

جداسازی طیفی و مکانی تصاویر ابرطیفی با استفاده

از PCA و تبدیل Semi-NMF

حبیب علیزاده و حسن قاسمیان

گروه مخابرات، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران

چکیده

در سال‌های اخیر جداسازی داده‌های سنجش از دور با استفاده از تجزیه ماتریس نامنفی (Nonnegative Matrix Factorization) مورد توجه قرار گرفته است و برای بهبود کارایی آن، به تابع هزینه اقلیدسی قیدهای کمکی می‌افزاید. چالش اصلی در این میان معرفی قیدهایی است که بتواند نتایج بهتری را استخراج کند. همیستگی بین باندهای تصاویر ابرطیفی مسئله‌ای است که کمتر مورد توجه الگوریتم‌های جداسازی قرار گرفته است. این مسئله در جداسازی کلاس‌های مشابه بیشتر مشخص می‌شود. در این مقاله یک روش جدید برای جداسازی داده‌های ابرطیفی سنجش از دور با استفاده از تجزیه ماتریس نیمه‌نامنفی (Semi-NMF) و تحلیل مؤلفه‌های اصلی پیشنهاد کردیم. در روش پیشنهادی جداسازی طیفی و مکانی به صورت همزمان انجام می‌شود و تمام محدودیت‌های فیزیکی مسئله برآسان مدل مخلوط خطی اعمال می‌شود. همچنین، علاوه بر محدودیت‌های فیزیکی، از ویژگی داده‌های ابرطیفی در فرآیند جداسازی بهره‌برداری شده است. تنکی ضرایب فراوانی یکی از ویژگی‌های مهم داده‌های ابرطیفی است که در این مقاله ما با استفاده از ماتریس ALS به درون تابع هزینه اعمال گرداییم. در روش پیشنهادی قواعد به روزرسانی الگوریتم با استفاده از روش ALS به دست آمده است. در بخش انتهاي اين مقاله از داده‌های ابرطیفی مصنوعی و واقعی به منظور بررسی کارآمدی الگوریتم پیشنهادی استفاده شده است. نتایج به دست آمده برتری الگوریتم پیشنهادی را در مقایسه با برخی از الگوریتم‌ها جداسازی نشان می‌دهد.

وازگان کلیدی : تصاویر ابرطیفی، جداسازی داده‌های سنجش از دور، جداسازی کور منابع ، تجزیه ماتریس نیمه‌نامنفی ، تجزیه و تحلیل مؤلفه‌های اصلی.

پیش از بهره‌برداری از این تصاویر در زمینه‌های مختلف علمی و کاربردی، باید چنین پیکسل‌هایی به مواد تشکیل‌دهنده خود تفکیک شوند. جداسازی داده‌های ابرطیفی قصد دارد چنین پیکسل‌هایی را طی دو مرحله تفکیک نماید:

(۱) جداسازی طیفی: استخراج امضای طیفی خالص عناصر کوچک‌تر از ابعاد پیکسل در صحنه که به اصطلاح امضای انتهایی^۲ می‌نامند.

(۲) جداسازی مکانی: تعیین درصد حضور هر یک از اعضای انتهایی در هر پیکسل که به اصطلاح ضرایب فراوانی^۳ نامیده می‌شود.

۱- مقدمه

با توسعه فناوری سنجیده‌های سنجش از دور، کاربردهای نظامی و غیرنظامی تصاویر ابرطیفی بیش از پیش مورد توجه قرار گرفته است. تصاویر ابرطیفی، مجموعه تصاویر از منطقه مورد مطالعه در صدها باند طیفی مجاور و نزدیک به هم است. این تصاویر حاوی اطلاعات طیفی و مکانی از عناصر و پدیده‌های مختلف موجود در صحنه است. تصاویر ابرطیفی برخلاف برخورداری از قدرت تفکیک طیفی بسیار بالا، قدرت تفکیک مکانی پایین دارند. به همین دلیل عناصر کوچک‌تر از ابعاد یک پیکسل، باعث تشکیل پیکسل‌هایی با امضای طیفی^۱ مخلوط می‌شوند. بنابراین

داده‌های ابرطیفی به تابع هزینه افزوده شده است. از جمله آنها می‌توان به CNMF (Pauca, et al., 2006) و MiniDisCoNMF (Huck, 2010) MiniDisCoNMF، قید همواری به تابع هزینه افزوده شده است و تنها برای استخراج اعضای انتهایی مورد استفاده قرار می‌گیرد و تخمین ضرایب فراوانی با استفاده از روش‌های دیگر همچون FCLS انجام می‌شود (Pauca, et al., 2006). در الگوریتم FCLS انجام می‌شود (Huck, 2010) که فرایندی MiniDisCoNMF، قید کاهش پراکنده‌گی طیفی به تابع هزینه افزوده شده است و پارامترهای تنظیم آن از طریق هزینه افزوده شده است. این الگوریتم، جداسازی طیفی و مکانی بسیار طولانی است. این الگوریتم، جداسازی طیفی و مکانی را به صورت همزمان انجام می‌دهد؛ ولی با این وجود، تخمین ضرایب فراوانی به خوبی استخراج اعضای انتهایی انجام نمی‌شود و این مسأله به سبب قیدی است که به تابع هزینه افزوده شده است، زیرا قید افزوده شده تنها برای استخراج اعضای انتهایی مناسب است.

یکی از ویژگی‌های داده‌های ابرطیفی، همبستگی بین باندهای تصاویر است. واضح است که این ویژگی امکان شناسایی امضای طیفی طبقه‌های مختلف را دشوار می‌کند و این مسأله در طراحی الگوریتم‌ها جداسازی مورد توجه قرار نگرفته است. در این مقاله ما یک شیوه جدید برای جداسازی داده‌های ابرطیفی سنجش از دور با استفاده از تجزیه ماتریس نامنفی (Semi-NMF) (Cichocki, et al., 2009) و تحلیل مؤلفه‌های اصلی^۲ (ICA)، پیشنهاد کردیم. همچنین برای غلبه بر مسأله جواب غیر یکتا، دو قید مطابق با ویژگی داده‌های ابرطیفی به تابع هزینه افزوده ایم. روش پیشنهادی علاوه بر افزایش صحت، می‌تواند سرعت اجرای فرایند جداسازی را نیز افزایش دهد.

بخش‌های بعدی مقاله به صورت زیر بیان شده است. از آنجایی که روش پیشنهادی براساس مدل مخلوط خطی است و از این مدل در تبدیل PCA نیز استفاده می‌شود، ابتدا مدل مخلوط خطی و تبدیل PCA به عنوان پیش‌زمینه در بخش دو توصیف می‌شود و سپس در بخش سه روش پیشنهادی با استفاده از PCA و Semi-NMF و قیدهای کمکی افزوده شده به تابع هزینه تشریح می‌شود. در بخش چهار روش پیشنهادی با استفاده از داده‌های ابرطیفی مصنوعی و واقعی ارزیابی می‌شود. بخش‌های پنجم و ششم به ترتیب شامل نتیجه‌گیری و مراجع هستند.

² Projected Gradient

³ Principal Component Analysis (PCA)

در سال‌های اخیر روش‌های زیادی برای جداسازی داده‌های ابرطیفی در قالب به کارگیری ویژگی‌های هندسی و آماری ارائه شده است. روش‌هایی از جمله N-FINDER (Nascimento and Dias, 2005) و (Boardman, 1993) PPI (Winter, 1999) هندسی جهت جداسازی طیفی و استخراج اعضای انتهایی بهره می‌گیرند. اما این روش‌ها فقط برای تصاویری قابل به کارگیری هستند که دارای حداقل یک پیکسل خالص از هر عضو انتهایی باشند. همچنین این روش‌ها تنها اعضای انتهایی را استخراج می‌کنند و جداسازی مکانی و تخمین ضرایب فراوانی توسط الگوریتم‌های دیگری همچون FCLS (Heinz and Chang, 2001) و SAM (Yuhas, et al., 1992) (Kwon and Nasrabadi, 2005) انجام می‌شود.

الگوریتم‌های آماری از ویژگی‌های آماری داده‌ها در استخراج اعضای انتهایی و تخمین ضرایب فراوانی استفاده می‌کنند. این الگوریتم‌ها با به کارگیری ویژگی‌های آماری، بر بسیاری از محدودیت‌ها و نقص‌های روش‌هایی که بر پایه ویژگی‌های هندسی قرار دارند، غلبه می‌کند. تحلیل مؤلفه‌های مستقل (ICA)، معروف‌ترین روش در این زمینه است که موفقیت‌هایی در جداسازی منابع به صورت کور و تنها با استفاده از فرض استقلال آماری و غیر گوسی بودن کسب کرده است. با این وجود، در خصوص جداسازی تصاویر ابرطیفی، فرض استقلال آماری در حال مصالحه با محدودیت‌های فیزیکی مدل مخلوط خطی است و این مسأله، صحت نتایج داده‌های جداسازی شده با استفاده از این روش را زیر سؤال برده است (Nascimento and Dias, 2005).

تجزیه ماتریس نامنفی^۱ یکی دیگر از روش‌های جدید در حوزه جداسازی کور منابع است که ماتریس داده‌ها را به صورت حاصل‌ضرب دو ماتریس نامنفی تفکیک می‌کند (Berry, et al., 2007). مهم‌ترین مسأله در به کارگیری الگوریتم NMF در جداسازی داده‌های سنجش از دور عدم وجود پاسخ یکتا به سبب وجود کمینه‌های محلی ناشی از غیر محدب بودن تابع هزینه است. از جمله راه کارهای کاهش اثر این نقص افزودن قیدهای کمکی مطابق با ویژگی داده‌ها به تابع هزینه است (Liu, et al., 2004).

در خصوص جداسازی داده‌های سنجش از دور، الگوریتم‌هایی متعددی با استفاده از NMF طراحی شده است. در هر یک از آنها قیدهایی براساس ویژگی‌های

¹ Nonnegative Matrix Factorization(NMF)

که نقشهٔ فراوانی^۳ نامیده می‌شود. هر یک از این نقشه‌ها، فراوانی یک عضو انتهایی در همه پیکسل‌های تصویر را نشان می‌دهد. $E = (e_1, e_2, \dots, e_N) \in R^{L \times N}$ ماتریس خطا و نوافه در همه پیکسل‌های تصویر را مدل می‌کند.

مدل مخلوط خطی تنها به معادله^(۲) محدود نمی‌شود و دارای محدودیت‌های فیزیکی است. این محدودیت‌ها باید در طراحی الگوریتم جداسازی در نظر گرفته شوند و به گونه‌ای اعمال شوند که نتایج به دست آمده بنیان فیزیکی مسأله را نقض نکند، در غیر این صورت نتایج به دست آمده از صحت کافی برخوردار نخواهد بود. این محدودیت‌های فیزیکی عبارتند از:

(۱) همه عناصر ماتریس اعضای انتهایی M باید نامنفی باشند. این قید بر پایه این حقیقت استوار است که درصد انحرافی بازتابی از همه پدیده‌های روی زمین مقادیر نامنفی هستند و مقادیر منفی نامفهوم و بی‌معنی خواهد بود.

(۲) همه عناصر ماتریس ضرایب فراوانی A باید نامنفی باشند، یعنی:

$$A_{ij} \geq 0 \quad (1 \leq i \leq P, 1 \leq j \leq N) \quad (۳)$$

این محدودیت را به‌اصطلاح قید نامنفی‌بودن فراوانی^۴ می‌گویند (Iordache, et al., 2011) و براساس این واقعیت است که مقادیر فراوانی، پارامترهای فیزیکی نامنفی هستند. (۳) مجموع هر یک از بردارهای فراوانی باید برابر واحد باشد، یعنی:

$$\sum_{i=1}^P A_{ij} = 1 \quad (1 \leq j \leq N) \quad (۴)$$

این محدودیت را به‌اصطلاح، قید مجموع واحد^۵ می‌نامند (Iordache, et al., 2011) و بر این مبنای است که باید مجموع سطح‌های پوشیده‌شده توسط هر یک از اعضای انتهایی، برابر با سطح پیکسل باشد.

اگر درون یک ماده چندین بازتاب وجود نداشته باشد یعنی بازتاب از یک ماده دوباره توسط همان ماده بازتاب نشود و از طرف دیگر بین مواد درون پیکسل بازتاب‌های متقابل وجود نداشته باشد و مقایس مخلوط‌شدنی در مقابل اندازه مواد بسیار بزرگ باشد (Lucas, et al., 2004)، می‌توان مخلوط طیفی را به‌صورت خطی در نظر گرفت و از LMM برای جداسازی استفاده کرد.

۲- پیش‌زمینه

از آنجایی که از مفاهیم مدل مخلوط خطی^۶ و تحلیل مؤلفه‌های اصلی در فرایند جداسازی روش پیشنهادی استفاده شده است LMM و PCA به ترتیب در بخش ۱-۲ و ۲-۲ توصیف می‌گردند.

۲-۱- مدل مخلوط خطی

بسیاری از روش‌های جداسازی داده‌های ابرطبیفی براساس مدل LMM عمل می‌کنند. این مدل اولین بار در سال ۱۹۷۱ معرفی شد (Horwitz, et al., 1971) و براساس آن، امضای طیفی هر پیکسل به‌وسیله ترکیب خطی وزن داری از امضای طیفی اعضای انتهایی در صحنه مدل می‌شود. وزن هر یک از اعضای انتهایی، با سطحی از پیکسل که توسط $x_i \in R^{L \times 1}$ عضو انتهایی پوشانده شده متناظر است. اگر بردار امضای طیفی یک پیکسل با L باند طیفی باشد، براساس LMM به‌صورت زیر مدل خواهد شد:

$$x_i = MA_i + e_i \quad (۱)$$

که $M = [M_1, M_2, \dots, M_P] \in R^{L \times P}$ ماتریس اعضای انتهایی است و $M_j \in R^{L \times 1}, 1 \leq j \leq P$ بردار امضای طیفی زأین عضو انتهایی می‌باشد. تعداد اعضای انتهایی موجود در تصویر است. $A_i = (A_{1i}, A_{2i}, \dots, A_{Pi})^T \in R^{P \times 1}$ بردار وزن یا بردار فراوانی اعضای انتهایی در پیکسل آم است و همچنین، $e_i \in R^{L \times 1}$ بردار نوافه و خطا در پیکسل آم می‌باشد.

می‌توان رابطه^(۱) را برای همه N پیکسل تصویر به‌صورت زیر توسعه داد:

$$X = MA + E \quad (۲)$$

که در رابطه فوق $X \in R^{L \times N}$ ماتریس شامل بردار امضای طیفی همه پیکسل‌های تصویر است که به‌صورت ستونی مرتب شده‌اند. $A = (A_1, A_2, \dots, A_N) \in R^{P \times N}$ ماتریس ضرایب فراوانی است که شامل بردارهای فراوانی همه پیکسل‌های تصویر است. اگر هر یک از سطرهای ماتریس فراوانی A به یک ماتریس متناظر با محل اصلی پیکسل‌های تصویر، تبدیل شوند، تصویری به دست می‌آید

¹ Linear Mixing Model (LMM)

² Abundance Map

³ Abundance Nonnegative Constraint (ANC)

⁴ Abundance Sum-to-one Constraint (ASC)

۲-۲- تبدیل مؤلفه‌های اصلی

از تبدیل مؤلفه‌های اصلی آنها در فرایند جداسازی استفاده می‌شود. در روش پیشنهادی، این تبدیل به منظور کاهش ویژگی‌های داده‌ها مورد استفاده قرار نمی‌گیرد؛ بلکه با استفاده از این تبدیل خطی، تصاویر به فضای ناهمبسته نگاشته می‌شوند. در این فضای داده‌ها ناهمبسته و معتمد خواهد شد و همچنین امضا طیفی هر یک از طبقه‌ها به فضای نگاشته می‌شود که تمایز آنها نمایان تر خواهد شد و موجب کارآمدتر شدن فرایند جداسازی می‌شود. بنابراین، پیش از اینکه داده‌ها به طور مستقیم در فرایند جداسازی قرار گیرند باید توسط تبدیل PCA به فضای جدید نگاشته شوند و سپس داده‌های جدید در فضای جدید توسط الگوریتم جداسازی تفکیک شوند. همه مؤلفه‌های اصلی در فرایند جداسازی مورد استفاده قرار می‌گیرند. از آنجایی که جداسازی داده‌ها براساس مدل مخلوط خطی انجام می‌گیرد باید مدل LMM را با استفاده از تبدیل PCA، به صورت زیر

تغییر داد:

$$Y = BX = BMA = SA \quad (6)$$

$$S = BM \quad (7)$$

که در معادله (7)، Y داده‌های جدید حاصل از تبدیل PCA است و S تبدیل PCA ماتریس اعضای انتهایی M است که درنتیجه اعمال تبدیل PCA داده‌ها جدید است. بنابراین با تبدیل PCA داده‌های ابرطیفی، تنها ماتریس اعضای انتهایی M تغییر می‌کند، ولی همچنان ماتریس ضرایب فراوانی A بدون تغییر باقی می‌ماند. پس از تفکیک داده‌های جدید توسط الگوریتم جداسازی ماتریس S به دست خواهد آمد و با عکس تبدیل PCA، ماتریس اعضای انتهایی دوباره به صورت زیر به دست خواهد آمد:

$$M = B^{-1} S. \quad (8)$$

در رابطه (6) دو ماتریس S و A مجهول هستند و باید با استفاده از ماتریس Y تعیین شوند. این مسئله به طور کامل با موضوع جداسازی کور منابع منطبق است. در این مقاله ما از یک روش جدید در حوزه جداسازی کور منابع به نام تجزیه ماتریس نیمه نامنفی (Semi-NMF) برای تفکیک ماتریس Y به دو ماتریس S و A استفاده کردی‌ایم که در بخش ۲-۳ چگونگی این تفکیک تشریح شده است.

۳- روش پیشنهادی

۱-۳- توسعه تبدیل PCA با مدل LMM

همبستگی بین باندهای تصاویر ابرطیفی یکی از ویژگی‌های این مجموعه داده‌هاست. وجود چنین ویژگی بین باندهای تصاویر ابرطیفی امکان شناسایی امضا طیفی طبقه‌های مختلف را دشوار می‌کند. این مسئله را می‌توان در طبقه‌بندی طبقه‌ها و الگوریتم‌های جداسازی ابرطیفی که از امضا طیفی هر پیکسل در فرایند طبقه‌بندی و تفکیک داده‌ها استفاده می‌کنند بهوضوح مشاهده کرد. این مسئله در جداسازی داده‌های ابرطیفی زمانی بیشتر آشکار می‌شود که طبقه‌ها دارای امضاهای طیفی مشابه باشند؛ اما با روش پیشنهادی می‌توان تا حد زیادی این مسئله را بهبود داد. در روش پیشنهادی به جای استفاده مستقیم از داده‌های طیفی،

دوسنیه



دارای تعداد زیادی کمینه محلی است. کمینه‌های محلی باعث می‌شود که نتایج جداولی یکتا نباشد یعنی با تکرار دوباره الگوریتم جداولی، ماتریس A و S متفاوتی حاصل می‌شود. مهم‌ترین راه حل این مسئله افزودن قیدهای کمکی مطابق با ویژگی داده‌های ابرطیفی است (Liu, et al., 2004). قیدهایی افزوده شده باید به گونه‌ای انتخاب شوند که بتواند تا حد زیادی مسئله تعداد زیاد کمینه‌های محلی را بهبود دهند.

در این مقاله ما دو قید مطابق با ویژگی داده‌های ابرطیفی به تابع هزینه افزوده‌ایم. قیدهایی که به تابع هزینه رابطه (۱۰) افزوده می‌شود به ماتریس ضرایب فراوانی اعمال می‌گردد که باید با ویژگی‌های داده‌ها مطابقت داشته باشد. با توجه به اینکه ماتریس Y دارای مقادیر معین است و تنها ماتریس A و S مجهول می‌باشند، هرچه مقادیر ماتریس A با دقت بیشتری استخراج شود نتایج حاصل در ماتریس S نیز از دقت بیشتری برخوردار خواهد بود.

یکی از ویژگی‌های داده‌های ابرطیفی این است که در تشکیل امضای طیفی یک پیکسل مخلوط، به طور حتم همه اعضای انتهایی شرکت ندارند و تنها عناصر و پدیده‌هایی که درون پیکسل قرار دارند در تشکیل امضای طیفی شرکت دارند. این ویژگی در ماتریس ضرایب فراوانی باید بدین صورت نمایان شود که برخی از عناصر ضرایب فراوانی ماتریس امضای انتهایی که در خارج از پیکسل مخلوط قرار دارند. به عبارت دیگر ماتریس امضای انتهایی تنک خواهد بود. برای اعمال این ویژگی ما از ماتریس nsNMF استفاده کردیم (Cichocki, et al., 2009). در این روش تنکی ماتریس فراوانی توسط ماتریس زیر کنترل می‌شود.

$$C = (1-\theta)I + \frac{\theta}{P} \mathbf{1} \quad (11)$$

در رابطه بالا C یک ماتریس ثابت است و روشی ریاضی است که برای کنترل تنکی ماتریس A استفاده می‌شود. I ماتریس همانی با ابعاد $P \times P$ است و $\mathbf{1}$ ماتریسی با ابعاد $P \times 1$ و دارای مؤلفه‌هایی برابر عدد یک است. همچنین برای کنترل میزان تنکی ماتریس فراوانی از θ ($0 \leq \theta \leq 1$) استفاده می‌شود، به گونه‌ای که با افزایش θ از صفر تا یک، تنکی ماتریس فراوانی افزایش می‌یابد. ماتریس C به صورت یک عامل ضرب شونده به ماتریس فراوانی A افزوده می‌شود.

۲-۳- جداولی تبدیل PCA داده‌ها با استفاده

از Semi-NMF محدودشده

ماتریس X شامل مقادیر انعکاس همه پیکسل‌های مجموعه تصاویر ابرطیفی است. انعکاس یک پارامتر فیزیکی همواره نامنفی است. تبدیل PCA ماتریس داده‌های ابرطیفی X موجب می‌شود ماتریس Y شامل مقادیر نامنفی و منفی شود، یعنی:

$$Y_{\pm} = S_{\pm} A_{\pm} \quad (9)$$

که زیرنویس \pm بیان می‌کند ماتریس شامل مقادیر نامنفی و منفی است و همچنین زیرنویس $+$ بیان می‌کند، ماتریس فقط شامل مقادیر نامنفی است. همچنین در رابطه بالا ماتریس‌های S و A مجهول هستند و تنها ماتریس Y مشخص و معین است. برای تفکیک ماتریس Y به حاصل ضرب دو ماتریس S و A، الگوریتم NMF روشی ریاضی نوین می‌باشد. روش‌های جداولی ابرطیفی MiniDisCoNMF و CNMF NMF برای تفکیک داده‌های ابرطیفی استفاده می‌کند. شرط به کارگیری NMF در تفکیک داده‌ها، نامنفی بودن هر سه ماتریس Y, S و A است. اما در روش پیشنهادی تنها ماتریس A نامنفی است و ماتریس Y و S شامل مقادیر منفی و نامنفی است. این ویژگی ما را به سوی الگوریتم جداولی هدایت می‌کند که بتواند ماتریس شامل مقادیر نامنفی و منفی را به حاصل ضرب دو ماتریس تفکیک کند که یکی از ماتریس‌ها شامل مقادیر نامنفی و منفی باشد و ماتریس دیگر تنها شامل مقادیر نامنفی باشد. ما برای این منظور از الگوریتم Semi-NMF (Cichocki, et al., 2009) استفاده کردیم.

این الگوریتم نوع توسعه یافته الگوریتم NMF است و از طریق یک مدل خطی و با استفاده از تابع هزینه اقلیدسی زیر، ماتریس داده‌های Y شامل مقادیر نامنفی و منفی را به دو ماتریس A و S تفکیک می‌کند:

$$\begin{aligned} CF(S, A) &= \frac{1}{2} \|Y - SA\|_F^2 \\ &= \frac{1}{2} \text{trace}((Y - SA)^T (Y - SA)) \end{aligned} \quad (10)$$

که در معادله بالا $\| \cdot \|_F$ نرم Frobenius است. اما این الگوریتم همانند NMF به سبب غیرمحاسبه بودن تابع هزینه،

$$S \leftarrow (YA^TC^T)(CAA^TC^T)^{-1} \quad (16)$$

$$A \leftarrow \left[(C^TS^TSC^T + \alpha I)^{-1} (C^TS^TY) \right]_+ \quad (17)$$

که در رابطه بالا، و عملگر $[.,]$ ، همه مؤلفه‌های غیر منفی ماتریس را صفر می‌کند و موجب نامنفی شدن عناصر ماتریس A می‌شود تا قید فیزیکی ANC LMM در نظر گرفته شود. پس از عکس تبدیل PCA ماتریس S، ماتریس اعضای انتهایی M به دست خواهد آمد.

برای اعمال قید فیزیکی ASC از روش ارائه شده در (Heinz and Chang, 2001) استفاده شده است، یعنی:

$$Y \leftarrow \begin{bmatrix} Y \\ \delta_{1T} \end{bmatrix}, S \leftarrow \begin{bmatrix} S \\ \delta_{1T} \end{bmatrix} \quad (18)$$

که در رابطه بالا^۱ بردار سطری شامل تنها عدد یک است و پارامتر δ مجموع ضرایب فراوانی را کنترل می‌کند.

۳-۳- مقداردهی اولیه و معیار توقف الگوریتم

ماتریس‌های S و A در الگوریتم Semi-NMF به صورت تصادفی مقداردهی اولیه می‌شوند و این موضوع می‌تواند روی نتایج به دست آمده مؤثر باشد، چون انتخاب مقدار اولیه در الگوریتم‌های یادگیری از اهمیت ویژه‌ای برخوردار است. انتخاب مقدار اولیه مناسب می‌تواند در اجرای سریع الگوریتم کمک فراوانی می‌کند، همچنین می‌تواند تا حد زیادی از به دست آمدن نتایج مربوط به کمینه‌های محلی جلوگیری کند. به همین دلیل برخی از پژوهش‌گران، پژوهش‌های خود را بر روی انتخاب مؤثر مقادیر اولیه متمرکز کرده‌اند که از بین آنها می‌توان به مقداردهی اولیه براساس SVD (Boutsidis and Gallopoulos, 2008)، (Wild, et al., 2004) spherical K-clustering اما هیچ یک از روش‌ها برای داده‌های ابرطیفی مناسب نیستند.

در این مقاله ما از یک روش ساده برای مقداردهی اولیه ماتریس S استفاده کردیم. در این روش تعداد P داده به صورت تصادفی و یکی یکی از ماتریس X، به‌گونه‌ای انتخاب می‌شود که ضریب همبستگی بین همه آنها از مقدار آستانه‌ای که از قبل تعیین شده است، کمتر باشد و سپس داده‌های متاظر آنها در ماتریس Y به عنوان مقادیر اولیه ماتریس S در نظر گرفته می‌شود. بنابراین با استفاده از این

ضرایب فراوانی از یک پیکسل به پیکسل بعدی به صورت آنی تغییر نمی‌کند و در اکثر مواقع تغییر آنها به‌آرامی و نرمی انجام می‌شود. این ویژگی به این سبب است که حضور عناصر و پدیده‌ها از یک پیکسل به پیکسل دیگر به سرعت تغییر نمی‌کند. این ویژگی را می‌توان به صورت همواری ضرایب فراوانی از یک پیکسل به پیکسل بعدی توصیف کرد که باید به ماتریس ضرایب فراوانی اعمال شود. برای اعمال این ویژگی از روش سنتی نرم Frobenius ارائه شده در الگوریتم CNMF استفاده شده است که به صورت زیر تعریف می‌شود (Pauca, et al., 2006)

$$\|A\|_F^2 = \text{trace}(A^T A) = \sum_{i=1}^P \sum_{j=1}^N |A_{ij}|^2 \quad (12)$$

بنابراین با معرفی قیدهای تنکی و همواری ضرایب فراوانی به تابع هزینه اقلیدسی، با تابع هزینه توسعه یافته جدید زیر روبه‌رو خواهیم شد:

$$CF(S, A) = \frac{1}{2} \|Y - SCA\|_F^2 + \alpha J(A) \quad (13)$$

$$J(A) = \|A\|_F^2 \quad (14)$$

که در رابطه (13) C ماتریس nsNMF است که به درون تابع هزینه اقلیدسی به منظور اعمال تنکی روی بردارهای فراوانی ماتریس فراوانی افزوده شده است و $J(A)$ قید همواری ماتریس ضرایب فراوانی است که به وسیله پارامتر تنظیم α به تابع هزینه افزوده شده است. برای محاسبه قاعده به روزرسانی ما از الگوریتم یادگیری ALS^۱ استفاده کردیم. شیوه Gradient Descent نیازی به طول گام و تنظیم آن ندارد و سریع‌تر عمل می‌کند (Comon and Jutten, 2010)، (Berry, 2007) و از ویژگی مهم دیگر آن کمک به تنکی ماتریس فراوانی است. الگوریتم یادگیری ALS با مساوی صفر قراردادن گرادیان تابع هزینه نسبت به هر یک از ماتریس‌های S و A با فرض ثابت‌سودن دیگری، قواعد به روزرسانی ماتریس‌ها را محاسبه می‌کند، یعنی:

$$\frac{\partial CF(S, A)}{\partial S} = 0, \quad \frac{\partial CF(S, A)}{\partial A} = 0 \quad (15)$$

با حل معادلات بالا، قواعد به روزرسانی ماتریس‌های S و A به صورت زیر به دست می‌آید:

^۱ Alternating Least Square

همچنین SID که یک معیار براساس نظریه اطلاعات است برای کامین عضو انتهایی استخراج شده به صورت زیر تعریف می‌شود (Chang, 2003):

$$SID_k = D(\hat{M}_k | M_k) + D(M_k | \hat{M}_k) \quad (21)$$

که در رابطه بالا $D(M_k | \hat{M}_k)$ آنتروپی نسبی \hat{M}_k با توجه به M_k می‌باشد و به صورت زیر تعریف می‌شود:

$$D(M_k | \hat{M}_k) = \sum_{i=1}^L p_i \log\left(\frac{p_i}{q_i}\right) \quad (22)$$

$$p_i = \frac{M_{ik}}{\sum_{j=1}^L M_{jk}} \quad (23)$$

$$q_i = \frac{\hat{M}_{ik}}{\sum_{j=1}^L \hat{M}_{jk}}$$

مقادیر این معیار هر چه به صفر نزدیک‌تر باشد شباهت نتایج استخراج شده بیشتر خواهد بود. معیار RMSE برای k امین نقشه فراوانی به صورت زیر تعریف می‌شود (Plaza, 2004):

$$RMSE_k = \sqrt{\frac{1}{N} \sum_{j=1}^N (A_{kj} - \hat{A}_{kj})^2} \quad (24)$$

در رابطه بالا \hat{A}_{kj} و A_{kj} به ترتیب ضریب نسبت فراوانی k امین عضو انتهایی در پیکسل j است و N تعداد کل پیکسل‌های تصویر است.

۴- نتایج داده‌های مصنوعی ابرطیفی

برای تولید تصویر ابرطیفی مصنوعی، امضای طیفی چهار ماده از کتابخانه طیفی USGS انتخاب شده و تصویری با پیکسل‌های مخلوط با ابعاد 100×100 تولید می‌شود. در انتخاب مواد به شباهت امضای طیفی توجه شده است. به منظور بررسی کارآمدی الگوریتم پیشنهادی در فرایند جداولی داده‌ها، سه آزمایش طراحی شده است. هر یک از آزمایش‌های یکی از عوامل مهم را مورد بررسی قرار می‌دهد. در این بخش روش پیشنهادی با دو الگوریتم جداولی NMF و CNMF و MiniDisCoNMF که براساس

طراحی شده‌اند مقایسه می‌شود.

آزمایش اول، تأثیر قدرت نوافه در فرایند جداولی داده‌ها با استفاده از روش پیشنهادی را مورد بررسی قرار می‌دهد. وجود نوافه گوسی موجب تغییر شکل امضای طیفی

روش، امضای طیفی پیکسل‌هایی به عنوان مقدار اولیه انتخاب می‌شود که کمترین ارتباط خطی با یکدیگر دارند، به عبارت دیگر مقدار اولیه ماتریس k ، پیکسل‌هایی از مواد مختلف و تا حد ممکن خالص خواهند بود. با تعیین مقدار اولیه ماتریس اعضای انتهایی، مقدار اولیه ماتریس A به صورت زیر با استفاده از مقادیر اولیه ماتریس S به دست می‌آید:

$$A = S^{-1}Y. \quad (19)$$

از آنجایی کهتابع هزینه در هر تکرار به صورت نزولی در حال کاهش است، تفاضل تابع هزینه در دو حلقه تکرار متوالی به عنوان معیاری برای توقف الگوریتم پیشنهادی استفاده شده است. اگر تفاضل‌ها در تعداد معینی از تکرارها از مقدار خطأ مورد نظر کمتر باشد می‌توان الگوریتم را متوقف کرد در غیر این صورت الگوریتم بعد از حداقل تعداد تکرار متوقف خواهد شد.

۴- نتایج آزمایشی

در این بخش از مجموعه داده‌های ابرطیفی مصنوعی و واقعی برای بررسی کارایی الگوریتم پیشنهادی استفاده شده است. برای مقایسه کمی نتایج، از سه معیار ¹SAD، ²SID و ³RMSE استفاده شده است. این معیارها، سه معیار متدالول در مطالعات جداولی داده‌های سنجش از دور هستند که برای اندازه‌گیری میزان شباهت نتایج به دست آمده و مقادیر مرجع مورد استفاده قرار می‌گیرند. معیار SAD برای k امین عضو انتهایی استخراج شده به صورت زیر تعریف می‌شود

: (Keshava and Mustard, 2002)

$$SAD_k = \cos^{-1}\left(\frac{\hat{M}_k^T M_k}{|\hat{M}_k^T| \cdot |M_k|}\right) \quad (20)$$

که در رابطه بالا، \hat{M}_k و M_k به ترتیب عضو انتهایی استخراج شده و عضو انتهایی مرجع می‌باشد و نماد $||$ عملگر نرم برداری است. هر چه مقدار SAD کمتر باشد، شباهت عضو انتهایی استخراج شده بیشتر خواهد بود. این معیار بیشتر در مواردی مناسب است که عضو انتهایی استخراج شده دارای ابهام مقیاس باشد، چون ضریب‌ها تأثیری در عبارت کسری درون رابطه (20) ندارد.

¹ Spectral Angel Distance

² Spectral Information Divergence

³ Root Mean Square Error

یک پیکسل خالص امکان‌پذیر است. در این آزمایش توانایی الگوریتم‌ها در جداسازی داده‌های مورد بررسی قرار می‌گیرند که تعداد پیکسل‌های مختلط آن‌ها در مقایسه با پیکسل‌های خالص بسیار زیاد باشند. در این آزمایش SNR ثابت برابر 40dB در نظر گرفته شده است.

شکل (۲-۴) (الف) و (۲-۴) (ب) به ترتیب مربوط به اعضاً انتهایی استخراج شده و ضرایب فراوانی است. همان‌طور که در شکل (۲-۴) نشان داده شده است، الگوریتم‌ها به‌ویژه الگوریتم پیشنهادی در استخراج اعضاً انتهایی کمتر تحت تأثیر افزایش تعداد پیکسل‌های مخلوط قرار گرفته‌اند و مقادیر متوسط معیارهای RMSE و SAD کمتری به خود اختصاص داده است؛ اما با افزایش تعداد پیکسل‌های مختلط متوسط مقادیر معیارها در حال افزایش است و نشان می‌دهد که کارایی الگوریتم‌ها در حال کاهش است. اما الگوریتم پیشنهادی نسبت به دو روش دیگر کمتر متاثر شده است و افت و خیز کمتری دارد؛ به‌ویژه وقتی که هیچ پیکسل خالصی در تصویر وجود ندارد و متوسط RMSE با شبیه‌سازی افزایش می‌یابد؛ اما روش پیشنهادی با شبیه کمتری افزایش یافته است.

آزمایش سوم، تأثیر قیدهای پیشنهادی افزوده شده و تبدیل PCA در جداسازی داده‌ها را مورد بررسی قرار می‌دهد و کارایی الگوریتم به‌صورت جداگانه با حضور و عدم حضور قیدها و تبدیل PCA، متوسط معیارها بررسی شده است.

در این آزمایش فرایند جداسازی همراه با قیدها پیشنهادی و بدون آنها و همچنین با وجود و عدم وجود تبدیل PCA، ده بار تکرار شده و در هر بار اجرا معیار SAD و آن‌ها محاسبه شده است. در جدول (۱-۴)، متوسط معیارها و انحراف استاندارد آنها نشان داده شده است. نتایج این جدول نشان می‌دهد که قیدهای پیشنهادی و تبدیل PCA علاوه‌بر افزایش صحت نتایج به دست آمده، انحراف استاندارد آنها را نیز کاهش داده است و نتایج استخراج شده از قابلیت اطمینان بیشتری برخوردارند. جداسازی داده‌ها بدون قیدها و با حضور و عدم حضور تبدیل PCA نیز در جدول (۱-۴) نشان داده شده است. نتایج نشان می‌دهد که

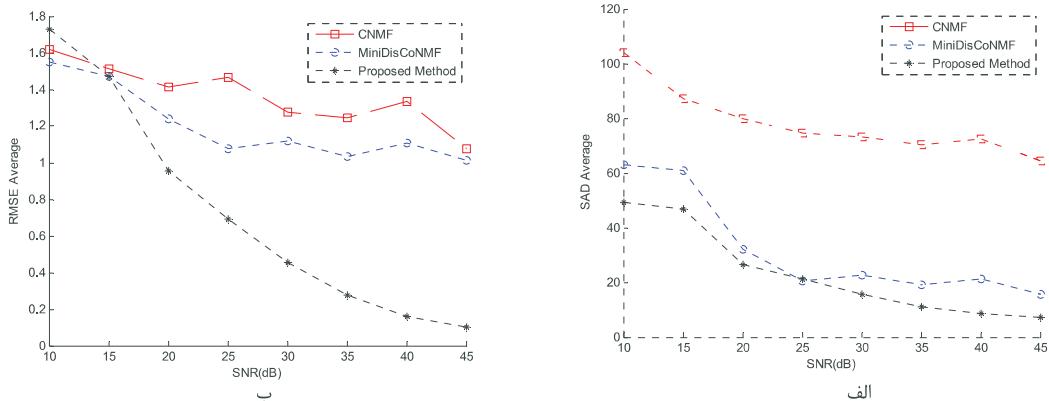
پیکسل‌ها می‌شود و زمانی که امضای طیفی هر یک از پیکسل‌ها به‌صورت تصادفی توسط نویه گوسی تغییر شکل یافته باشند، آزمون مناسبی برای بررسی و مقایسه کارایی الگوریتم‌ها است. برای این منظور داده‌های شبیه‌سازی شده به نویه گوسی با متوسط صفر و با SNR معین آشسته می‌شوند. در این حالت SNR به‌صورت زیر توصیف می‌شود:

$$SNR_{dB} = 10 \log \frac{E\{x_i^T x_i\}}{E\{n T_n\}} \quad (25)$$

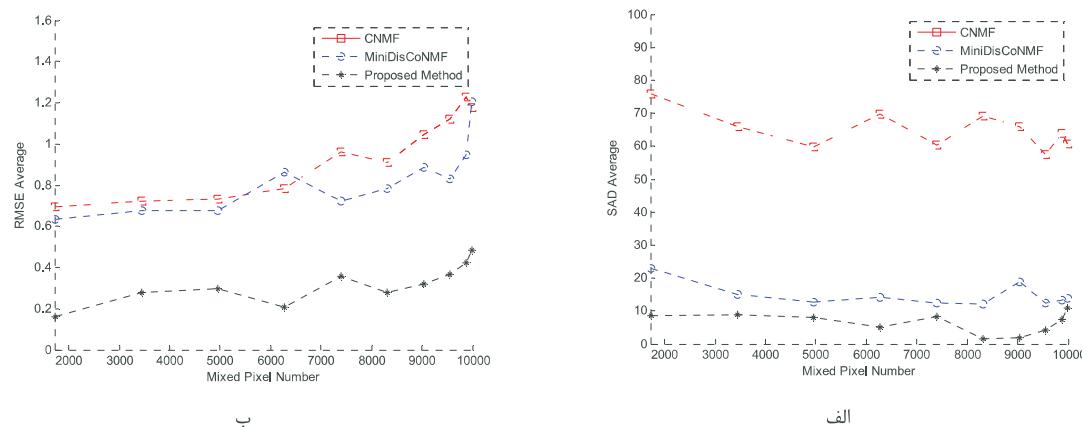
که در رابطه بالا x_i و n به ترتیب مشاهدات و نویه مربوط به یک پیکسل تصویر می‌باشد و همچنین $\{E\}$ نماد عملگر امید ریاضی است.

شکل (۱-۴) (الف) و (۱-۴) (ب) به ترتیب مربوط به مقایسه اعضاً انتهایی استخراج شده و ضرایب فراوانی است. همان‌طور که انتظار داشتیم و در شکل (۱-۴) می‌توان مشاهده کرد، با افزایش SNR میزان کارایی الگوریتم‌ها در حال افزایش است؛ اما کارایی الگوریتم پیشنهادی در استخراج اعضاً انتهایی و تخمین ضرایب فراوانی بهتر از دو الگوریتم دیگر است و منحنی‌های مربوط به روش پیشنهادی مقادیر پایین‌تر از دو روش دیگر به خود اختصاص می‌دهد. الگوریتم پیشنهادی برخلاف دو الگوریتم دیگر متناسب با افزایش مقدار SNR، با کمترین میزان افت و خیز در حال افزایش است و پایداری الگوریتم و قیدهای پیشنهادی در شرایط نویزی متفاوت را نشان می‌دهد. با وجود این‌که الگوریتم MiniDisCoNMF نتایجی نزدیک به الگوریتم پیشنهادی در استخراج اعضاً انتهایی دارد، ولی تخمین ضرایب فراوانی به‌خوبی استخراج اعضاً انتهایی انجام نمی‌شود؛ این مسئله می‌تواند به‌دلیل معرفی تنها یک قید به تابع هزینه باشد.

آزمایش دوم، تأثیر تعداد پیکسل‌های مخلوط تصویر در استخراج اعضاً انتهایی و نقشهٔ فراوانی مورد بررسی قرار می‌دهد. برخی از الگوریتم‌های جداسازی همچون VCA (فقط برای تصاویری قابل Nascimento and Dias, 2005) به کارگیری هستند که حداقل یک پیکسل خالص از هر عضو انتهایی داشته باشند؛ اما روش پیشنهادی مشروط به وجود چنین شرطی نیست و فرایند جداسازی حتی بدون وجود



(شکل ۴-۱): مقایسه کارایی الگوریتم‌ها بر حسب SNR های مختلف. الف) SAD. ب) RMSE



(شکل ۴-۲): مقایسه کارایی الگوریتم‌ها بر حسب تعداد پیکسل‌های مختلف. الف) SAD. ب) RMSE

(جدول ۴-۱): متوسط و انحراف استاندارد معیارهای SAD و RMSE نتایج استخراج شده با توجه به حضور و عدم حضور قیدها و تبدیل PCA

	جداسازی با روشن پیشنهادی	جداسازی با قیدها و بدون PCA	جداسازی بدون قیدها و با PCA	جداسازی بدون قیدها و بدون PCA
SAD Mean	۸/۶۵	۲۰/۱۲	۶۲/۶۱	۵۶/۸۴
SAD S.D	۰/۰۸	۰/۱۲	۱۵/۰۴	۱۱/۳۲
RMSE Mean	۰/۱۶	۱/۰۱	۱/۰۳	۱/۰۶
RMSE S.D	۰/۰۰۱	۰/۰۱	۰/۲۵	۰/۲۲

مرجع مقایسه شده‌اند. اعدادی که ضخیم نشان داده شده‌اند نتایج بهتر در مقایسه با سایر روش‌ها نشان می‌دهند. در شکل (۴-۴) نقشهٔ فراوانی تخمینی توسط هر یک از روش‌ها نشان داده شده است.

امضای طیفی Road و Water در این تصویر شباهت بسیار زیادی به یکدیگر دارند و با وجود این‌که در تصویر پیکسل مخلوطی بین این دو عضو انتهایی وجود ندارد، هیچ یک از روش‌های جداسازی به جز روش پیشنهادی، نتوانسته آنها را از یکدیگر تفکیک کنند. با دقت نظر در جدول‌های (۴-۳) و (۴-۴) ستون (الف) متوجه خواهیم شد که الگوریتم‌های جداسازی VCA و CNMF نتوانسته‌اند، عضو انتهایی Road را به صورت مجزا استخراج کنند و با این‌که امپای طیفی عضو انتهایی Water استخراج شده‌اما در نقشهٔ فراوانی آنها همراه با عضو انتهایی Road نشان داده شده است. الگوریتم MiniDisCoNMF با وجود این‌که عضو انتهایی Road را با شباهت کمتری نسبت به روش پیشنهادی استخراج کرده، اما نقشهٔ فراوانی آن به خوبی این مناطق را نشان نداده است و نقشهٔ فراوانی Water همانند دو الگوریتم دیگر همراه با عضو انتهایی Road نشان داده شده است. در این تصویر علاوه بر شباهت امپای طیفی Road و Water، امپای طیفی Tree و Grass نیز به یکدیگر شباهت دارند و با دقت نظر در جدول‌های (۲-۴) و (۴-۴) و شکل (۴-۴) ستون‌های (ب) و (ج) متوجه خواهیم شد که روش پیشنهادی بهوضوح بین این دو طبقه تمایز قرار داده است و موفق به جداسازی آنها از یکدیگر شده است. روش پیشنهادی علاوه بر استخراج همه اعضای انتهایی و تمایز بین امپای طیفی اعضای انتهایی و تخمین نقشهٔ فراوانی آنها، از متوسط نتایج بهتری نیز برخوردار است و این نتیجه در جدول‌های (۲-۴) و (۴-۳) نشان داده شده است.

مجموعه داده ابرطیفی دوم توسط سنجنده^۲ و از ایالت Indiana در سال ۱۹۹۲ بدست آمده است که نمایش رنگی آن در شکل (۵-۴) نشان داده شده است. این داده ابرطیفی شامل مجموعه تصاویری با ابعاد ۱۴۵×۱۴۵ در ۲۲۰ باند طیفی است که برخی از باندهای نویه‌ای حذف شده‌اند. دقت طیفی و مکانی آن به ترتیب ۱۰nm و ۱۷m است. این داده ابرطیفی شامل ۱۷ طبقه است (Landgrebe, 1998). با ادغام برخی از طبقه‌ها، تعداد طبقه‌ها به ۵ طبقه کاهش می‌یابد. در جدول‌های (۴-۴) و (۵-۴)



(شکل ۴-۴): نمایش رنگی باندهای ۸۰ و ۱۰۲۵ از داده Washington

تبديل PCA بر روی تعداد کمینه‌های محلی تأثیری چندانی ندارد و همچنان انحراف معیار نتایج به دست آمده بسیار زیاد است. در این آزمایش الگوریتم با قیدها و بدون تبدیل PCA نیز مورد بررسی قرار گرفت. نتایج به دست آمده بهبود معیارها را نشان می‌دهد؛ اما نتایج نشان می‌دهد که همراهی قیدها و تبدیل PCA با یکدیگر تأثیر زیادی در خروجی الگوریتم پیشنهادی دارد.

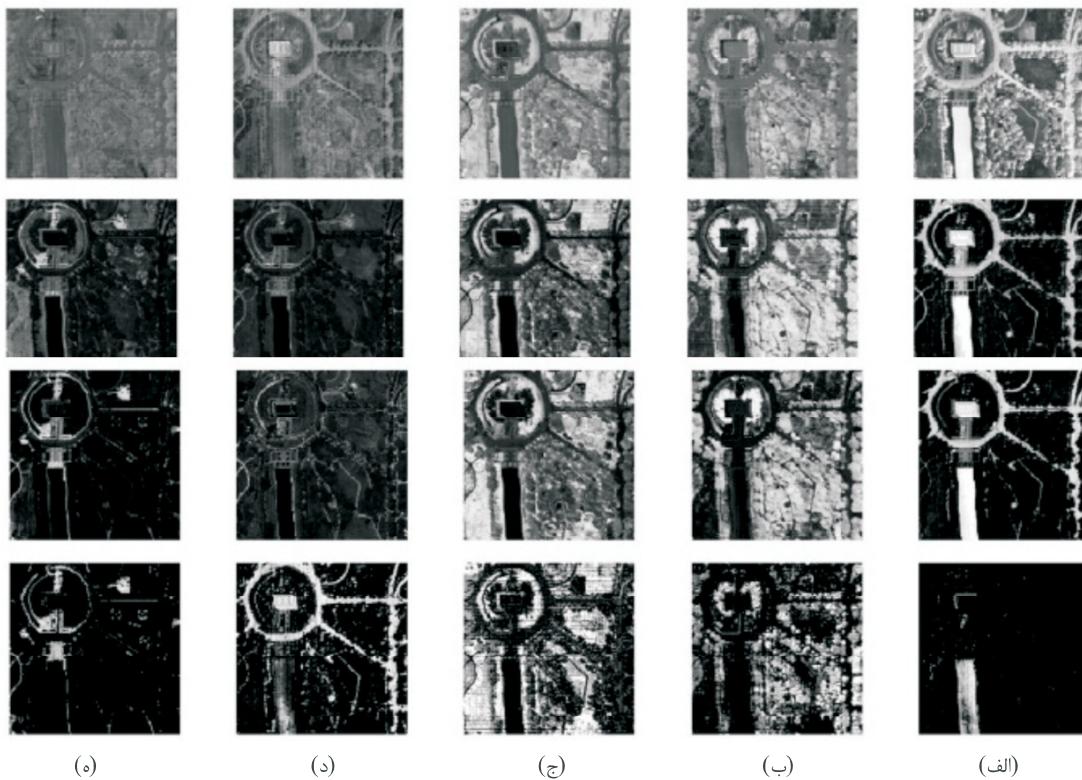
۲-۴- نتایج داده‌های واقعی ابرطیفی

در این بخش الگوریتم پیشنهادی با استفاده از دو مجموعه داده ابرطیفی واقعی ارزیابی شده است. همچنین علاوه بر الگوریتم‌های MiniDisCoNMF و CNMF جداسازی VCA که بر پایه ویژگی‌های هندسی عمل می‌کند برای مقایسه نتایج استخراج شده استفاده شده است. از آنجایی که برای این داده‌ها اعضای انتهایی مرجع وجود ندارد، تعدادی پیکسل نمونه از هر طبقه با استفاده از Ground Truth انتخاب شده و پس از متوسط‌گیری، به عنوان اعضای انتهایی مرجع در نظر گرفته می‌شوند تا با اعضای انتهایی استخراج شده مورد مقایسه قرار گیرد.

مجموعه داده ابرطیفی اول توسط سنجنده^۱ و از HYDICE به دست آمده است. بخشی از این داده ابرطیفی که شامل مجموعه تصاویری با ابعاد ۱۵۰×۱۵۰ است، مورد استفاده قرار گرفته که در شکل (۳-۴) تصویر رنگی از آن نشان داده شده است. این داده دارای ۲۱۰ باند طیفی در محدوده $0.4-2.4 \mu\text{m}$ است. پیش از استفاده از این داده برخی از باندهای نویه‌ای حذف شده‌اند. این داده ابرطیفی شامل ۵ طبقه است، که در جدول‌های (۲-۴) و (۳-۴) اعضای انتهایی استخراج شده و

² Airborne Visible/Infrared Imaging Spectrometer

¹ Hyperspectral Digital Imagery Collection Experiment



(شکل ۴-۴) نقشه فراوانی هر یک از اعضای انتهایی استخراج شده توسط MiniDisCoNMF، CNMF، VCA و روش پیشنهادی (به ترتیب از بالا به پایین). الف. Roof- ۰ .Road- ۵ .Grass- ۶ .Tree- ۷ .Water- ۸ .ب- ۹ .ج- ۱۰ .د- ۱۱ .ه- ۱۲ .الف)

(جدول ۲-۴) مقایسه معیار SAD اعضای انتهایی استخراج شده و مرجع

Washington Data	Roof	Road	Grass	Tree	Water	Average
VCA -FCLS	۱۸/۱۴	—	۸/۸۶	۳/۴۲	۲۲/۸۱	—
CNMF	۱۷/۸۲	—	۹/۵۰	۱۴/۷۲	۲۱/۰۱	—
MiniDisCoNMF	۱۵/۶۶	۱۹/۶۳	۱۰/۹۲	۴/۵۹	۲۵/۳۵	۷۷/۱۷
Proposed Method	۴/۰۷	۱۵/۳۱	۱/۰۹	۳/۴۵	۱۶/۱۵	۴۰/۰۸

(جدول ۳-۴) مقایسه معیار SID اعضای انتهایی استخراج شده و مرجع

Washington Data	Roof	Road	Grass	Tree	Water	Average
VCA -FCLS	۰/۱۳	—	۰/۰۶	۰/۰۲	۰/۱۴	—
CNMF	۰/۱۰	—	۰/۰۷	۰/۱۰	۰/۱۳	—
MiniDisCoNMF	۰/۰۷	۰/۱۰	۰/۰۷	۰/۰۳	۰/۲۱	۰/۵۰
Proposed Method	۰/۰۰۶	۰/۰۸	۰/۰۰۰۵	۰/۰۱	۰/۰۸	۰/۱۸

روش پیشنهادی توجه ویژه‌ای به همبستگی بین باندهای تصاویر ابرطیفی شده است و برای کاهش اثر آن بهویژه در طبقه‌های مشابه، تبدیل مؤلفه‌های اصلی مورد استفاده قرار گرفته است، همچنین با استفاده از این تبدیل خطی امضا طیفی هر یک از طبقه‌ها به فضای نگاشته می‌شود که تمایز آنها نمایان‌تر خواهد شد و جداسازی اعضای انتهایی مشابه بهتر انجام خواهد شد. برای بهبود کارایی-Semi-NMF و کاهش جواب‌های غیر یکتا، ما قید nsNMF را به تابع هزینه برای تنکی ضرایب فراوانی ارائه کردیم. با استفاده از این قید و به همراه قید همواری علاوه‌بر بهبود نتایج استخراج شده واریانس آن‌ها نیز کاهش می‌یابد. همچنین برای الگوریتم پیشنهادی روشی ساده برای مقداردهی اولیه پیشنهاد شد. عدم نیاز به حداقل یک پیکسل خالص از هر عضو انتهایی و استخراج اعضای انتهایی و ضرایب فراوانی به صورت همزمان از ویژگی‌های الگوریتم پیشنهادی در مقایسه با الگوریتم VCA است. توسعه مدل مخلوط خطی (LMM) با استفاده از تبدیل ساده PAC و همچنین استفاده از Semi-NMF به همراه قید nsNMF برای تنکی از جمله ویژگی‌های الگوریتم پیشنهادی در مقایسه با الگوریتم‌های MinDisCoNMF و CNM است. روش پیشنهادی با استفاده از داده‌های ابرطیفی مصنوعی و واقعی مورد بررسی و مقایسه با الگوریتم‌های جداسازی بالا قرار گرفت و نتایج استخراج شده برتری روش پیشنهادی را نشان می‌دهد.

۶- مراجع

Berry MichaelW, Browne Murray, Langville Amy N., V. Paul Pauca, Robert J. Plemmons , Algorithms and applications for approximate nonnegative matrix factorization , Computational Statistics & Data Analysis 52 ,2007, pp. 155 – 173.

Boardman Joseph W., Automating spectral unmixing of AVIRIS data using convex geometry concepts, in Summaries 4th Annu. JPL Airborne Geoscience Workshop, 1993, vol. 1, JPL Pub. 93-26, pp. 11–14.

Boutsidis C. and Gallopoulos E., SVD based initialization: A head start for nonnegative matrix factorization, Pattern Recognit. , Apr. 2008 vol. 41, no. 4, pp. 1350–1362.

Cichocki Andrzej, Zdunek Rafal, Anh Huy Phan., Amari Shun-ichi, Nonnegative Matrix and Tensor Factorizations,: Application to Exploration Multi-Way Data Analysis and Blind Source Separation, First edition , Singapore ,United Kingdom : John Wiley & Sons, Ltd,2009.



(شک-۴) نمایش رنگی باندهای ۱۵۰,۹۰ و ۱۰۰ از داده Indiana

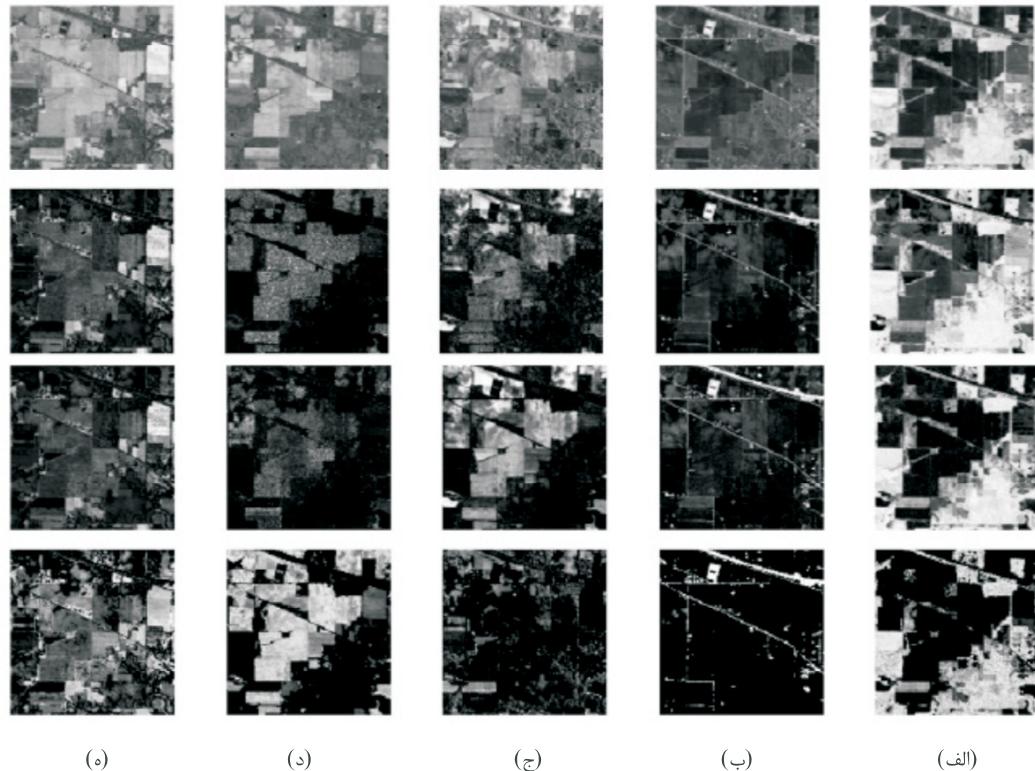
نتایج مقایسه اعضای انتهایی استخراج شده و مرجع نشان داده شده است. اعدادی که ضخیم نشان داده شده‌اند نتایج بهتر در مقایسه با سایر روش‌ها را نشان می‌دهند.

در شکل (۶-۴) نیز نقشه فراوانی تخمینی توسط هر یک از روش‌ها نشان داده شده است. امضا طیفی اعضا انتهایی E3 و E4 شباهت بسیار زیادی به یکدیگر دارند. همچنین در تصاویر پیکسل‌های مختلط زیادی در مرز بین آنها وجود دارند که متعلق به هر دو عضو انتهایی باشند؛ حتی در بعضی از پیکسل‌ها، اعضا انتهایی دیگری نیز حضور دارند. با این وجود الگوریتم‌های CNMF و MinDisCoNMF در تفکیک این دو طبقه موقوفیت چندانی نداشته‌اند؛ ولی الگوریتم پیشنهادی توفیق بیشتری داشته است. این موضوع را بسراحتی می‌توان از شکل (۶-۴) ستون‌های (ج) و (د) که به ترتیب مربوط به نقشه ضرایب فراوانی هستند مشاهده کرد. در نقشه ضرایب فراوانی مربوط به الگوریتم‌های CNMF و MinDisCoNMF وجود عضو انتهایی E4 در نقشه ضرایب فراوانی E3 کامل مشهود است؛ اما در نقشه ضرایب فراوانی الگوریتم پیشنهادی این دو عضو انتهایی تا حد بسیار زیادی متمایز شده‌اند. این نتیجه را تا حدودی می‌توان در نقشه ضرایب فراوانی الگوریتم VCA نیز مشاهده کرد. نتایج جدول‌های (۴-۴) و (۵-۴) این موضوع را تأیید می‌کنند. متوسط نتایج به دست آمده نشان می‌دهد، الگوریتم پیشنهادی با اختلاف زیادی بهتر از روش‌های دیگر داده‌های ابرطیفی را تفکیک کرده است.

۵- نتیجه‌گیری

در این مقاله یک روش جدید برای جداسازی طیفی و مکانی تصاویر ابرطیفی با استفاده از تجزیه ماتریس نیمه‌نامنفی و براساس تبدیل مؤلفه‌های اصلی داده‌ها ارائه شده است. در

فصل نهم
پنجم



(شکل ۴-۶) نقشه فراوانی هر یک از اعضای انتها بی استخراج شده توسط MiniDisCoNMF، CNMF، VCA و روش پیشنهادی (به ترتیب از بالا به پایین). (الف) E1 - ج E3 - د E4 - ه E5 - (ب)

(جدول ۴-۴) مقایسه معیار SAD اعضای انتها بی استخراج شده و مرجع

Indiana Data	E5	E4	E3	E2	E1	Average
VCA -FCLS	۲۱/۸۸	۶/۰۴	۱۶/۸۶	۳/۴۶	۱۱/۴۱	۵۹/۶۷
CNMF	۲۱/۶۶	۲۸/۹۴	۱۸/۴۸	۸/۵۶	۳۱/۴۲	۱۰۹/۰۸
MiniDisCoNMF	۱۰/۷۸	۱۲/۷۹	۱۲/۵۹	۷/۴۴	۳۱/۴۳	۷۵/۰۶
Proposed Method	۵/۳۲	۳/۵۵	۱۳/۱۳	۵/۵۸	۷/۹۷	۳۵/۵۶

(جدول ۴-۵) مقایسه معیار SID اعضای انتها بی استخراج شده و مرجع

Indiana Data	E5	E4	E3	E2	E1	Average
VCA -FCLS	۰/۰۱۶	۰/۰۰۵	۰/۲۱	۰/۱۱	۰/۳۳	۰/۷۳
CNMF	۰/۱۳	۰/۲۳	۰/۱۳	۰/۰۵	۰/۴۱	۰/۹۸
MiniDisCoNMF	۰/۰۴	۰/۰۷	۰/۰۶۶	۰/۰۳	۰/۴۱	۰/۶۴
Proposed Method	۰/۰۱۱	۰/۰۰۵	۰/۰۶۵	۰/۰۱	۰/۰۲	۰/۱۲

Wild S., Curry J., and Dougherty A., Improving non-negative matrix factorization through structured initialization, Pattern Recognit. , Nov. 2004, vol. 37, no. 11, pp. 2217–2232.

Winter Michael E., N-find: an algorithm for fast autonomous spectral endmember determination in hyperspectral data, in: Proceedings of the SPIE conference on Imaging Spectrometry V, 1999, vol.3753, pp.266-275.

Yuhas R. H., Goetz A. F. H., and Boardman J. W., Discrimination among semi-arid landscape endmember using the spectral angle mapper (SAM) algorithm, in Proc. Summaries3rd Annu. JPL Airborne Geosci. Workshop, 1992, vol. 1, pp. 147–149.



حبيب على زاده تحصیلات خود را در
قطعه کارشناسی مهندسی مخابرات
دانشگاه امام حسین (ع) در سال ۱۳۸۷
و در قطعه کارشناسی ارشد مهندسی
مخابرات دانشگاه تربیت مدرس در سال
۱۳۹۰ به پایان رسانده است. حوزه
تحقیقاتی وی عبارتند از: پردازش تصویر، پردازش تصاویر
سنگش از دور، جداسازی کور منابع.
نشانی رایانمۀ ایشان عبارت است از:

Habib.alizadeh@modares.ac.ir



حسن قاسمیان تحصیلات مقطع
کارشناسی ارشد و دکترای خود را در
مهندسی مخابرات و کامپیوتر دانشگاه
آمریکا بهترتبی در سال های
۱۳۶۳ و ۱۳۶۷ به اتمام رساند. ایشان
مدرک کارشناسی خود را در مهندسی
مخابرات از دانشگاه مخابرات در سال ۱۳۵۹ دریافت کرد.
وی در حال حاضر استاد گروه مهندسی مخابرات و مهندسی
پژوهشی دانشگاه تربیت مدرس است. زمینه های پژوهشی
مورد علاقه ایشان پردازش و آنالیز اطلاعات تصاویر و
سیگنال های دیجیتال، پردازش تصاویر سنگش از دور است.
نشانی رایانمۀ ایشان عبارت است از:

ghassemi@modares.ac.ir

Comon Pierre and Jutten Christian, Handbook of Blind Source Separation, 1st edition USA ,Elsevier, 2010.

Heinz Daniel C. and Chang Chein-I, Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery, IEEE Trans. Geosci. Remote Sens. , Mar. 2001, vol.39, no.3, pp.529-545

Horwitz H. M., Nalepk R. F., Hyde P. D., and Morgensten J. P., Estimating the proportions of objects within a single resolution element of a multispectral scanner, in Proc. 7th Int. Symp. Remote Sens. Environ, Ann Arbor, MI, 1971, pp. 1307–1320.

Huck Alexis, Guillaume Mireille., Jacques.Blan-Talon, Minimum Dispersion Constrained Non-negative Matrix Factorization to Unmix Hyperspectral Data, IEEE Trans. Geosci. Remote Sens., June 2010, vol. 48, no. 6, pp. 2590–2602.

Iordache Marian-Daniel, Bioucas-Dias José M., Antonio Plaza, Sparse Unmixing of Hyperspectral Data, IEEE Trans. Geosci. Remote Sens., June. 2011, vol. 49, no. 6, pp. 2014–2039

Keshava Nirmal and Mustard John F., Spectral unmixing, IEEE Signal Process. Mag., Jan. 2002 vol. 19, no. 1, pp. 44–57.

Kwon Heesung and Nasrabadi N. M., Kernel orthogonal subspace projection for hyperspectral signal classification, IEEE Trans. Geosci. Remote Sens., Dec. 2005, vol. 43, no. 12, pp. 2952–2962.

Liu Wen., Zheng N., and Li X., Review on nonnegative matrix factorization, Inst. Artif. Intell. Robot, Xi'an Jiaotong Univ., Xi'an, China, Tech. Rep., Sep. 2004.

Nascimento José M. P. and Dias José M. Bioucas, Does independent component analysis play a role in unmixing hyperspectral data?, IEEE Trans. Geosci. Remote Sens., Jan. 2005, vol. 43, no. 1, pp. 175–187.

Nascimento José M. P. and Dias José M. Bioucas, Vertex component analysis: A fast algorithm to unmix hyperspectral data, IEEE Trans. Geosci. Remote Sens, Apr. 2005., vol. 43, no. 4, pp. 898–910.

Pauca V. Paul, Piper J., Plemmons Robert J., Non-negative matrix factorization for spectral data analysis , Linear Algebra and its Applications,2006 , pp.29–47

Plaza A., Martinez P., Perez R., and Plaza J., A quantitative and comparative analysis of endmember extraction algorithms from Hyperspectral data, IEEE Trans. Geosci. Remote Sens., Mar. 2004, vol. 42, no. 3, pp. 650–663.

فصل نهم

