

روشی نوین به منظور طبقه‌بندی داده‌های چند بازگشتی لایدار با استفاده از اطلاعات هندسی مجاورتی و فضای پدیده

علیرضا صفدری‌نژاد^۱

محمدجواد ولدان‌زوج^۲

مهدی مختارزاده^۳

تاریخ پذیرش مقاله: ۹۴/۱۱/۲۸

تاریخ دریافت مقاله: ۹۴/۰۶/۱۱

چکیده

داده‌های اخذ شده توسط سیستم‌های لیزر اسکنر هوایی به دلیل برخورداری از مزایایی نظیر دقت هندسی نسبتاً بالا و تراکم مکانی بالای نقاط، اطلاعات هندسی متنوع و منحصر به فردی از سطوح فیزیکی عوارض فراهم می‌آورند. طبقه‌بندی و تفکیک داده‌های ابر نقطه به عوارض سازنده‌ی محیط، نقش مهمی در روند مدلسازی سه‌بعدی عوارض ایفا می‌کند. در مقاله پیش رو، مسأله‌ی تفکیک ابرنقاط بعنوان یک فرایند طبقه‌بندی نظارت‌شده مدنظر قرار گرفته شده است؛ روند اجرایی در روش پیشنهادی مبتنی بر سه گام بوده که در گام نخست، برای هر نقطه از ابرنقاط مجموعه‌ای از ویژگی‌ها مبتنی بر تحلیل‌های مجاورتی تولید می‌گردد. در گام دوم، ویژگی‌های بهینه به کمک داده‌های آموزشی و فضای پدیده استخراج شده و در نهایت، طی یک الگوریتم خوشه‌بندی، با استفاده از ویژگی‌های استخراج شده، داده‌های ابر نقطه به کلاس‌های مدنظر طبقه‌بندی می‌گردند. از این روش بمنظور طبقه‌بندی ابر نقاط چندبازگشتی لایدار مربوط به یک منطقه‌ی شهری استفاده شد که نتایج طبقه‌بندی، دقت کلی معادل ۹۳/۱۵ درصد و ضریب کاپای ۰/۸۹ را نشان دادند.

واژه‌های کلیدی: لیزر اسکنر هوایی، خوشه‌بندی، تحلیل مجاورت، انتخاب ویژگی، فضای پدیده، فضای ویژگی

۱. دانشجوی دکترای مهندسی نقشه‌برداری، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی خواجه‌نصیرالدین طوسی safdari_nezhad@sina.kntu.ac.ir

۲. استادیار گروه فتوگرامتری و سنجش از دور، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی خواجه‌نصیرالدین طوسی m_mokhtarzade@kntu.ac.ir

۳. استاد گروه فتوگرامتری و سنجش از دور، دانشکده مهندسی ژئوماتیک، دانشگاه صنعتی خواجه‌نصیرالدین طوسی valadanzouj@kntu.ac.ir

۱- مقدمه

نظارت شده مد نظر قرار گرفت. در فرآیند خوشه‌بندی، یک مجموعه داده ناهمگن با ویژگی‌های متفاوت، به چندین زیرمجموعه همگن که هر یک دارای ویژگی‌های مشابهی هستند، تقسیم‌بندی می‌شوند. (Richards & et al, 2006) بدیهی است که برای نیل به این هدف، نخست باید به تعریف «ویژگی» و درگام بعد، اندازه‌گیری «معیارهای شباهت بین ویژگی‌ها» پرداخت.

با توجه به ماهیت ابر نقاط لایدار (مجموعه‌ای از نقاط سه‌بعدی)، هر نقطه با سه ویژگی (مؤلفه‌های مختصات سه‌بعدی) شناخته شده که به تنهایی کفایت لازم به منظور تفکیک ابرنقاط به کلاس‌های سازنده در محیط را برخوردار نخواهند بود. بدیهی است که موقعیت قرارگیری هر نقطه در محیط پیرامون خود، اطلاعات مناسبی به منظور تشخیص کلاس هر نقطه را فراهم می‌آورد.

از این رو، در روش پیشنهادی، پس از تعریف چند سطح همسایگی در فضای دوبعدی و سه بعدی، براساس تحلیل‌های مجاورتی، مجموعه‌ای از ویژگی‌های آماری و هندسی برای تمام نقاط موجود در ابر نقاط محاسبه شدند. در ادامه، انتخاب ویژگی‌های بهینه در بین ویژگی‌های استخراجی بکمک فضای پدیده^۳ (Richards & et al, 2006) صورت پذیرفت که در نهایت بعد از شناسایی بردار ویژگی‌های^۴ بهینه، از یک خوشه‌بندی مبتنی بر فاصله‌ی اقلیدسی در فضای ویژگی به منظور طبقه‌بندی نقاط موجود در ابر نقاط استفاده شد. با این مقدمه، مقاله حاضر در چهار بخش اصلی تدوین شده است. در بخش نخست، مقدمه‌ای در ارتباط با اهمیت خوشه‌بندی داده‌های لایدار و خلاصه‌ای از روند اجرایی این تحقیق ارائه گردید. در بخش دوم، پس از معرفی داده‌های مورد استفاده و روش‌های خوشه‌بندی نقاط، روش پیشنهادی با جزئیات بیشتر تشریح شده است. نتایج پیاده‌سازی در بخش سوم ارائه شده و در نهایت، در بخش چهارم و پنجم به بحث در خصوص نتایج بدست آمده و ارائه پیشنهاداتی برای تحقیقات آتی پرداخته شده است.

لیزر اسکنرهای هوایی^۱ به عنوان ابزارهایی نوین در حوزه تهیه اطلاعات سه بعدی از سطح زمین محسوب می‌شوند. (Ghanma, 2006) اصول اندازه‌گیری این سیستم‌های فعال مبتنی بر اندازه‌گیری فاصله بوده و طی فرآیند نمونه‌برداری، ابر نقاط نامنظمی از سطح زمین را جمع‌آوری می‌کنند. ذات و ماهیت داده‌های این سیستم نقطه‌ای بوده و قابلیت ثبت چند بازگشت از پالس ارسال را برخوردارند.

کاربردهای متنوعی از داده‌های لایدار در حوزه‌های علوم ژئوماتیک توسعه یافته که مواردی همچون: ۱- مدیریت بحران و تحلیل‌های تعجیلی مثل تخمین خسارات ناشی از زلزله و سیل، (Alharthy & et al, 2004; 213-218) ۲- تولید انواع مدل‌های رقومی ارتفاعی (Axelsson, 2000 & Ghanma, 2006) ۳- استفاده در الگوریتم‌های متنوع در حوزه استخراج اتوماتیک عوارض، (Alharthy & et al, 2004, 213-218) ۴- بررسی زیست‌توده‌های گیاهی (Zimble & et al, 2003, 171-182) و ... نمونه‌هایی از کاربردهای این داده‌ها بحساب می‌آیند.

با این وجود، بواسطه‌ی: ۱- تنوع و پیچیدگی عوارض سه‌بعدی (طبیعی و مصنوعی)، ۲- حجم و تراکم بالای ابر نقاط و همچنین ۳- کاربردهای مد نظر از این داده‌ها، لازم است قبل و یا حتی در مرحله استخراج اطلاعات، این داده‌ها به نحو مطلوبی سازماندهی شوند. اغلب روش‌های سازماندهی ابر نقاط لایدار^۲، به نحوی سعی در تفکیک نقاط مربوط به عوارض مشابه در گروه‌های مجزا (Filin, 2006: 71-80) دارند.

ایده‌ی اصلی این مقاله، پاسخگویی به این سؤال بود که آیا، صرفاً با در نظر گرفتن مختصات سه‌بعدی نقاط در ابر نقاط لایدار و بدون استفاده از سایر داده‌های کمکی (تصویر، نقشه، مدل عوارض و ...) می‌توان ابر نقاط خام لایدار را به صورت نظارت‌شده طبقه‌بندی کرد؟

با این هدف، تفکیک نقاط مربوط به عوارض مختلف موجود در ابر نقاط لایدار، به عنوان یک مسأله خوشه‌بندی

3- Prototype Space

4- Features

1- Airborne Laser Scanner (LiDar)

2- LiDar Point Clouds

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی ()
روشی نوین به منظور طبقه‌بندی داده‌های چندبازگشتی ... / ۱۷

همچنین فضای محاسباتی مورد بررسی قرار گرفته‌اند. از نقطه نظر روش‌های طبقه‌بندی، استخراج زمین لخت در مرحله اول (فیلترگذاری)، و سپس جداسازی ساختمان و درخت در نقاط غیرزمینی حاصل از مرحله نخست را می‌توان به عنوان یکی از روش‌های رایج در این زمینه بشمار برد. ایده اصلی استخراج زمین لخت، فیلترگذاری ابر نقاط خام لایدار و حذف نقاط غیرزمینی است. تاکنون فیلترهای متنوعی برای این منظور توسعه یافته که برای آشنایی با جزئیات فیلترها خواننده می‌تواند به منابع شماره ۸ تا ۱۳ مراجعه نماید. معمولاً، تفکیک سایر عوارض بر روی مجموعه نقاط غیرزمینی حاصل از فیلترگذاری صورت می‌پذیرد. روش‌های مختلفی برای این منظور توسعه یافته، که در برخی از این روش‌ها، تنها از نقاط لایدار استفاده می‌شود؛ و در برخی دیگر، داده‌های لایدار به همراه داده‌های اضافی نظیر تصاویر و نقشه‌ها مورد پردازش قرار می‌گیرند. (منابع شماره ۱۴ تا ۱۸)

از نقطه نظر فضای محاسباتی، پردازش داده‌های لایدار با دو رویکرد متفاوت صورت می‌پذیرند. در رویکرد اول، طی فرآیند نمونه‌برداری مجدد و تکنیک‌های درون‌یابی مکانی، مجموعه نقاط سه‌بعدی به یک شبکه منظم نگاشت شده و این شبکه مبنای تحلیل‌های آتی قرار می‌گیرد. (Filin & etal, 2006: 71-80)

بدیهی است که در این رویکرد، استفاده از روش‌های پردازش تصاویر رقومی و تشخیص الگو با سهولت بیشتری میسر بوده و این سهولت اجرا را می‌توان از مزایای این رویکرد قلمداد نمود. اما نمونه‌برداری مجدد نقاط در مواردی (نظیر نقاط واقع بر بدنه ساختمان‌ها، نقاط ثبت شده در بالا و پائین پل‌ها و غیره) موجب از دست رفتن بخشی از اطلاعات شده که تا حدی امکان استفاده حداکثری از پتانسیل داده‌های سه‌بعدی لایدار را محدود می‌سازد.

اما در رویکرد دوم، ماهیت نامنظم داده‌های سه‌بعدی حفظ شده و تمامی تحلیل‌ها با حفظ اصالت و ماهیت خام داده‌ها صورت می‌گیرد. با توجه به اینکه در این رویکرد هیچ نوع تغییری در ماهیت داده‌ها صورت نمی‌پذیرد، مشکل

۲- مواد و روش‌ها

۲-۱- داده‌های مورد استفاده

داده‌های چند بازگشتی مورد استفاده در این تحقیق مربوط به منطقه‌ای شهری با وسعت تقریبی ۱۲۰۰۰ مترمربع بوده که شامل سه نوع عارضه کلی ساختمان، درخت و سطح طبیعی زمین می‌باشد. داده‌های مذکور توسط یک لیزراسکنر هوایی اخذ شده که بطور متوسط تراکم نمونه‌برداری معادل نیم متر را برخوردار است (نگاره ۱). سقف ساختمان‌ها در این منطقه از دو نوع شیب‌دار و مسطح بوده که سقف‌های شیب‌دار در زوایای مختلف در منطقه حضور دارند. از سوی دیگر، بواسطه چند بازگشتی بودن داده‌ها، بازگشت‌های متعددی از حجم درختان ثبت شده است.

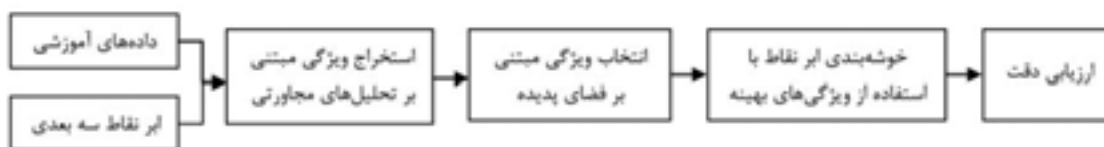


نگاره ۱: نمای کلی از منطقه‌ی مورد مطالعه

۲-۲- روش‌های خوشه‌بندی ابر نقاط لایدار

در یک منطقه‌ی شهری، ابر نقاط نامنظم اخذ شده توسط لیزر اسکنرهای هوایی شامل نقاط بازگشتی از عوارض دست‌ساز بشر (نظیر ساختمان‌ها، ماشین‌ها و غیره) و سطوح طبیعی (مانند سطح زمین، درختان و غیره) می‌باشند. طبقه‌بندی داده‌های ابر نقطه‌ای بعنوان یکی از ضروریات فرایند مدل‌سازی سه‌بعدی بشمار می‌رود. بعبارت دیگر، لازم است که پیش از هرگونه مدل‌سازی و استفاده عملی از این نقاط در یک سیستم اطلاعات مکانی، در وهله نخست نقاط مربوط به عوارض مختلف از هم تفکیک شوند. (Alharthy & etal, 2004: 213-218)

در این بخش، بطور مختصر روش‌های خوشه‌بندی داده‌های لایدار از دو نقطه نظر روش‌های طبقه‌بندی و



نگاره ۲: روند کلی متدولوژی پیشنهادی

جدول ۱: ویژگی‌های مجاورتی تولید شده به ازای نقاط همسایه در هر سطح همسایگی

همسایگی سه‌بعدی					همسایگی دوبعدی					ویژگی
۱۰۰ تایی	۷۵ تایی	۵۰ تایی	۲۵ تایی	۱۰ تایی	۱۰۰ تایی	۷۵ تایی	۵۰ تایی	۲۵ تایی	۱۰ تایی	
*	*	*	*	*	*	*	*	*	*	انحراف معیار مؤلفه ارتفاعی نقاط همسایگی
*	*	*	*	*	*	*	*	*	*	دامنه تغییرات مؤلفه ارتفاعی نقاط در همسایگی
*	*	*	*	*	*	*	*	*	*	میانگین مؤلفه ارتفاعی نقاط
*	*	*	*	*						همبستگی مؤلفه‌های مسطحاتی نقاط
*	*	*	*	*	*	*	*	*	*	فاصله از دورترین نقطه همسایگی (وسعت همسایگی)
*	*	*	*	*						انحراف معیار فواصل نقاط در همسایگی (فاصله سه‌بعدی)
*	*	*	*	*	*	*	*	*	*	زاویه زنبیتی نرمال صفحه برازش یافته به نقاط همسایگی
*	*	*	*	*	*	*	*	*	*	RMSE نقاط همسایگی از صفحه برازش یافته به آنها
*	*	*	*	*	*	*	*	*	*	دامنه اختلاف ارتفاع نقاط همسایه از صفحه برازش یافته
*	*	*	*	*	*	*	*	*	*	فاصله نقطه مورد نظر از مرکز ثقل همسایگی (میانگین)

میان نقاط پیرامون خود (اطلاعات توپولوژیک مربوط به هر نقطه) قابل تشخیص خواهد بود.

از این رو، بنظر می‌رسد تولید اطلاعاتی که بنحوی گویای روابط توپولوژیک مربوط به هر نقطه در ابر نقاط باشند، کمک شایانی به شناسایی کلاس نقاط نمایند. در همین راستا، در گام نخست، برای هر یک از نقاط موجود در ابر نقاط، مجموعه‌ای از ویژگی‌های مبتنی بر تحلیل‌های مجاورتی تولید می‌گردد. هدف از این کار، آشکارسازی الگوهای مستتر توپولوژیکی در ابر نقاط است که به خودی خود و تنها با استفاده از مختصات سه‌بعدی آنها قابل شناسایی نخواهند بود.

برای این هدف و به منظور تولید ویژگی‌های مجاورتی، به ازای هر نقطه از ابر نقاط پنج سطح همسایگی شامل ۱۰، ۲۵، ۵۰، ۷۵ و ۱۰۰ نقطه نزدیک در نظر گرفته شد. برای یافتن نقاط همسایه در هر سطح همسایگی به طور مجزا،

حذف ناخواسته بخشی از داده‌ها نیز به خودی خود منتفی خواهد بود. حجم بالای پردازشی و همچنین پیچیدگی‌های اجرایی در تحلیل‌های مکانی داده‌ها در رویکرد دوم را می‌توان بعنوان مشکلات این رویکرد بحساب آورد.

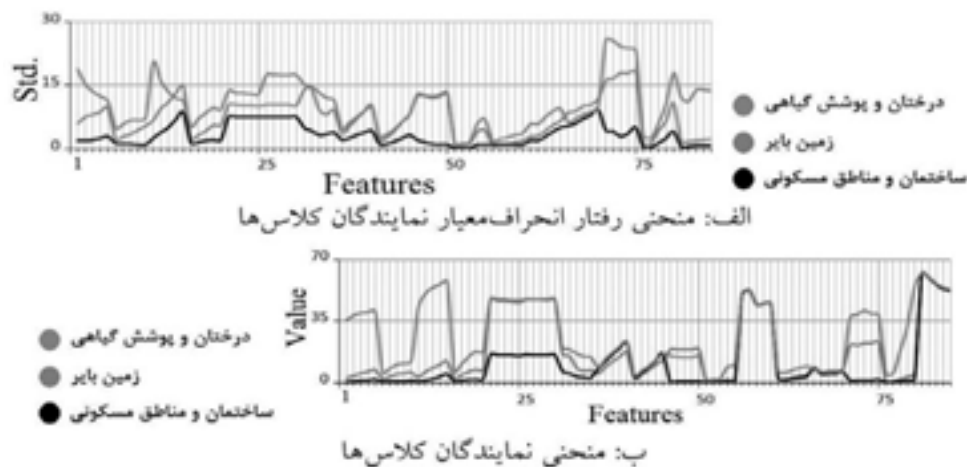
۲-۳- روش پیشنهادی

در تحقیق حاضر، خوشه‌بندی ابر نقاط نامنظم سه‌بعدی اخذ شده توسط سیستم لیزر اسکنر هوایی، طی یک فرآیند سه مرحله‌ای و به صورت نظارت‌شده صورت گرفته است. روند کلی اجرای روش پیشنهادی را می‌توان در نگاره (۲) مشاهده نمود.

۲-۳-۱- استخراج ویژگی‌های مجاورتی

از نقطه نظر تشخیص بصری عوارض موجود در یک ابر نقطه، کلاس هر نقطه بواسطه‌ی موقعیت قرارگیری آن در

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی ()
روشی نوین به منظور طبقه‌بندی داده‌های چندبازگشتی ... / ۱۹



نگاره ۳: بردار ویژگی نمایندگان داده‌های آموزشی انحراف معیار نمایندگان هر کلاس

کلاس‌های مورد نظر شناسایی شوند. نحوه انتخاب ویژگی‌های بهینه، در بخش بعد مورد بررسی قرار خواهد گرفت.

۲-۳-۲- انتخاب ویژگی‌های بهینه

از آنجایی که اجرای یک طبقه‌بندی نظارت شده بعنوان هدف غایی این تحقیق در نظر گرفته شده بود، تعدادی نقطه از ابر نقاط برای هر کلاس به عنوان داده‌های آموزشی انتخاب شدند. در نگاره (۳) بردار ویژگی نمایندگان داده‌های آموزشی هر کلاس (میانگین داده‌های آموزشی) و بردار انحراف معیار نمایندگان هر کلاس در فضای Signature ترسیم شده است.

همانطور که در نگاره (۳) قابل مشاهده است، برخی از ویژگی‌ها پراکندگی بیشتری را نسبت به نماینده‌ی کلاس برخوردار بوده و برخی از ویژگی‌ها تمرکز بالایی را نسبت به نمایندگان کلاس از خود نشان می‌دهند. از سوی دیگر با نگاه به منحنی نمایندگان کلاس می‌توان دریافت که برخی از ویژگی‌ها قابلیت تفکیک کلاس‌های مد نظر را نداشته و مقادیر نمایندگان کلاس‌ها برای آنها کاملاً برهم منطبقند. اما در مقابل ویژگی‌هایی نیز وجود داشته که نمایندگان هر کلاس در آن ویژگی مقادیر متفاوتی را به خود اختصاص داده و تفکیک پذیرند. بدیهی است که انتخاب همه‌ی ویژگی‌های استخراج شده در روند طبقه‌بندی، دستاوردی

از فاصله اقلیدسی دوعدی و سه‌بعدی بین نقاط استفاده گردید. سپس به ازای هر نقطه، مجموعه‌ای از ویژگی‌های هندسی و آماری بوسیله‌ی نقاط همسایه تولید گردید. جدول (۱) ویژگی‌های مجاورتی تولید شده در هر سطح همسایگی را نشان می‌دهد.

بدیهی است که می‌توان طیف گسترده‌تری از ویژگی‌های مجاورتی را بکمک تحلیل‌های هندسی و مجاورتی استخراج نمود؛ که شناسایی، طراحی و مدلسازی این ویژگی‌ها می‌تواند بعنوان یک موضوع تحقیقاتی مجزا مورد کاوش محققین قرار گیرد. با این حال، بنظر می‌رسد با توجه به خصوصیات منطقه‌ی مورد مطالعه، ویژگی‌های استخراج شده، شاخص‌های مناسبی به منظور تمایز توپولوژیکی نقاط در کلاس‌های مد نظر باشند. بر این اساس مطابق جدول (۱)، در مجموع ۹۰ ویژگی ساده هندسی متمایز برای هر نقطه تولید گردید.

نکته‌ی مهم در بکارگیری شاخص‌های تولیدی، عدم هم‌مقیاس بودن مقادیر ویژگی‌های تولیدی بوده که این امر در صحت طبقه‌بندی کننده‌های مبتنی بر محاسبه‌ی فاصله اقلیدسی در فضای ویژگی تأثیرگذار خواهد بود. به همین دلیل، به منظور هم‌مقیاس شدن محورهای فضای ویژگی، تمامی ویژگی‌های استخراجی بطور مستقل در بازه‌ی یکسانی نرمال شدند. پیش از استفاده از این ویژگی‌ها در فرآیند خوشه‌بندی، لازم است ویژگی‌های مناسب برای جدا نمودن

در مقاله حاضر، برای انتخاب ویژگی‌های بهینه، افزون بر فاصله ویژگی‌ها از قطر اصلی فضای پدیده دو شرط دیگر نیز مورد توجه قرار گرفت: ۱- انحراف معیار ویژگی‌ها برای هر کلاس؛ که به نحوی معیاری از تعمیم‌پذیری رفتار تفکیک‌پذیری ویژگی در خصوص داده‌های مورد طبقه‌بندی خواهد بود؛ و ۲- فاصله ویژگی‌ها از یکدیگر در فضای پدیده؛ با توجه به این واقعیت که موقعیت مشابه دو ویژگی در فضای پدیده به معنای رفتار تفکیک‌پذیری مشابه آنها در ارتباط با کلاس‌های مورد طبقه‌بندی است، حتی اگر فاصله این دو ویژگی از قطر اصلی فضای پدیده زیاد باشد باز هم می‌توان یکی از آنها را حذف نمود. به عبارت دیگر، ویژگی‌های نزدیک به هم در فضای پدیده حتی اگر فاصله زیادی از قطر اصلی فضای پدیده داشته باشند، موجب رفتار تفکیک‌پذیری متمایزی نخواهند شد (همان).

میزان بهینه بودن هر ویژگی به طور منفرد در فضای پدیده، از حاصل تقسیم فاصله‌ی ویژگی تا قطر اصلی فضای پدیده به شاخص پراکندگی ویژگی (متوسط انحراف معیار ویژگی به ازای کلاس‌های طبقه‌بندی) بدست خواهد آمد. اما این شاخص، ملاک مناسبی برای انتخاب چندین ویژگی در فضای پدیده نیست. علت این امر نیز بواسطه‌ی عدم لحاظ نمودن فاصله‌ی بین ویژگی‌ها در فضای پدیده خواهد بود. بر این اساس در روش پیشنهادی، اولین ویژگی بر اساس بیشینه‌ی شاخص منفرد انتخاب شده؛ و در انتخاب ویژگی‌های بعدی، فاصله‌ی ویژگی کاندید شده نسبت به ویژگی‌های انتخاب شده‌ی قبلی نیز تأثیرگذار خواهد بود. روند اجرایی انتخاب ویژگی در فضای پدیده در ادامه تشریح شده است.

در گام نخست اولین ویژگی بهینه براساس بیشینه رابطه

$$\vec{F}_1 = \frac{\vec{d}}{STD_{mean}} \quad (1) \quad \text{قابل شناسایی خواهد بود:}$$

که در آن، \vec{d} بردار فاصله ویژگی‌ها از قطر اصلی فضای پدیده و STD_{mean} بردار متوسط انحراف معیار ویژگی‌ها (هر ویژگی به ازای هر کلاس دارای یک انحراف معیار خواهد بود)، و \vec{F}_1 بردار فاصله وزندار ویژگی‌ها از قطر اصلی

جز کاهش دقت طبقه‌بندی نداشته باشند. از این روی، اجرای فرایند انتخاب ویژگی‌های بهینه به نحو مؤثری بر دقت نتایج تأثیرگذار خواهد بود.

انتخاب ویژگی^۱ در فرآیند طبقه‌بندی از سه جنبه حائز اهمیت است. از طرفی، افزایش ویژگی‌های مورد استفاده در فرآیند طبقه‌بندی موجب افزایش هزینه محاسباتی طبقه‌بندی می‌گردد، و این امر به نوبه خود موجب کاهش کارایی طبقه‌بندی کننده می‌شود. از طرف دیگر، با افزایش ویژگی‌های مورد استفاده در طبقه‌بندی، برای پرهیز از وقوع پدیده Hughes، نیاز به داده‌های آموزشی نیز افزایش می‌یابد. (Richards & etal, 2006) در نهایت، وجود ویژگی‌هایی که تفکیک‌پذیری مناسبی میان کلاس‌های مورد نظر ایجاد نمی‌کنند، می‌توانند موجب برآورد نادرست فراصفحه جداکننده^۲ کلاس‌ها و کاهش دقت طبقه‌بندی گردند. تکنیک‌های متنوعی جهت انتخاب ویژگی در پردازش‌های سنجش از دوری توسعه یافته‌اند. (مجردی، ۱۳۸۸) مجردی در سال ۲۰۰۹ یک روش نوین و کارآمد به منظور انتخاب ویژگی از تصاویر ابرطیفی ابداع نمود. وی فضای مورد بحث را «فضای پدیده» نامید و کارایی و کفایت آنرا در تصاویر ابرطیفی مورد ارزیابی قرار داد (همان). در این تحقیق نیز، از مفهوم بکار رفته در فضای پدیده بمنظور انتخاب ویژگی از تصاویر ابرطیفی، بمنظور انتخاب ویژگی‌های بهینه در این تحقیق استفاده شده است.

در فضای پدیده، هر کلاس به منزله یکی از محورهای سیستم مختصات متعامد تعریف کننده فضا، و هر ویژگی مشخص کننده یک نقطه در این فضا خواهد بود. به این ترتیب، فاصله نمایندگان کلاس‌ها از قطر اصلی این فضا می‌تواند به عنوان معیاری از تفکیک‌پذیری ویژگی‌ها در ارتباط با کلاس‌های مورد نظر تعبیر شود. در این حالت، هر چه فاصله یک ویژگی از قطر اصلی فضای پدیده بیشتر باشد، این ویژگی تفکیک‌پذیری بیشتری میان کلاس‌های مورد نظر ایجاد خواهد نمود.

1- Feature Selection

2- Discriminate Hyper-Plane

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی ()
روشی نوین به منظور طبقه‌بندی داده‌های چندبازگشتی ... / ۲۱

۳- پیاده‌سازی و نتایج

از مجموع ۹۰ ویژگی تولید شده برای هر نقطه، چهار ویژگی بعنوان ویژگی‌های بهینه شناسایی شدند. سلول ویژگی‌های بهینه در جدول (۱) با تفاوت در رنگ نمایش داده شده‌اند. بواسطه عواملی مثل: ۱- اطمینان از تفکیک پذیری ویژگی‌ها و ۲- یکسان بودن مقیاس ویژگی‌ها، از الگوریتم خوشه‌بندی نظارت نشده K-Means به عنوان الگوریتم طبقه‌بندی استفاده شده است. نگاره (۵) نتایج طبقه‌بندی ابر نقاط را نشان می‌دهد.

در نهایت، عملکرد طبقه‌بندی کننده در سطح داده‌های چک مستقل مورد ارزیابی قرار گرفت (جدول ۲)، که نتایج ارزیابی حاکی از دقت کلی ۹۳/۱۵ درصد و ضریب کاپا ۰/۸۹ در سطح داده‌های چک بود. براساس نتایج به دست آمده، این روش دقت خوبی را در جداسازی نقاط مربوط به دو کلاس زمین و ساختمان نشان داد.

جدول ۲: ماتریس ابهام نتایج خوشه‌بندی

مجموع	زمین	درخت	ساختمان	
۲۸۶۰	۴۵	۱۹۱	۲۶۲۴	ساختمان
۲۰۳۰	۷	۱۸۶۶	۱۵۷	درخت
۱۲۴۲	۱۲۲۲	۲۰	۰	زمین
	۱۲۷۴	۲۰۷۷	۲۷۸۱	مجموع

نتایج مطلوب در طبقه‌بندی، نشان دهنده‌ی این موضوع بوده که ویژگی‌های انتخاب شده به حدی تفکیک پذیر هستند که یک خوشه بندی ساده‌ی مبتنی بر فاصله‌ی اقلیدسی در فضای ویژگی به نحو مؤثری منجر به تفکیک کلاس‌های مورد نظر شده است.

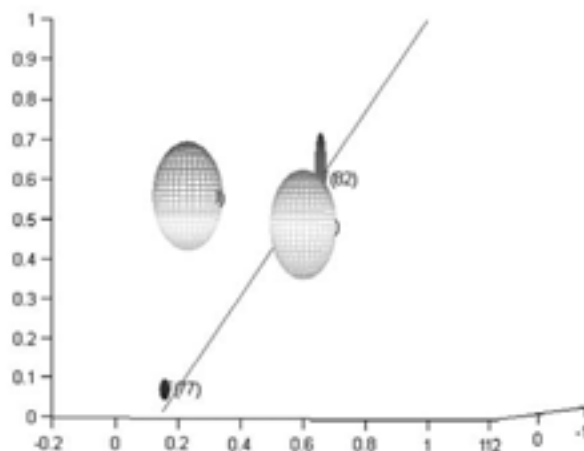
از طرفی این نکته را باید در نظر داشت که استفاده از طبقه بندی کننده‌ی K-Means در این روش نشان‌دهنده نظارت نشده بودن نیست؛ بلکه داده‌های آموزشی در این الگوریتم در انتخاب ویژگی حضور داشته که منجر به شناسایی ویژگی‌های تفکیک‌پذیر شده‌اند.

فضای پدیده می‌باشد. طبیعتاً در انتخاب نخستین ویژگی بهینه فاصله از سایر ویژگی‌ها مفهومی ندارد، به همین جهت این پارامتر در رابطه (۱) مشاهده نمی‌شود. پس از مشخص شدن نخستین ویژگی بهینه، برای انتخاب سایر ویژگی‌ها یک معیار بهینگی به فرم رابطه (۲) پیشنهاد گردید:

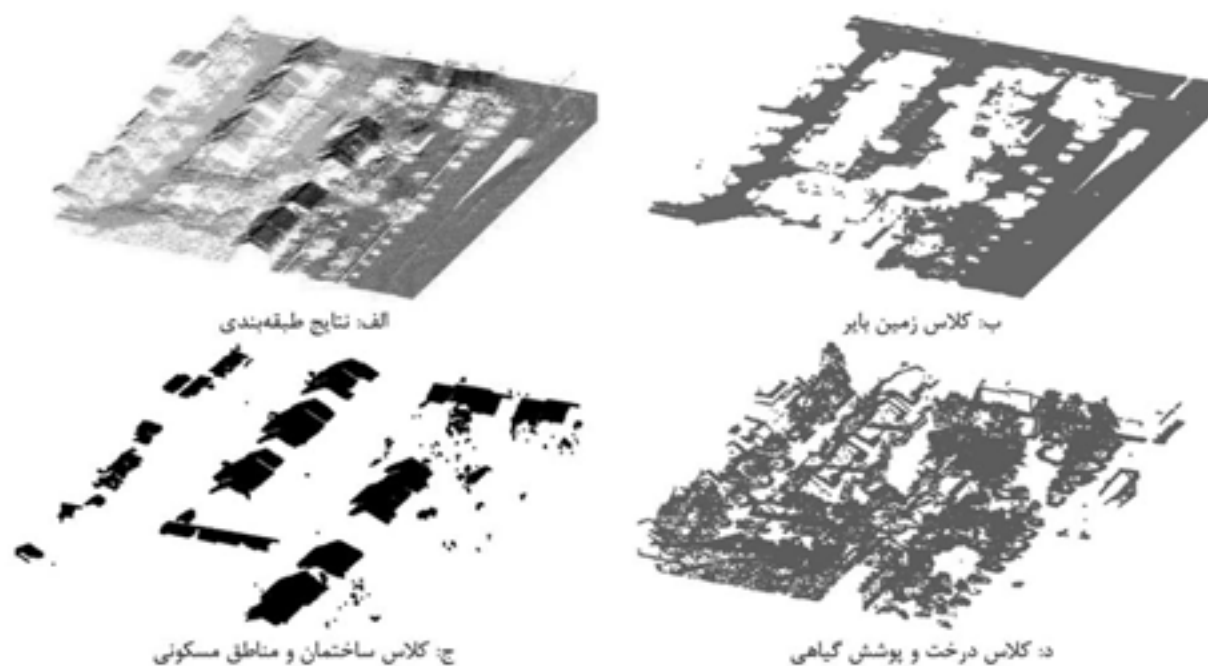
$$\vec{F}_i = \max \left(\vec{F}_1 \bullet \frac{\sum_{k=1}^{i-1} \vec{d}_k}{i-1} \right), \quad i = 2, 3, \dots, n \quad (2)$$

که در این رابطه، \vec{d}_k بردار فاصله ویژگی‌ها از ویژگی انتخاب شده در مراحل قبل (1, 2, ..., i-1) و n تعداد ویژگی‌هایی است که قرار است انتخاب شوند. لازم به توضیح است که در این رابطه، عملگر "•" نشان دهنده ضرب درایه به درایه می‌باشد.

در نهایت، طی یک فرایند افزایشی در انتخاب ویژگی، و بر اساس دقت نتایج طبقه‌بندی، تعداد ویژگی‌های بهینه شناسایی می‌گردند. بر اساس مطالب گفته شده، چهار ویژگی بهینه از بین ۹۰ ویژگی تولیدی انتخاب شدند. فضای پدیده ویژگی‌های مذکور و همچنین بیضوی پراکندگی آنها در نگاره (۴) ترسیم شده است. همانطور که مشاهده می‌شود تفکیک‌پذیری بین کلاسی و داخل کلاسی خوبی بین ویژگی‌های انتخابی برقرار می‌باشد.



نگاره ۴: ویژگی‌های انتخاب شده در فضای پدیده



نگاره ۵: نتایج طبقه‌بندی ابر نقاط توسط الگوریتم پیشنهادی

۴- نتیجه‌گیری

و بعضاً پیچیده‌تر بتوان دقت خوشه‌بندی را در این نواحی نیز بهبود بخشید؛ که این موضوع می‌تواند به عنوان یکی از افق‌های مطالعات آتی در این زمینه به حساب آید.

در این مقاله روشی ساده و کارآمد جهت تفکیک و طبقه‌بندی ابر نقاط LiDAR ارائه شد. نتایج این تحقیق نقش مهم اطلاعات همجواری نقاط در تعیین کلاس‌های طبقه‌بندی را به اثبات رساند.

۵- پیشنهادات

با توجه به نتایج و تجربیات اجرایی بدست آمده در این تحقیق، مواردی همچون: ۱- طراحی و اجرای الگوریتم‌های حذف نویز در نتایج طبقه‌بندی بمنظور افزایش صحت نتایج، ۲- توسعه و طراحی ویژگی‌های غنی‌تر از نقطه‌نظر محتوای اطلاعاتی مبتنی بر همجواری نقاط در ابر نقاط سه بعدی، ۳- اجرای روندی سلسله مراتبی بمنظور طبقه‌بندی ابر نقاط، مبتنی بر روش پیشنهادی، ۴- استفاده از اطلاعات کمکی مثل Intensity پالس بازگشتی و شمارنده‌ی بازگشت در روند طبقه‌بندی، ۵- توسعه‌ی الگوریتم‌های مبتنی بر پردازش موازی در روند تولید ویژگی‌های مجاورتی و ۶- ارزیابی سطح کارآمدی این متد در مناطق پیچیده و با تنوع کلاس‌های سازنده را می‌توان بعنوان افق‌های تحقیقی در پژوهش‌های آتی مطرح نمود.

با توجه به توانایی این روش در شناسایی ویژگی‌های مجاورتی بهینه بر اساس ویژگی‌های منطقه و داده‌های آموزشی، انتظار می‌رود روش پیشنهادی از تعمیم‌پذیری مناسبی در مناطق مختلف برخوردار باشد.

با توجه به حصول دقت‌های بالا در جداسازی نقاط زمین لخت و ساختمان، این روش می‌تواند به عنوان جایگزین روش‌های فیلترگذاری در کاربردهایی نظیر تولید مدل رقومی ارتفاعی و مدلسازی سه‌بعدی ساختمان‌ها مورد استفاده قرار گیرد. در مورد نقاط مربوط به درختان، با توجه به پیچیدگی ارتباط هندسی نقاط همسایه، احتمال می‌رود وجود خطاهای اندک در این نواحی به دلیل ساده‌انگاری طراحی ویژگی‌های مورد استفاده در فرآیند خوشه‌بندی باشد. از اینرو، انتظار می‌رود با تولید ویژگی‌های مؤثرتر

فصلنامه علمی - پژوهشی اطلاعات جغرافیایی ()
روشی نوین به منظور طبقه‌بندی داده‌های چندبازگشتی ... / ۲۳

from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(1-2), August, 85-101.

12- Shan, J. and Sampath, A., 2005. Urban DEM generation from raw lidar data: A labelling algorithm and its performance. *Photogrammetric Engineering and Remote Sensing*, 71(2), 217-226.

13- Zhang, K. and Whitman, D., 2005. Comparison of three algorithms for filtering airborne LiDAR data. *Photogrammetric Engineering and Remote Sensing*, 71(3), 313-324.

14- Haala, N. and Brenner, C., 1999. Virtual city models from laser altimeter and 2D map data. *Photogrammetric Engineering and Remote Sensing*, 65(7), 787-795.

15- Vosselman, G. and Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. *International Archives of Photogrammetry and Remote Sensing*, Vol. 34(3W4), pp. 37-43.

16- Overby, J., Bodum, L., Kjems, E., and Iisoe, P.M., 2004. Automatic 3D building reconstruction from airborne laser scanning and cadastral data using Hough transform. *International Archives of Photogrammetry and Remote Sensing*, Vol. 35, Part. B3, pp. 296-301.

17- Hofmann, A.D., Maas, H.-G., and Streilein, A., 2002. Knowledge-based building detection based on laser scanner data and topographic map information. Available at <http://www.isprs.org/commission3/proceedings02/papers/paper025.pdf>.

18- Schwalbe, E., Maas, H.-G., and Seidel, F., 2005. 3D building model generation from airborne laser scanner data using 2D GIS data and orthogonal point cloud projections. *International Archives of Photogrammetry and Remote Sensing*, Vol. 36, Part 3/W19, 209-214.

منابع و مأخذ

- ۱- مجردی، برات، ۱۳۸۸. استخراج ویژگی به منظور طبقه‌بندی تصاویر ابرطیفی، رساله دکتری، دانشکده ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه نصیرالدین طوسی.
- 2- Ghanma, M., 2006. Integration of photogrammetry and LiDAR. PhD thesis, Department of Geomatics Engineering, University of Calgary, April 2006.
- 3- Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models. *IAPRS*, XXXIII B4/1, Istanbul, Turkey.
- 4- Alharthy, A. and Bethel, J., 2004. Detailed building reconstruction from airborne laser data using a moving surface method. *International Archives of the Photogrammetry and Remote Sensing*, Vol. XXXV-B3, pp. 213-218.
- 5- Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., and Gerard, P.D., 2003. Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment*, 87(2-3), 171-182.
- 6- Filin, S., Pfeifer, N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. *ISPRS Journal of Photogrammetry & Remote Sensing* 60 (2006) 71-80.
- 7- Richards, J.A., and Jia, X., 2006. *Remote Sensing Digital Image Analysis: An Introduction*. Fourth Edition, Springer Publication (Germany).
- 8- Kilian, J., Haala, N., and English, M., 1996. Capture and evaluation of airborne laser scanner data, *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXI, Part B3, Vienna, pp. 383-388.
- 9- Kraus, K. and Pfeifer, N., 1998. Determination of terrain models in wooded areas with aerial laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(4), 193-203.
- 10- Vosselman, G., 2000. Slope based altering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, Vol. 33, Part B3/2, Amsterdam, the Netherlands, pp. 935-942.
- 11- Sithole, G. and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction