بر اساس نقشه‌های زمین‌شناسی و معدنی، مناطق آلتراسیون در منطقه سرچشمه شامل پتاسیک، بیوتیت، فیلیک، آرژیلیک و پروپیلیتیک هستند [۶]. قابل ذکر است که به طور معمول تمام مناطق آلتراسیون مربوط به مس پورفیری از چنین الگویی تبعیت می‌کنند.

گونه‌های دگرسانی از قبیل پروپیلیتیک، فیلیک، آرژیلیک، سیلیسیک و آرژیلیک پیشرفته در محدوده مورد بررسی مشاهده شده‌اند.

منطقه دره زار نیز به‌طور کلی از سنگ‌های دیوریت و گرانودیوریت تشکیل شده است و از لحاظ کانی‌شناسی مشتمل بر پتاسیک، فیلیک، آرژیلیک و پروپیلیتیک می‌باشد. نقشه زمین‌شناسی منطقه مورد مطالعه در شکل شماره ۵ آورده شده است.



شکل ۵- نقشه زمین‌شناسی منطقه مورد مطالعه

منطقه هم‌پوشی دو طیف‌سنج مرئی و مادون‌قرمز قرار دارند که باندهای ۷۷ و ۷۸ به دلیل سطح نویز بالا حذف شدند. در نهایت پس از حذف کردن باندهای مربوط به جذب بخار آب ۱۶۳ باند به منظور انجام سایر پردازش‌ها باقی ماند [۲۸].

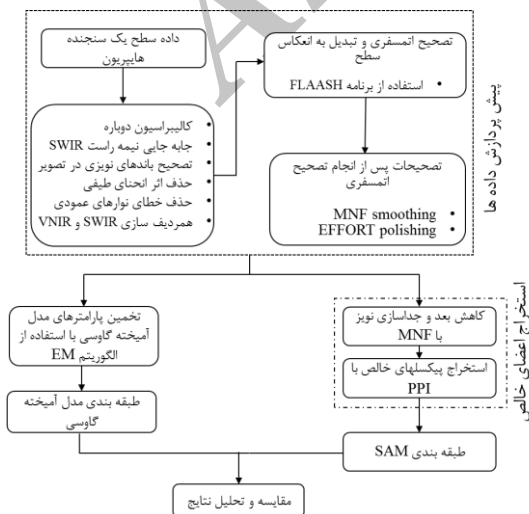
۴- روش پیشنهادی

۴-۱- پیش پردازش داده هایپریون

داده مورد استفاده در این پژوهش تصویر هایپریون در سطح پردازشی LIR است، که در این سطح داده، فقط تصحیح رادیومتری انجام شده است. رویکرد پیش‌پردازش شامل مراحل است که در ادامه بررسی می‌شود. قابل ذکر است که انجام تصحیح هندسی به دلیل تغییر دادن درجات روشنایی پیکسل‌ها، پس از اعمال الگوریتم‌های پردازشی مربوطه و کسب نتایج طبقه بندی، به کمک تصاویر هایپریون سطح LIGst که از نظر هندسی تصحیح شده بودند، انجام شد.

۴-۱-۱- حذف باندهای کالیبره نشده و جذب آب

داده LIR هایپریون شامل ۲۴۲ باند طیفی می‌باشد که فقط ۱۹۸ باند آن دارای مقادیر غیر صفر است. به‌علاوه در بین باندهای غیر صفر چهار باند ۵۶، ۵۷، ۷۷ و ۷۸ در



شکل ۶- نمای کلی پیاده‌سازی روش پیشنهادی

۴-۱-۲- تصحیح جابجایی مکانی

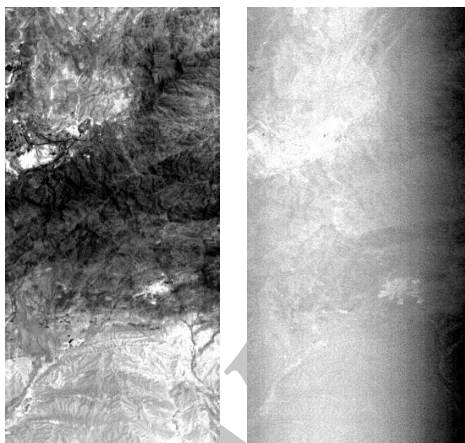
در تصاویر هایپریون از نظر مکانی به اندازه یک پیکسل جابه‌جایی در محدوده مرئی در عرض تصویر به سمت راست اتفاق می‌افتد. که برای تصحیح آن، به اندازه یک پیکسل داده را به سمت چپ جابه‌جا می‌کنیم. با انجام این کار ستون اول که دارای مقادیر صفر بود به ستون آخری منتقل می‌شود و چون این ستون ارزشی ندارد، می‌توان آن را حذف کرد. همچنین در محدوده SWIR ستون آخری دارای مقادیر صفر بوده و همانند محدوده مرئی، این ستون را نیز حذف می‌کنیم تا از نظر هندسی دو محدوده طیفی سازگار باشند. همچنین در محدوده SWIR ستون آخری دارای مقادیر صفر بوده و همانند محدوده مرئی، این ستون را نیز حذف می‌کنیم تا از نظر هندسی دو محدوده طیفی سازگار باشند و همچنین در SWIR ستون ۱۲۷ به بعد، جابجایی یک پیکسلی در راستای طول تصویر رخ می‌دهد که برای رفع این مشکل ستون ۱ تا ۱۲۸ را از ستون ۱۲۹ تا ۲۵۶ با subset مکانی جدا کرده و دوباره با موزائیک کردن دو محدوده، این جابجایی نیز برطرف می‌شود [۲۹].

۴-۱-۳- تصحیح خطای نوار نوار شدگی

در سنجنده‌های پوش‌بروم از قبیل هایپریون خطای نوار نوار شدگی به دلیل کالیبراسیون نامناسب آشکارسازها و نوسان‌های زمانی در پاسخ آن‌ها، پدید می‌آید. این خطا در راستای حرکت سنجنده بوده و به صورت نوارهایی پیوسته یا متناوب در طول تصویر به عنوان نویز نواری ظاهر می‌شود که به پیکسل‌های چنین نوارهایی، پیکسل‌های غیر عادی اطلاق می‌شود. طبق مطالعات هان و همکاران پیکسل‌های غیر عادی به صورت نوارهای پیوسته با مقادیر درجه روشنایی غیر عادی و خیلی پایین بوده که معمولاً دارای مقدار صفر هستند و تحت عنوان Bad line ها از آن‌ها یاد می‌شود و یا به‌صورت نوارهای پیوسته و متناوبی نمایان می‌شوند که مقادیر پیکسل‌هایشان نسبت به ستون‌های مجاور خیلی پایین است. در این پژوهش به منظور یافتن Bad line ها از کد نویسی در محیط متلب یاری گرفته شد و مکان هر Bad line در هر باند مشخص شد و سپس به کمک میانگین‌گیری از ستون‌های مجاور این خطا برطرف گردید. در ادامه برای حذف خطای حالت دوم که منجر به راه راه شدن تصویر می‌شد از ابزار vertical stripe removal موجود در نرم‌افزار ENVI بهره گرفته شد [۲۸].

۴-۱-۴- تصحیح خطای انحنای طیفی

یکی دیگر از مشکلات و خطاهایی که در پردازش تصاویر هایپریون ایجاد اشکال می‌کند، اثر انحنای طیفی می‌باشد. اثر انحنای طیفی نوعی شیفت طیفی می‌باشد که به دلیل فاصله گرفتن طول موج های across track از طول موج مرکزی پدید می‌آید. اثر مذکور به‌صورت مجزا در تک تک باندها آشکار نیست بنابراین برای بررسی این اثر روی داده‌های مورد استفاده، شاخصی لازم خواهد بود. از طرفی برآورد مقدار smile هایپریون به دلیل متغیر بودن آن از یک صحنه به صحنه دیگر قبل از پرتاب ممکن نیست. یکی از روش‌های متداول برای اثبات حضور اثر انحنای طیفی در تصاویر هایپریون، وجود یک گرادپان روشنایی در تصویر با مقدار ویژه اول در فضای MNF می‌باشد. همانطور که از شکل (۷-الف) پیداست داده هایپریون دارای خطای انحنای طیفی می‌باشد که به منظور رفع خطای یاد شده از ابزار Cross track illumination correction موجود در نرم افزار ENVI یاری گرفته شد [۳۰].

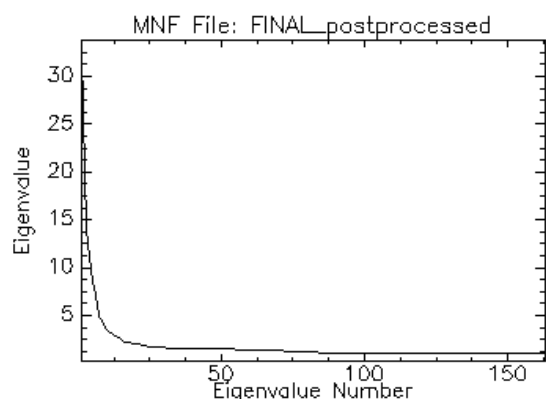


شکل ۷- الف) تصویر MNF1 قبل از حذف خطای Smile
ب) تصویر MNF1 بعد از حذف خطای Smile

۴-۱-۵- تصحیح خطای اتمسفری

انجام تصحیح اتمسفری یک مرحله پیش‌پردازشی مهم در کسب اطلاعات درباره سطح زمین است. روش‌های ساده و پیچیده فراوانی وجود دارند که هدفشان تصحیح اثرات اتمسفر روی تصاویر است. نرم افزارهای FLAASH، ACORN، ATCOR از نرم افزارهای تصحیح اتمسفری هستند که از کد انتقال تابش MODTRAN برای برآورد پارامترهای اتمسفری استفاده می‌کنند [۳۱]. در این

اندیس خلوص پیکسلی^۲ به منظور استخراج پیکسل‌های خالص استفاده شد [۶]. اندیس خلوص پیکسلی از طریق تکرار و تصویر پلات پراکندگی n بعدی بر روی یک بردار واحد با امتداد تصادفی محاسبه می‌شود. تعداد کل دفعاتی که هر پیکسل به عنوان پیکسل انتهایی ظاهر می‌شود محاسبه شده و در نهایت یک تصویر خلوص پیکسلی ایجاد می‌گردد که در آن ارزش هر پیکسل متناسب با تعداد دفعاتی است که آن پیکسل به عنوان پیکسل انتهایی ظاهر شده است. از آن جایی که این روش معمولاً بر روی نتیجه MNF که در آن باندهای حاوی نوفه حذف شده‌اند اجرا می‌شود، ابتدا الگوریتم MNF که در حقیقت یک روش اصلاح شده آنالیز مؤلفه‌های اصلی می‌باشد، بر روی تصاویر پیش پردازش شده هایپریون اعمال شد و سپس با روش PPI پیکسل‌هایی که از نظر طیفی خالص بودند، از ۲۰ تصویر اول MNF مشخص شدند.



شکل ۹- نمایش مقادیر ویژه حاصل از اعمال تبدیل MNF

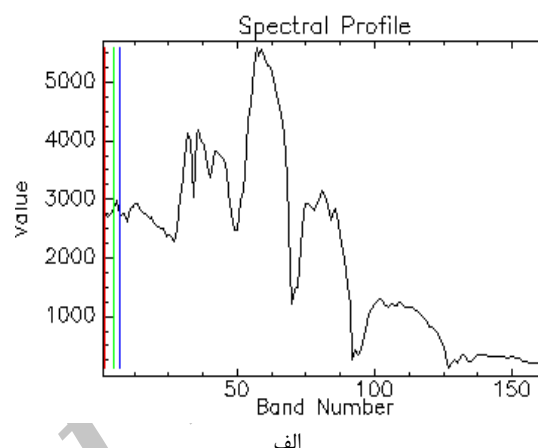
۳-۴- اعمال الگوریتم شناسایی هدف

این بخش در دو مرحله صورت می‌گیرد. در مرحله اول از الگوریتم SAM به منظور شناسایی هدف استفاده شده است. پس از استخراج طیف‌های مربوطه، الگوریتم مورد نظر روی تصویر پیش پردازش شده هایپریون اعمال شد. اجرای این الگوریتم با استفاده از نرم افزار ENVI صورت گرفته است.

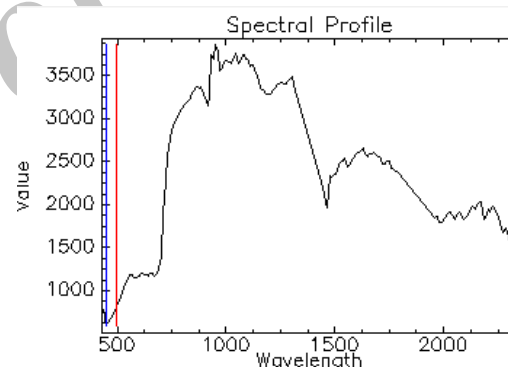
در مرحله دوم از مدل آمیخته‌ی گاوسی برای خوشه بندی کانی‌ها استفاده می‌شود. تخمین پارامترهای این مدل با استفاده از الگوریتم EM انجام گرفته است. برنامه نویسی و اجرای این الگوریتم در نرم افزار متلب (MATLAB) انجام شده است.

^۲ Pixel Purity Index (PPI)

پژوهش از برنامه FLAASH موجود در نرم‌افزار ENVI به جهت دستیابی به رفلکتانس استفاده شده است [۳۲]. به منظور بررسی نتیجه اعمال تصحیح مذکور، طیف پوشش گیاهی مورد ارزیابی قرار گرفت. با توجه به نشانه طیفی پوشش گیاهی بعد از تصحیح که مشابه با نشانه طیفی کتابخانه‌ای پوشش گیاهی بوده، بنابراین الگوریتم FLAASH در تصحیح اتمسفری منطقه موثر بوده است.



الف



ب

شکل ۸- الف) طیف پوشش گیاهی قبل از تصحیح اتمسفری (ب) طیف پوشش گیاهی بعد از تصحیح اتمسفری

۴-۲- استخراج نشانه طیفی هدف

برای اجرای آن دسته از روش‌های آشکارسازی هدف که نیاز به اطلاعات طیفی دارند، باید امضای طیفی هدف موردنظر مشخص باشد. برای این منظور روش‌های مختلفی از قبیل استفاده از طیف سنج و برداشت‌های میدانی، یاری گرفتن از کتابخانه طیفی، تفسیر بصری و کمک گرفتن از روش یافتن پیکسل‌های خالص وجود دارد. در بین روش‌های موجود در این پژوهش از دو پردازش حداقل سهم نویز^۱ و

^۱ Minimum Noise Fraction (MNF)

۵- نتایج و بحث

همانگونه که در بخش ۴ توضیح داده شد، الگوریتم‌های پیشنهادی در دو مرحله بر روی داده‌های مورد نظر پیاده‌سازی می‌شوند. در مرحله اول الگوریتم SAM روی داده‌های فراطیفی اعمال شد و سپس از مدل آمیخته‌ی گاوسی به منظور خوشه‌بندی اطلاعات معدنی استفاده شده است. برای ارزیابی نتایج روش‌های پیشنهادی، به منظور برآورد صحت و دقت روش خوشه‌بندی و شناسایی هدف از ماتریس خطا به کمک داده‌های واقعیت زمینی استفاده شد.

اگرچه استفاده از ماتریس خطا می‌تواند معیارهای مختلفی را برای ارزیابی طبقه‌بندی ارائه کند، در این پژوهش از مولفه‌های ضریب کاپا و دقت کلی که در رابطه‌های ۶ و ۷ آورده شده است، به عنوان پرکاربردترین معیارهای ارزیابی، مورد استفاده قرار می‌گیرند.

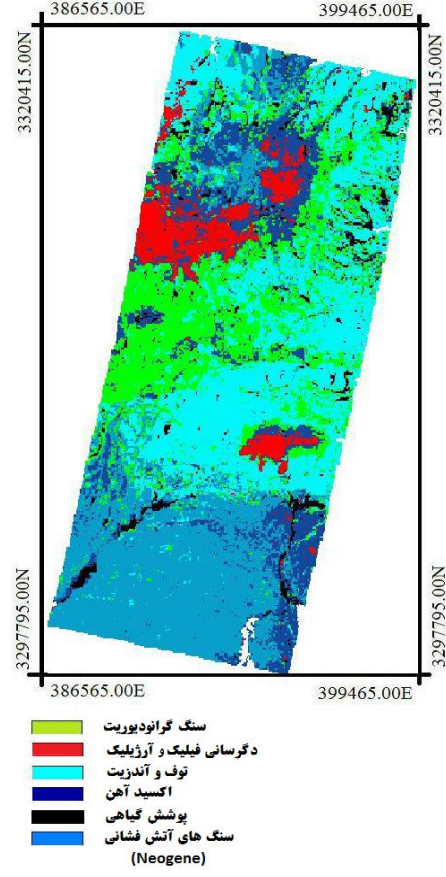
$$K_{hat} = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} \times x_{+i})} \quad (6)$$

$$OA = \frac{\sum_{i=1}^r x_{ii}}{N} \quad (7)$$

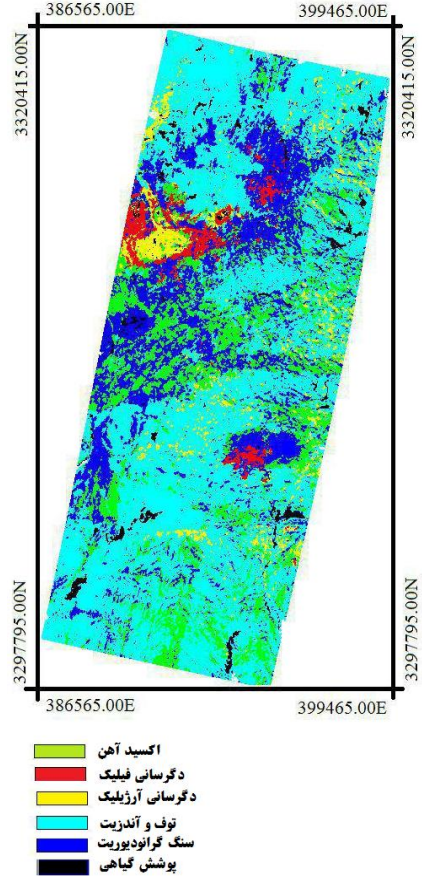
در جدول ۱ میزان دقت کلی و ضریب کاپا مربوط به الگوریتم‌های پیشنهادی، برای داده فراطیفی هایپریون و داده شبیه‌سازی شده آورده شده است.

جدول ۱- مقادیر پارامترهای دقت کلی و ضریب کاپا بر روی دو داده فراطیفی و شبیه‌سازی

الگوریتم اجرا شده	داده تصویری	دقت کلی (درصد)	ضریب کاپا
GMM-EM	شبیه سازی	۱۰۰	۱
	hyperion	۸۰.۰۰	۰.۷۱
SAM	شبیه سازی	۹۸.۱۲	۰.۹۸
	hyperion	۸۲.۰۰	۰.۷۵



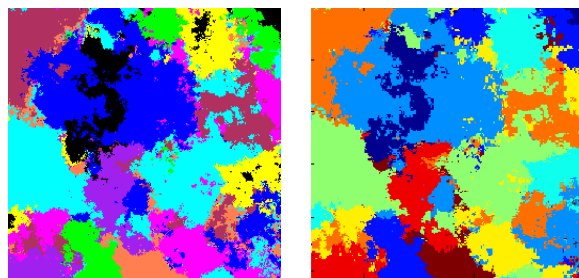
شکل ۱۰- نقشه طبقه‌بندی شده با استفاده از مدل آمیخته گاوسی



شکل ۱۱- نقشه طبقه‌بندی شده با استفاده از الگوریتم SAM

بطور متوسط در نواحی مرکزی مس سرچشمه بالاتر از سریدون می‌باشد و در نتیجه این منطقه در تصویر طبقه بندی شده نهایی در کلاس جداگانه‌ای قرار گرفته است. در تصویر خروجی مدل آمیخته گاوسی به دلیل ماهیت غیر نظارت شده آن، کلاس‌های کلی‌تر از هم متمایز شده‌اند که منجر به تفکیک کلاس توف و آندزیت از سنگ‌های آتش فشانی (Ng) شده است، در صورتی که چنین تمایزی در خروجی الگوریتم SAM دیده نمی‌شود. از سوی دیگر به دلیل نواحی جذبی مشابه کانی‌های آلتراسیون فیلک و آرژیلیک (نواحی جذبی ۲۲۱۰-۲۱۷۰) الگوریتم مدل آمیخته گاوسی این دو ناحیه را با هم به عنوان یک کلاس مشخص کرده، اما در روش SAM به دلیل معرفی طیف‌های جداگانه برای هر یک از دو آلتراسیون مذکور سعی بر تفکیک این دو ناحیه از هم شده است.

متاسفانه ارزیابی دقیق‌تر بدلیل عدم وجود داده‌های با وضوح بالاتر و همینطور عدم دسترسی به نقشه زمین شناسی دقیق از مناطقی مثل منطقه دره زار، سریدون و سرچشمه امکان پذیر نبود. به عبارت دیگر به دلیل اندازه پیکسل‌های بزرگ سنجنده هایپریون و در نتیجه وجود پیکسل‌های مختلط، در طبقه بندی نهایی بعضی کلاس‌ها در هم ادغام شده و به درستی از هم تفکیک نمی‌شوند. همچنین قابل ذکر است که عملیات زمین مرجع کردن نیز به دلیل دقت مکانی پایین سنجنده هایپریون با خطا همراه می‌باشد. در نهایت، پیشنهاد می‌شود به منظور بهبود خروجی‌ها و کاربرد بهتر تصاویر ماهواره‌ای در زمین شناسی و معدن، وضوح مکانی نیز به موازات وضوح طیفی بالا رود که برای رسیدن به این هدف می‌توان از روش‌های تلفیق تصاویر ماهواره‌ای با دقت مکانی بالا و دقت طیفی بالا بهره گرفت. به علاوه، استفاده از داده‌های ژئوفیزیکی نیز می‌تواند کمک کننده باشد.



شکل ۱۲- الف) نقشه طبقه بندی شده با استفاده از مدل آمیخته گاوسی (ب) نقشه طبقه بندی شده با استفاده از الگوریتم SAM

۶- نتیجه گیری

بهره گیری از داده‌های دورسنجی به عنوان تنها منبع برای شناخت دگرسانی‌ها که یکی از مهمترین عوامل کنترل کننده و شناخت مناطق پتانسیل دار معدنی به شمار می‌آیند، بسیار حائز اهمیت است و استفاده از این داده‌ها به منظور جدا کردن واحدهای زمین شناسی کمک قابل توجهی در حذف خطاها و دقیق تر نمودن داده‌های مکانی نقشه‌های زمین شناسی دارد. در این پژوهش توانایی سنجنده فراطیفی هایپریون و الگوریتم‌های SAM و مدل آمیخته گاوسی به منظور نگاشت مناطق آلتراسیون مورد آزمایش قرار گرفت. با استفاده از امضاهای طیفی استخراج شده به عنوان نماینده کلاس‌های مختلف، کلاس‌های طیفی دگرسانی فیلک، آرژیلیک و پوشش گیاهی، به کمک الگوریتم نقشه بردار زاویه طیفی، برای منطقه به خوبی تشخیص داده شدند. نواحی آتره مشخص شده نشان می‌دهد که نواحی اطراف پیت معدن مس سرچشمه از نظر لیتولوژی و سطح استخراجی با دره زار یکسان هستند. از طرف دیگر، چون مس سرچشمه در فاز استخراج بوده از سطح استخراجی هوازده گذشته و به سطح سولفیدی رسیده بنابراین کانی های اکسیدی آهن در آن نسبت به سریدون که در سطح هوازده قرار دارد کاهش یافته و در نتیجه سطح انعکاس

مراجع

- [1] Van der Meer, Freek D., et al. "Multi-and hyperspectral geologic remote sensing: A review." *International Journal of Applied Earth Observation and Geoinformation* 14.1 (2012): 112-128.
- [2] Zhang, Xianfeng, Micha Pazner, and Norman Duke. "Lithologic and mineral information extraction for gold exploration using ASTER data in the south Chocolate Mountains (California)." *ISPRS Journal of Photogrammetry and Remote Sensing* 62.4 (2007): 271-282.
- [3] Gabr, Safwat, Abduwasit Ghulam, and Timothy Kusky. "Detecting areas of high-potential gold mineralization using ASTER data." *Ore Geology Reviews* 38.1 (2010): 59-69
- [4] Nasreddine Bournas, Nasreddine, et al. "Integrating Airborne Geophysics with Landsat TM Imagery and Geology for Assessing the Mineral Potential of the lullemeden Basin, Niger." 2014 SEG Annual Meeting. Society of Exploration Geophysicists, 2014.

- [5] Bedini, Enton. "Mineral mapping in the Kap Simpson complex, central East Greenland, using HyMap and ASTER remote sensing data." *Advances in Space Research* 47.1 (2011): 60-73.
- [6] Zadeh, Mahdieh Hosseinjani, et al. "Sub-pixel mineral mapping of a porphyry copper belt using EO-1 Hyperion data." *Advances in Space Research* 53.3 (2014): 440-451.
- [7] Martinez, Wendy L., and Angel R. Martinez. *Computational statistics handbook with MATLAB*. Vol. 22. CRC press, 2007.
- [8] Pearson, K. (1894). *Contributions to the mathematical theory of evolution source*, Philosophical Transactions of the Royal Society of London, 185, 71-110.
- [9] Melnykov, V. and Maitra, R. (2010). Finite mixture models and model-based clustering. *Statistics Surveys*, 4, 80–116.
- [10] Greenspan, H., Goldberger, J. and Eshe, I. (2001). Mixture model for face-color modeling and segmentation. *Pattern Recognition Letters*, 22, 1525-1536.
- [11] Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). *Statistical Analysis of Finite of Mixture Distributions*. Willey, New York.
- [12] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1-38.
- [13] Reynolds, D.A. and Rose, R.C. (1995), Robust text-independent speaker identification using Gaussian mixture speaker models, *IEEE Transaction on Speech Audio Process*, 3 72–83.
- [14] Reynolds, D.A., Quatieri, T.F. and Dunn R.B. (2000). Speaker verification using adapted Gaussian mixture models. *Digital Signal Processing*, 10, 19–41.
- [15] Greenspan, H., Goldberger, J. and Eshe, I. (2001). Mixture model for face-color modeling and segmentation. *Pattern Recognition Letters*, 22, 1525-1536.
- [16] Titterington, D.M., Smith, A.F.M. and Makov, U.E. (1985). *Statistical Analysis of Finite of Mixture Distributions*. Willey, New York.
- [17] McKenna, S., Gong, S. and Raja, Y. (1998). Modelling facial color and identity with Gaussian mixtures. *Pattern Recognition*, 31, 1883-1892.
- [18] Kon, S. (1984). Models of stock returns a comparison. *The Journal of Finance*, 39, 147-165.
- [19] Zivkovic, Zoran. "Improved adaptive Gaussian mixture model for background subtraction." *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Vol. 2. IEEE, 2004.
- [20] Everitt, B. S. and D. J. Hand. 1981. *Finite Mixture Distributions*, London: Chapman and Hall.
- [21] M. Figueiredo and A. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, Mar. 2002
- [22] B. A. S. Hasan, and Q. J. Gan, "Sequential EM for unsupervised adaptive Gaussian mixture model based classifier," in *Proc. 6th Int. Conf. Machine Learning and Data Mining in Pattern Recognition*, Leipzig, Germany, 2009, pp. 96-106.
- [23] Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. "Maximum likelihood from incomplete data via the EM algorithm (with discussion)," *Journal of the Royal Statistical Society: B*, 39: pp. 1-38.
- [24] Redner, A. R. and H. F. Walker. 1984. "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, 26: pp. 195-239.
- [25] Preibe, C. E. 1994. "Adaptive mixture density estimation," *Journal of the American Statistical Association*, 89: pp. 796-806.
- [26] http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB
- [27] Abedi, Maysam, Ali Gholami, and Gholam-Hossain Norouzi. "A stable downward continuation of airborne magnetic data: A case study for mineral prospectivity mapping in Central Iran." *Computers & Geosciences* 52 (2013): 269-280.
- [28] Datt, Bisun, et al. "Preprocessing EO-1 Hyperion hyperspectral data to support the application of agricultural indexes." *Geoscience and Remote Sensing*, IEEE Transactions on 41.6 (2003): 1246-1259.
- [29] Farooq, S., and H. Govil. "Mapping Regolith and Gossan for Mineral Exploration in the Eastern Kumaon Himalaya, India using hyperion data and object oriented image classification." *Advances in Space Research* 53.12 (2014): 1676-1685.
- [30] Dadon, Alon, Eyal Ben-Dor, and Arnon Karnieli. "Use of derivative calculations and minimum noise fraction transform for detecting and correcting the spectral curvature effect (smile) in Hyperion images." *Geoscience and Remote Sensing*, IEEE Transactions on 48.6 (2010): 2603-2612.

- [31] Kruse, F. A. "Comparison of ATREM, ACORN, and FLAASH atmospheric corrections using low-altitude AVIRIS data of Boulder, CO." Summaries of 13th JPL Airborne Geoscience Workshop, Jet Propulsion Lab, Pasadena, CA. 2004.
- [32] HYPERION LEVEL 1GST (L1GST) PRODUCT OUTPUT FILES DATA FORMAT CONTROL BOOK (DFCB)

Archive of SID