

رویکردی نو در پیاده‌سازی الگوریتم‌های خلاصه‌سازی هندسی عوارض خطی با هدف کنترل بر میزان ساده‌سازی

امیر غلامی^۱، ابوالفضل قنبری^{۲*}، ندا کفاش چرندابی^۳

۱- دانشجوی ارشد سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده برنامه‌ریزی و علوم محیطی، دانشگاه تبریز

۲- استاد دانشکده برنامه‌ریزی و علوم محیطی، دانشگاه تبریز

۳- استادیار دانشکده فنی و مهندسی مهند، دانشگاه تبریز

تاریخ دریافت مقاله: ۱۴۰۳/۰۵/۰۵ تاریخ پذیرش مقاله: ۱۴۰۳/۰۸/۲۳

چکیده

ساده‌سازی هندسی عوارض خطی از اوایل سال ۱۹۰۰ میلادی مورد توجه جغرافی‌دانان و کارتوگراف‌ها قرار گرفت و با گذشت زمان و افزایش حجم تولید داده‌ها، بر اهمیت آن افزوده شد. اکثر الگوریتم‌های ارائه شده توسط محققان مختلف، بر مبنای حد آستانه عمل می‌کنند و کنترل خاصی بر میزان حجم ساده‌سازی هندسی ندارند. بدین ترتیب در این تحقیق، رویکردی ارائه شد که توسط آن می‌توان بدون آزمون و خطا، درصد یا تعداد کاهش نقاط اولیه عارضه خطی را تعیین نمود. این رویکرد به عنوان نمونه بر روی دو الگوریتم داگلاس-پوکر (DP) و ویسوالینگام-وایعت (VW) با هدف انتخاب برترین نقاط به تعداد دلخواه (n) پیاده شد و نتایج به دست آمده از آن توسط سه داده مختلف مورد ارزیابی قرار گرفت که نشان داد، هر یک از الگوریتم‌های پیشنهادی می‌تواند ۸۰ درصد از نقاط انتخابی توسط کارشناس نقشه بردار در برداشت عرصه یک زمین کشاورزی را با دقت ۹۸/۶۸ درصدی در محاسبه مساحت و خطای ۹۵ سانتی‌متری در تعیین موقعیت مطلق، به درستی انتخاب کند. نتایج بررسی کیفیت ساده‌سازی توسط الگوریتم‌های پیشنهادی مبتنی بر DP و VW، بر اساس سه شاخص مختلف، نشان می‌دهد که متوسط خطا توسط پلیگون‌های اسلایور، درصد شباهت انحنای متوسط و درصد شباهت تندی زوایا برای ۱۰، ۳۰، ۵۰، ۷۰ و ۹۰ درصد ساده‌سازی بر روی چهارده خط سیر جمع‌آوری شده، برای الگوریتم مبتنی بر DP به ترتیب برابر ۰/۵۰، ۷۲/۳۰ و ۷۱/۸۵ و برای الگوریتم مبتنی بر VW به ترتیب برابر ۰/۴۵، ۷۰/۳۴ و ۶۹/۸۴ بوده است. همچنین نشان داده شده، که حتی با بهینه‌سازی مقدار حد آستانه الگوریتم‌های DP و VW، توسط الگوریتم بهینه‌سازی ژنتیک، خطای نسبی آن‌ها در انتخاب برترین نقاط به تعداد مشخص، به صورت میانگین برای ۱۴ خط سیر جمع‌آوری شده، به ترتیب ۳۸/۲۱۲ و ۴۰/۲۲۸ درصد بیشتر از الگوریتم‌های پیشنهادی تحقیق در مدت زمان پردازشی برابر خواهد بود.

کلیدواژه‌ها: ساده‌سازی هندسی، عارضه خطی، خط سیر.

* نویسنده مکاتبه‌کننده: تبریز، بلوار ۲۹ بهمن، دانشگاه تبریز، دانشکده برنامه‌ریزی و علوم محیطی.

تلفن: ۰۹۱۴۴۰۱۷۴۹۰

۱- مقدمه

یکی از رویکردهایی که به منظور کاهش حجم داده‌های مکان مرجع استفاده می‌شود، ساده‌سازی هندسی است که روی عوارض خطی شامل عوارض چند خطی و چند ضلعی اعمال می‌گردد و منظور از آن، کاهش تعداد نقاط مورد نیاز برای ذخیره‌سازی عوارض خطی است [۱]. به منظور ساده‌سازی هندسی، روش‌های گوناگونی توسط محققان پیشنهاد شده است که با هدف تعمیم نقشه و نمایش عوارض در مقیاس‌های کوچکتر از مقیاس تهیه نقشه، کار خود را شروع کردند [۲].

با افزایش حجم داده‌های به دست آمده از منابع مختلفی از جمله نرم افزارهای مکان‌محور قابل نصب بر روی تلفن‌های همراه هوشمند، اهدافی از قبیل کاهش منابع مورد نیاز جهت ذخیره‌سازی و همچنین سهولت در انتقال و پردازش داده‌ها، بیشتر مورد توجه محققان قرار گرفت و همچنان پژوهشگران در حال ارائه الگوریتم‌های جدیدتر و پیشرفته‌تری با اهداف مختلف در این زمینه هستند [۳]. هریک از الگوریتم‌های ساده‌سازی هندسی، بسته به هدفی که برای آن طراحی شده‌اند، از منطق خاصی پیروی می‌کنند و خصوصیات هندسی گوناگونی را حین ساده‌سازی در نظر می‌گیرند که اکثر آن‌ها با در نظر گرفتن یک یا چند حد‌آستانه برای خصوصیت یا خصوصیات هندسی مورد نظر، به امر ساده‌سازی می‌پردازند [۴] و معمولاً از نظر میزان حجم ساده‌سازی، کنترل آن‌ها بر مقدار کاهش حجم داده‌ها در فرآیند ساده‌سازی نسبتاً کم است؛ در حالی که این امر می‌تواند بسیار مهم باشد. بنابراین تحقیق حاضر به دنبال رویکردی کلی برای در اختیار داشتن کنترل بیشتر بر میزان حجم ساده‌سازی می‌باشد. بدین منظور، به عنوان نمونه، دو الگوریتم شناخته شده و پرکاربرد داگلاس-پوکر (DP) و ویسوالینگام-وایعت (VW)، به نحوی متفاوت و متناسب با هدف مورد نظر، پیاده‌سازی خواهند شد؛ چرا که این الگوریتم‌ها همواره مورد توجه محققان است و پژوهشگران زیادی به دنبال بهبود آن‌ها هستند و یا از این الگوریتم‌ها برای مقایسه رویکردهای

پیشنهادی خود و ارزیابی آن‌ها استفاده می‌کنند [۵، ۶ و ۷].

ساده‌سازی هندسی در ابتدا با هدف تعمیم نقشه معرفی گشت و بعنوان یکی از بخش‌های اصلی این فرآیند به کار گرفته شد [۸]. پس از توسعه فناوری و استفاده از پردازش‌های رایانه‌ای، فرآیند ساده‌سازی هندسی از حالت دستی به خودکار تبدیل شد و اولین الگوریتم‌های ساده‌سازی هندسی خودکار ارائه شدند [۹]. در اولین مطالعات صورت گرفته در این راستا، تیلر و هرشی (۱۹۶۳ و ۱۹۶۵) رویکردهایی را ارائه دادند که بر پایه حد آستانه‌ای وابسته به مقدار پهنای خطوط ترسیم می‌بود. بدین ترتیب هر سه نقطه متوالی از اولین نقطه، شروع به بررسی می‌شدند و نقاطی که فاصله آن‌ها از خط واصل نقاط قبل و بعد از آن، کمتر از حد آستانه می‌بود، حذف می‌گشت [۱۰].

از دیگر رویکردهای نخستین ساده‌سازی هندسی عوارض خطی می‌توان n -امین نقطه را نام برد (۱۹۶۶ میلادی) که بر اساس آن در یک توالی شامل مختصات نقاط عوارض خطی، پس از تعیین مقدار n و انتخاب اولین نقطه، n -امین نقطه بعدی از رشته مختصات‌ها انتخاب می‌شد و سپس n -امین نقطه بعدی بعد از آن و به همین ترتیب تا آخرین نقطه، روند انتخاب نقاط تکرار می‌شد و خروجی ساده‌سازی شکل می‌گرفت [۱۱]. در این رویکرد کنترل حجم ساده‌سازی در دست کاربر است و مقدار حدودی حجم نهایی یک عارضه، پس از ساده‌سازی برابر با $1/n$ حجم اولیه است؛ اما این نکته حائز اهمیت است که این الگوریتم، منطق بسیار ساده‌ای دارد و صرفاً با در نظر گرفتن ترتیب نقاط به امر ساده‌سازی می‌پردازد.

در روندی هوشمندانه‌تر، لنگ (۱۹۶۹) در هر مرتبه از تصمیم‌گیری برای نگه داشتن نقاط، بیشتر از سه نقطه را نیز دخالت داد که بر اساس آن، پس از در نظر گرفتن یک حد آستانه از نوع فاصله، ابتدا فاصله اقلیدسی نقطه دوم از خط فرضی واصل نقطه اول به سوم محاسبه می‌شود و در صورت کمتر بودن مقدار آن

هر تکرار حذف کردند [۱۵].

الگوریتم‌های دیگری نیز ارائه شده‌اند که با بیش از یک حد آستانه کار می‌کنند؛ بعنوان نمونه جنکس (۱۹۸۳) الگوریتمی مبتنی بر سه حد آستانه معرفی نمود که در توالی سه تایی از نقاط، فاصله نقطه اول از نقطه دوم، فاصله نقطه اول از سوم و زاویه میان بردار واصل دو خط متوالی واصل آن‌ها را بررسی می‌کرد [۱۶].

با گذشت زمان و افزایش تنوع در کاربرد ساده‌سازی هندسی، رویکردهایی با منطق‌های پیچیده‌تر توسط محققان معرفی شدند. شن و همکاران (۲۰۱۸) از تکنیک‌های مورد استفاده در پردازش تصویر برای ساده‌سازی الهام گرفتند [۱۷]. سامسونو و یاکیموا (۲۰۲۰) به منظور بهینه نمودن فرایند ساده‌سازی هندسی از رگرسیون استفاده نمودند [۱۳]. ژائو و همکاران (۲۰۲۰) از تبدیل موجک برای کاهش جزئیات عوارض خطی بهره بردند [۱۸]. دو و همکاران (۲۰۲۲) خصوصیات هندسی موثر در ساده‌سازی را به عنوان ورودی در شبکه عصبی مصنوعی استفاده نمودند و از این طریق سعی کردند کیفیت ساده‌سازی را افزایش دهند [۱۹].

با وجود اینکه کاهش حجم مورد نیاز برای ذخیره‌سازی عوارض خطی یکی از اصلی‌ترین اهداف ساده‌سازی هندسی است و علی‌رغم تنوع بسیار زیاد در رویکردهای ارائه شده، در اکثر آن‌ها توانایی تعیین نسبی میزان کاهش حجم ذخیره‌سازی مورد نیاز برای عارضه خطی مورد نظر مقدور نیست و معمولاً بر اساس حد آستانه‌ای هندسی، ساده‌سازی انجام می‌پذیرد. به عنوان نمونه ممکن است در یک پایگاه داده مکانی شامل عوارض خطی، کاربر بخواهد حجم ذخیره‌سازی عوارض را ۵۰٪ کاهش دهد. در این صورت تعیین اینکه وارد کردن چه مقداری بعنوان حد آستانه می‌تواند منجر به کاهش ۵۰ درصدی حجم ذخیره‌سازی شود، خود یک مسئله بهینه‌سازی است، که حتی ممکن است نیاز باشد برای هر یک از عوارض موجود در پایگاه داده مذکور، مجزا محاسبه شود که به طبع، از نظر زمانی و

از حد آستانه، فاصله نقطه دوم و سوم از خط واصل نقاط اول و چهارم بررسی می‌گردد، اما در صورتی که حداقل یکی از نقاط در فاصله‌ای دورتر از خط واصل نقاط اول و آخر در حال بررسی باشد، آنگاه نقطه ماقبل آخری که در حال بررسی است نگه داشته می‌شود و مابقی نقاط حذف می‌شوند و این روند تا رسیدن به نقطه آخر ادامه می‌یابد [۵].

برخلاف الگوریتم‌هایی که در هر مرحله از کار خود تعداد محدودی از نقاط متوالی را به صورت محلی بررسی می‌کنند، داگلاس و پوکر (۱۹۷۳) الگوریتمی را مطرح نمودند که به طور سراسری عمل می‌کند. بدین صورت که از همان ابتدا فاصله تمامی نقاط مابین نقاط اول و آخر، از خط فرضی واصل آن دو محاسبه می‌شود و در حالتی که بیشترین فاصله، از حد آستانه از پیش تعیین شده بیشتر باشد، عارضه خطی از نقطه‌ای که بیشترین فاصله را دارا بوده است، به دو قسمت تقسیم می‌شود و برای هریک از این قسمت‌ها فرآیند مذکور تکرار می‌گردد و در غیر این صورت ساده‌سازی به اتمام می‌رسد [۱۲]. سپس جنکس (۱۹۸۱) یک الگوریتم مبتنی بر توالی هر سه نقطه متوالی معرفی نمود که در هر سه نقطه، فاصله نقطه میانی از خط واصل دو نقطه دیگر را با حد آستانه مقایسه می‌کرد [۱۳].

با این حال، فاصله تنها خصوصیت هندسی نبود که برای استفاده در ساده‌سازی معرفی شد. مکمستر (۱۹۸۳) به جای حد آستانه فاصله از حد آستانه زاویه استفاده نمود و در توالی سه تایی از نقاط، زاویه میان بردار واصل نقطه یک به سه و بردار واصل نقطه یک به دو را با مقدار حد آستانه مقایسه می‌کرد و نقطه میانی را در صورت کمتر بودن زاویه اندازه گیری شده از حد آستانه حذف می‌نمود [۱۴].

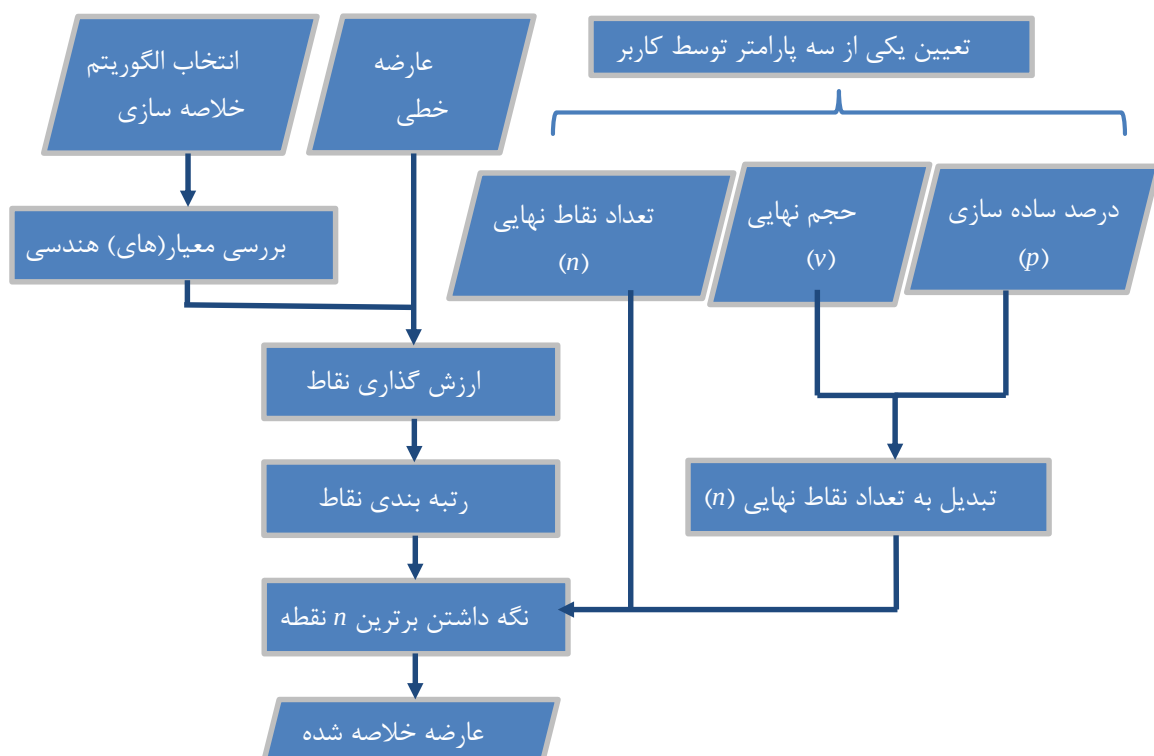
از دیگر خصوصیات هندسی که در امر ساده‌سازی مورد توجه محققان قرار گرفته است، مساحت می‌باشد که بعنوان نمونه، ویسوالینگام و وایعت (۱۹۹۳) با در نظر گرفتن مساحت مثلث‌های قابل تشکیل از تمامی سه نقطه متوالی، نقطه میانی کم مساحت‌ترین مثلث را در

اعمال گردد. پس از تعیین الگوریتم، روند کار و معیار یا معیارهای هندسی مورد استفاده در آن برای خلاصه‌سازی، مورد بررسی قرار می‌گیرد. سپس توسط این معیارها، هریک از نقاط عارضه خطی، به صورت نسبی ارزش‌گذاری می‌شوند و پس از رتبه‌بندی نقاط بر این اساس، برترین آن‌ها به تعداد دلخواه، که با سه روش متفاوت می‌تواند توسط کاربر تعیین گردد، نگه داشته می‌شوند و بقیه نقاط حذف می‌گردند. روند کلی مذکور، به طور خلاصه در شکل (۱) نمایش داده شده است و در ادامه، جزئیات بیشتر تشریح خواهند گشت.

محاسباتی بسیار هزینه‌بر می‌باشد. بدین ترتیب تحقیق حاضر به دنبال ارائه رویکردی کلی به منظور افزایش کنترل کاربر بر میزان کاهش حجم ذخیره‌سازی عوارض خطی، در ساده‌سازی هندسی است.

۲- مواد و روش‌ها

با توجه به هدف کلی این تحقیق که ارائه رویکردی کلی به منظور پیاده‌سازی متفاوت الگوریتم‌های ساده‌سازی هندسی عوارض خطی به منظور افزایش کنترل بر میزان کاهش حجم ذخیره‌سازی اطلاعات هندسی است، به عنوان نمونه از دو الگوریتم DP و VW استفاده خواهد شد تا تغییراتی در روند پردازشی آن‌ها



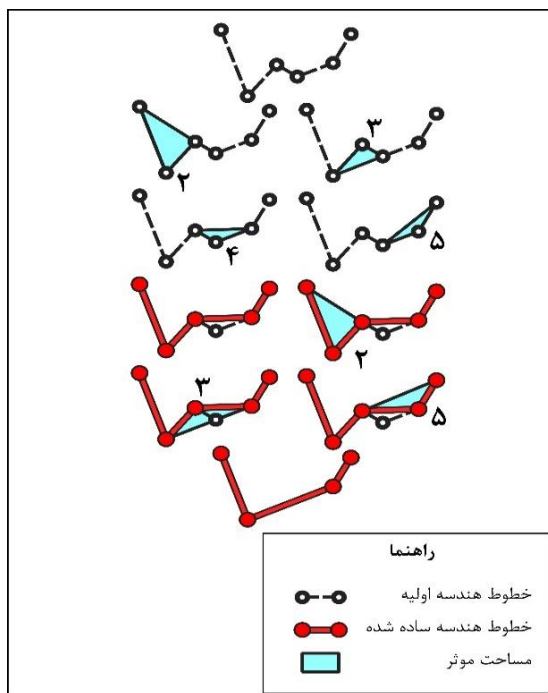
شکل ۱: روند کلی پیشنهادی بهینه‌سازی الگوریتم‌های خلاصه‌سازی هندسی با هدف کنترل بر حجم ساده‌سازی

بازگشتی عارضه تقسیم می‌شود. همان‌طور که در قسمت راست شکل (۲) نشان داده شده است، در مرحله نخست، اولین نقطه به آخرین نقطه وصل می‌شود (a) و فاصله نقاط میانی از خط واصل محاسبه

۲-۱- روند کار الگوریتم DP

مطابق با این روش، در ابتدا مجموعه‌ای مرتب از نقاط یا خطوط عارضه در نظر گرفته می‌شود و سپس بصورت

مؤثر محاسبه شده، از حد آستانه از پیش تعیین شده کمتر باشد، آنگاه نقطه مربوط به آن، کنار گذاشته می‌شود. این عملیات تا جایی ادامه می‌یابد که هیچ نقطه‌ای دارای مساحت مؤثر بزرگتر از حد آستانه از پیش تعیین شده نباشد. به عنوان یک مثال، مراحل کار VW در قسمت چپ شکل (۲) نمایش داده شده است. در مرحله اول، مساحت مؤثر برای نقاط یک تا پنج محاسبه شده و نقطه چهار که دارای کمترین مساحت مؤثر است و کمتر از حد آستانه از پیش تعیین شده می‌باشد کنار گذاشته شده است و سپس با تکرار این مراحل در غیاب نقطه چهارم، اینبار نقطه سه کنار گذاشته می‌شود و در نهایت هندسه ساده شده به دست می‌آید [۲۰].



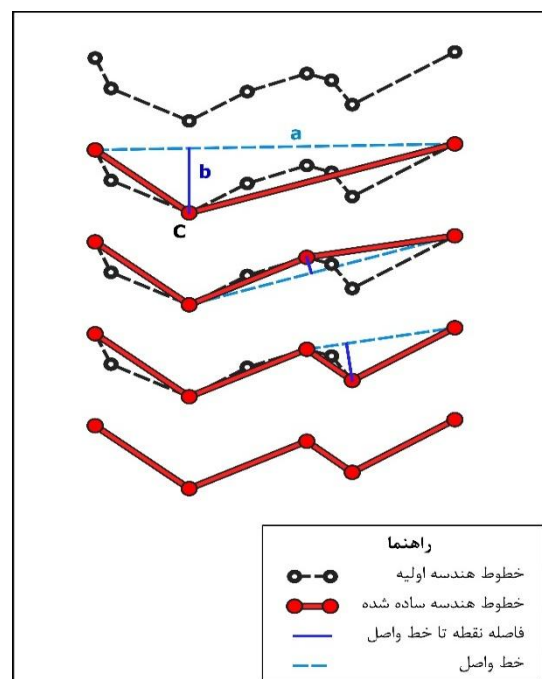
شکل ۲: روند کار الگوریتم DP (تصویر راست) و VW (تصویر چپ)

نهایی مورد نظر (n)، انتخاب درصد کاهش حجم ذخیره‌سازی (p) و انتخاب حجم نهایی پس از ساده‌سازی هندسی (v). در هر یک از روش‌های فوق، تعداد نهایی نقاط (n) یا به طور مستقیم توسط کاربر وارد می‌شود و یا به طور تقریبی قابل محاسبه می‌باشد.

می‌گردد. حال به ترتیب نقاط میانی بررسی می‌شوند و اگر فاصله نقطه‌ای که بیشترین فاصله را از خط دارد (فاصله b)، از حد آستانه مشخص شده بیشتر باشد، به عنوان نقطه لنگر (نقطه c) انتخاب می‌شود. سپس همین عملیات برای نقاط قبل و بعد از آن، به صورت مجزا انجام خواهد شد و در غیر اینصورت عملیات ساده‌سازی برای قسمت مورد نظر به اتمام خواهد رسید [۱۲].

۲-۲- روند کار الگوریتم VW

در روش VW ، ابتدا مساحت میان تمامی سه نقطه‌های پشت سر هم موجود در عارضه خطی محاسبه می‌شود و این مقادیر به عنوان مساحت مؤثر، به نقطه میانی هر مثلث نسبت داده می‌شود. سپس اگر کمترین مساحت



۲-۳- پارامترهای تعیین‌کننده حجم ساده‌سازی

کنترل بر میزان حجم ساده‌سازی هندسی، می‌تواند به چند روش مختلف توسط کاربر تعیین گردد که برخی از آن‌ها عبارت است از: انتخاب مستقیم تعداد نقاط

انتخاب مستقیم n برای خلاصه‌سازی هر عارضه به صورت جداگانه انجام می‌گیرد و در تحلیل جداگانه هر عارضه کاربرد بیشتری دارد.

اگر تعداد اولیه نقاط یک عارضه خطی برابر با N_0 و حجم اولیه مورد نیاز برای ذخیره‌سازی آن V_0 باشد، آنگاه مقدار n از طریق روابط (۱) و (۲) می‌تواند به طور تقریبی محاسبه گردد.

$$n = [N_0 * v/V_0] \quad \text{رابطه (۱)}$$

در رابطه (۱) از تقسیم حجم نهایی دلخواه مورد نیاز برای ذخیره‌سازی داده‌ها پس از ساده‌سازی (v)، بر حجم اولیه مورد نیاز برای ذخیره‌سازی (V_0)، یک نسبت به دست می‌آید که از ضرب آن در تعداد نقاط اولیه (N_0)، می‌توان تعداد نهایی نقاط را محاسبه نمود. از آنجا که تعداد نقاط باید به صورت عدد طبیعی باشد، از جزء صحیح عدد به دست آمده استفاده می‌گردد.

از رابطه (۱) به عنوان نمونه زمانی استفاده می‌گردد که کاربر نیاز به پاکسازی حجم معینی از داده‌های ذخیره شده برای تشکیل عارضه خطی را دارد.

$$n = [(100 - p) * N_0/100] \quad \text{رابطه (۲)}$$

در رابطه (۲) نیز به دلیل اینکه درصد کاهش حجم مورد نیاز برای ذخیره‌سازی نقاط عارضه خطی (p) به عنوان ورودی تعیین شده است، ابتدا p از ۱۰۰ کسر می‌گردد تا به درصد نقاط مورد نیاز برای نگهداری تبدیل شود. سپس با تقسیم بر صد، مقدار آن مابین صفر و یک قرار می‌گیرد و با ضرب مقدار به دست آمده در تعداد نقاط اولیه (N_0)، تعداد نقاط نهایی مورد نظر حاصل می‌گردد.

برخلاف تعیین مستقیم پارامتر n و یا تعیین غیر مستقیم آن از طریق v ، که به طور جداگانه برای هر عارضه تعیین می‌گردد، از طریق رابطه (۲) می‌توان درصد خلاصه‌سازی کل عوارض خطی یک پایگاه داده را نیز تعیین نمود. زمانی که کاربر قصد دارد کل حجم ذخیره‌سازی مورد نیاز برای یک پایگاه داده را به نسبت دلخواه کاهش دهد، این روش می‌تواند کارآمد باشد.

بنابراین در حقیقت قرار است که n نقطه از نقاط اولیه انتخاب شوند و طبیعتاً این کار زمانی ارزشمند است که این n نقطه، برترین نقاط ممکن برای تشکیل نسخه ساده شده عارضه خطی باشند. اینکه میزان برتری نسبی نقاط تشکیل‌دهنده عارضه خطی به چه صورت مشخص شود، می‌تواند برگرفته از منطق الگوریتم‌های موجود باشد. به عنوان نمونه می‌تواند بر اساس روش DP ، انتخاب برترین نقاط بر مبنای فاصله و یا بر اساس روش VW بر مبنای مساحت باشد. بدین ترتیب رویکرد کلی پیشنهادی این است که بر اساس معیار هندسی مورد استفاده در الگوریتم مورد نظر، تمامی نقاط یا بخش مورد نیاز از آن‌ها رتبه‌بندی شوند و سپس تعداد مورد نظری از آن‌ها (n) که بر مبنای شاخص هندسی مدنظر، برتر به حساب می‌آیند انتخاب شده و به عنوان خروجی برگردانده شوند.

۲-۴- انتخاب برترین n نقطه بر مبنای الگوریتم DP
همان‌طور که از پیش نیز گفته شد، در روش DP ، ابتدا نقطه اول به آخر متصل می‌شود و سپس فاصله تمامی نقاط مابین آن دو، تا خط تشکیل داده شده، توسط رابطه (۳) محاسبه می‌گردد [۲۱].

$$D = \frac{|mx_0 - y_0 + b|}{\sqrt{m^2 + 1}} \quad \text{رابطه (۳)}$$

با توجه به رابطه (۳) که معادله خط را نشان می‌دهد، ضریب y برابر با منفی یک می‌باشد و مقادیر m و b همان شیب خط و عرض از مبدا آن هستند و x_0 و y_0 نیز مؤلفه مختصات طول و عرض نقطه مورد نظر می‌باشند.

ابتدا و انتهای یک خط، شامل دو نقطه بسیار مهم است که حدود عوارض خطی را در طول آن‌ها مشخص می‌کند و به همین دلیل، پس از ساده‌سازی نیز، این دو نقطه باید باقی بمانند و در واقع می‌توان گفت مهم‌ترین نقاط هستند و اگر انتخاب برترین دو نقطه ($n=2$) مورد نظر باشد، آن دو نقطه همان نقاط ابتدا و انتهای عارضه خطی هستند. در غیر این صورت، همانگونه که در قسمت سمت راست شکل (۳) نشان داده شده است،

منتخب باقی بمانند. به عبارت دیگر، معیار هندسی مورد استفاده در منطق بکار گرفته شده برای طراحی این الگوریتم، مساحت مؤثر می‌باشد، که می‌تواند به عنوان اهمیت نسبی، به هر نقطه نسبت داده شود و بدین ترتیب، با منطقی که این الگوریتم از آن استفاده می‌کند، می‌توان گفت در هر مرحله، کم اهمیت‌ترین نقطه کنار گذاشته می‌شود. بنابراین با حذف نقاط، هر بار گزینه‌های انتخابی برای بهترین n نقطه کمتر می‌شوند تا اینکه جواب نهایی حاصل گردد و به همین دلیل می‌توان آنرا بازگشتی نامید.

حال اگر معیار توقف روند بازگشتی حذف نقاط تغییر یابد، می‌توان بهترین n نقطه را مطابق با منطق این الگوریتم شناسایی نمود. بدین صورت که روند حذف نقاط تا جایی ادامه یابد که تعداد نقاط باقی مانده، برابر با n شود و دیگر شرط بزرگتر شدن کوچک‌ترین مساحت مؤثر از حدآستانه از پیش تعیین شده، دلیل توقف تکرار مراحل برای حذف نقاط نباشد.

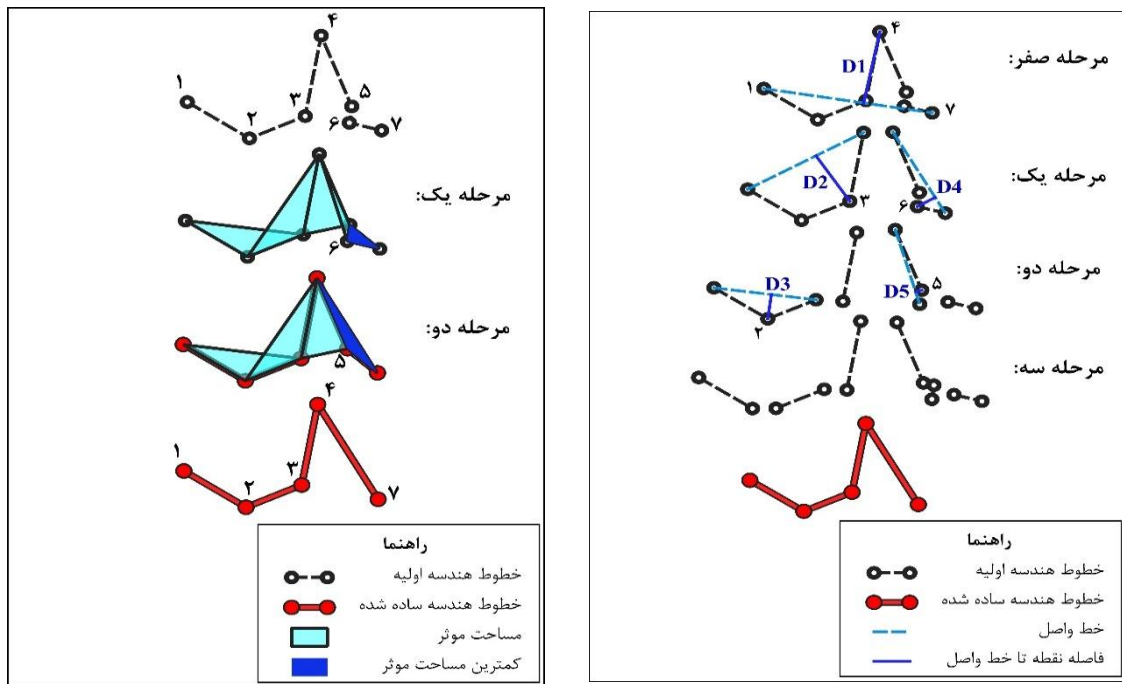
همان‌طور که در قسمت سمت چپ شکل (۳) مشاهده می‌شود، هر سه نقطه پشت سر هم تشکیل یک مثلث می‌دهند که مساحت آن، بعنوان مساحت مؤثر، به نقطه وسط (نقطه دوم در هر دنباله سه تایی از نقاط) نسبت داده می‌شود. در مرحله اول، نقطه ۶ که دارای کم‌ترین مساحت مؤثر می‌باشد کنار گذاشته می‌شود و در مرحله دوم نیز نقطه ۵ با همین منطق حذف می‌گردد. نهایتاً بهترین ۵ نقطه به عنوان جواب الگوریتم باقی خواهند ماند که اگر رتبه بندی آن‌ها نیز مدنظر باشد، می‌توان بر اساس مساحت مؤثر هر نقطه، این کار را انجام داد. میزان کاهش درصدی نقاط در این روش را نیز می‌توان توسط رابطه (۲) و مشابه با روش مبتنی بر DP تعیین نمود.

این دو نقطه توسط یک خط فرضی به یکدیگر متصل می‌شوند (خط چین آبی) و فاصله تمامی نقاط مابین آن‌ها تا خط فرضی، محاسبه خواهد شد. در این شرایط، نقطه‌ای که دارای بیشترین فاصله از خط فرضی باشد (نقطه چهارم)، به عنوان اولین نقطه لنگر انتخاب خواهد شد و مقدار فاصله آن از خط فرضی (فاصله D_1)، ذخیره می‌شود. در مرحله بعدی (مرحله یک)، دو خط فرضی به نحوی تشکیل می‌شوند که یکی از آن‌ها خط واصل میان نقطه اول و نقطه لنگر انتخاب شده در مرحله قبل و دیگری نیز خط واصل میان نقطه لنگر و نقطه آخر می‌باشد. مجدداً مقدار فواصل نقاط میانی، نسبت به خط واصل محاسبه خواهند شد و پس از انتخاب نقاط لنگر بر اساس بیشترین فاصله از خطوط فرضی واصل (فواصل D_2 و D_4)، مقدار این فواصل ذخیره خواهند شد و به نقاط مربوطه (نقاط سه و شش) نسبت داده خواهند شد.

این مراحل تا جایی ادامه خواهد یافت که تعداد نقاط هریک از خطوطی که در هر مرحله، در نقطه لنگر از هم جدا می‌شوند، به دو برسد و به هریک از نقاط، یک فاصله نسبت داده شده باشد. سپس نقاط بر اساس ارزش نسبی که به آن‌ها اطلاق شده است، از بزرگ به کوچک مرتب شوند و n نقطه اول، به عنوان مهم‌ترین یا به عبارتی برترین نقاط انتخاب گردند. در مثال مربوط به قسمت سمت راست شکل (۳)، نهایتاً به ترتیب نقاط ۱، ۷، ۴، ۳، ۲ و ۶ دارای اهمیت نسبی زیاد به کم خواهند بود و به عنوان مثال اگر انتخاب بهترین ۵ نقطه مدنظر باشد، آنگاه نقاط ۱، ۷، ۴، ۳ و ۲ انتخاب خواهند شد.

۲-۵- انتخاب برترین n نقطه بر مبنای الگوریتم VW

نکته قابل توجهی که به پیاده‌سازی الگوریتم VW با هدف کنترل میزان ساده‌سازی کمک می‌کند، این است که این رویکرد به صورت بازگشتی عمل می‌کند؛ بدین صورت که هربار نقطه‌ای که دارای کمترین مساحت مؤثر است، کنار گذاشته می‌شود تا اینکه بالاخره نقاط



شکل ۳: پیاده‌سازی متفاوت الگوریتم DP (تصویر راست) و VW (تصویر چپ)

فرکانسه‌های تارگت مدل ای-تی-۷۰ و همچنین گیرنده GNSS تلفن همراه هوشمند شیائومی مدل ردمی نوت ۱۱ استفاده گردید که پس از بررسی عملکرد آن‌ها در محدوده مطالعاتی، خطای مسطحاتی آن‌ها در ثبت هر نقطه به ترتیب حدود ۲ سانتی‌متر و ۳۰۰ سانتی‌متر برآورد شد.

۲-۶- جمع آوری داده

به منظور ارزیابی عملکرد الگوریتم‌های پیشنهادی در انتخاب برترین n نقطه از میان نقاط اولیه عارضه خطی، مطابق با آنچه در بخش روش‌های ارزیابی گفته خواهد شد، سه نوع بررسی مختلف انجام می‌شود. بدین منظور دو نوع داده مختلف جمع آوری می‌شود که یکی مربوط به برداشت مختصاتی عرصه یک زمین زراعی در حومه شهر تبریز توسط دو روش متفاوت (گیرنده ماهواره‌ای چندفرکانسه و نرم افزار ویکی‌لوک^۱ و دیگری خط سیرهای ثبت شده در دامنه کوه عینالی تبریز توسط نرم افزار ویکی‌لوک می‌باشد که در شکل (۴) قابل رؤیت است.

لازم به ذکر است که در جمع آوری داده‌های مکانی مربوط به عرصه زمین زراعی، از گیرنده ماهواره‌ای سامانه ناوبری ماهواره‌ای جهانی^۲ (GNSS) چند

^۳ Hi-Target

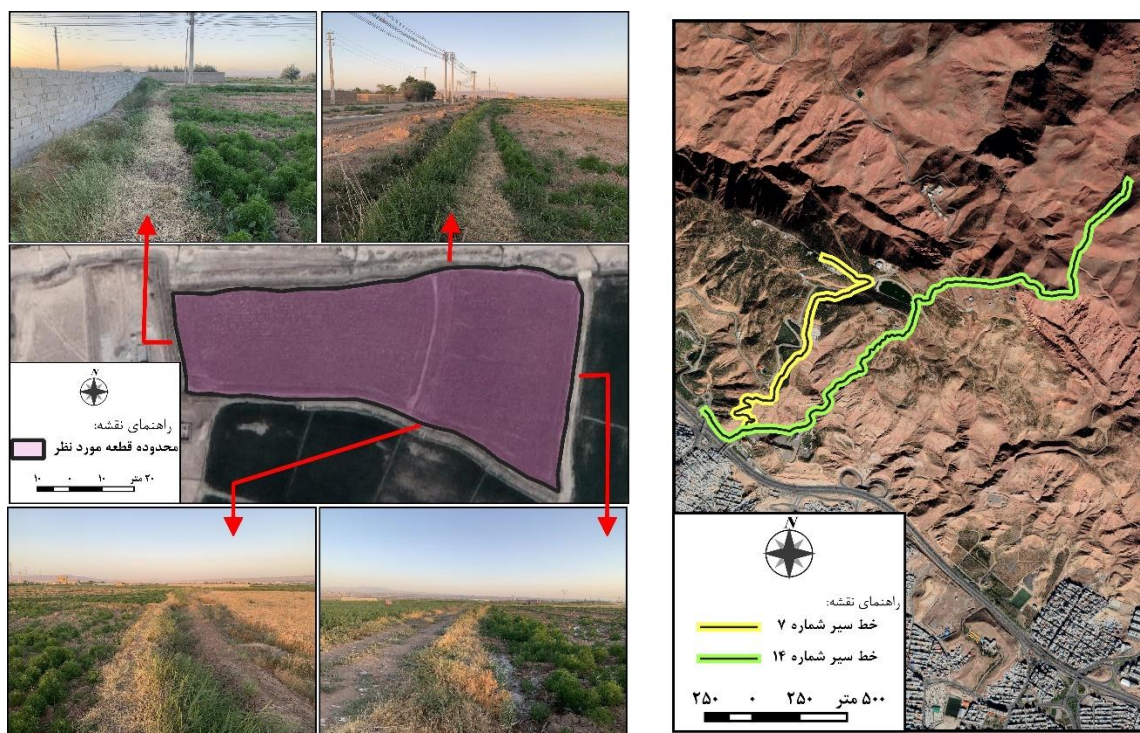
^۴ AT70

^۵ Xiaomi

^۶ Redmi Note 11

^۱ Wikiloc

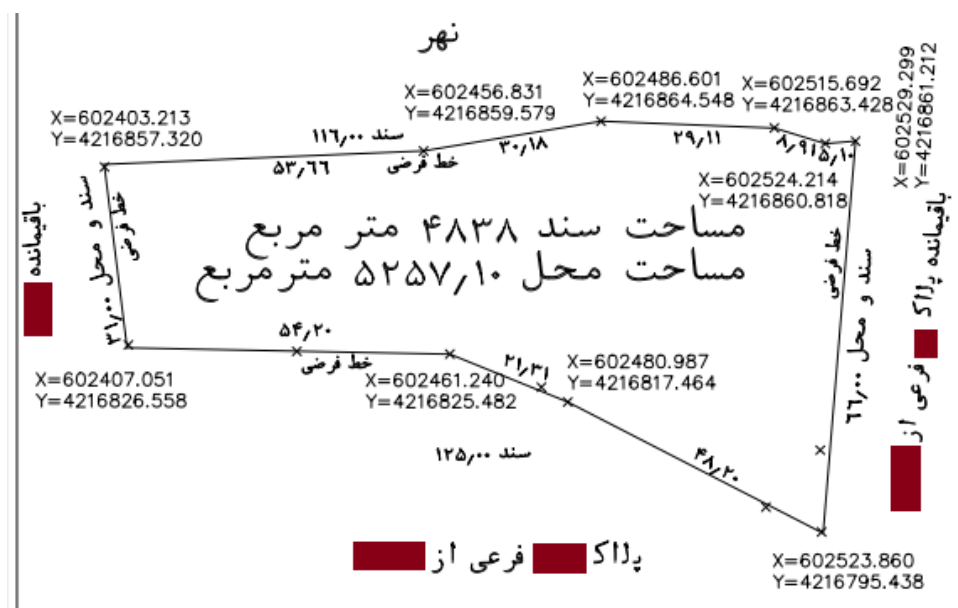
^۲ Global Navigation Satellite System



شکل ۴: دو خط سیر نمونه از دامنه کوه عینالی تبریز (تصویر راست) و عرصه یک ملک زراعی در حاشیه تبریز (تصویر چپ)

چندفرکانسه، تعداد ۱۰ نقطه (که مختصات آن‌ها نوشته شده است) به عنوان نقاط شکستگی حدود عرصه ملک مورد نظر، توسط کارشناس انتخاب شده است.

شکل (۵) نیز، نقشه ترسیمی کارشناس نقشه بردار را نشان می‌دهد که از میان ۱۴ نقطه (با ضربدر مشخص شده است) برداشت شده بوسیله گیرنده ماهواره‌ای



شکل ۵: بخشی از نقشه ثبتی ترسیم شده توسط کارشناس نقشه بردار سازمان ثبت اسناد و املاک کشور

۷-۲- روش‌های ارزیابی

ارزیابی الگوریتم‌های پیشنهادی تحقیق حاضر به سه روش مختلف انجام خواهد شد که در ادامه شرح داده شده است.

۲-۷-۱- بررسی عملکرد الگوریتم‌های پیشنهادی در انتخاب برترین نقاط نسبت به کارشناس نقشه بردار

سؤالی که مطرح می‌شود این است که آیا الگوریتم‌های پیشنهادی تحقیق حاضر، می‌توانند این ۱۰ نقطه انتخاب شده توسط کارشناس نقشه بردار را از میان ۱۴ نقطه موجود تشخیص دهند یا خیر. بدین منظور تمامی ۱۴ نقطه به عنوان ورودی به الگوریتم‌های مبتنی بر DP و VW داده شدند و ۱۰ نقطه برتر توسط آن‌ها مشخص شد تا معلوم شود که چه تعداد از نقاط انتخابی آن‌ها برابر با نقاط انتخابی توسط کارشناس نقشه بردار به عنوان نقاط شکست است.

همچنین با توجه به اینکه تعداد نقاطی که از میان آن‌ها باید ۱۰ نقطه انتخاب شود نسبتاً کم است (۱۴ نقطه)، آزمون مذکور به روش دیگری نیز انجام می‌گیرد. بدین صورت که ابتدا با فعال نمودن ثبت خط سیر در نرم افزار ویکی‌لوک و حرکت در محدوده زمین مورد مطالعه، یک خط سیر غیردقیق ثبت شود که برای ثبت آن از ریزتراشه‌ی $GNSS$ تلفن همراه هوشمند استفاده شده است و سپس الگوریتم‌های پیشنهادی، از میان تعداد نقاط متوالی نسبتاً زیادی که جهت ثبت خط سیر برداشت شده‌اند، تعداد ۱۰ نقطه را انتخاب کنند و در نهایت بررسی شود که چه تعدادی از این نقاط، نسبت به نقاط انتخاب شده توسط کارشناس نقشه بردار، که با دقت زیاد برداشت شده‌اند، نزدیک‌ترین هستند و خطای هر یک محاسبه گردد.

منظور از خطا، در واقع اختلاف میان داده‌های به دست آمده از روش برداشت دقیق (استفاده از گیرنده ماهواره‌ای چند فرکانسه) نسبت به روش برداشت غیر دقیق (استفاده از تلفن همراه هوشمند) است که این اختلاف، همان فاصله میان آن‌ها می‌باشد که در این

پژوهش به دو صورت مختلف اندازه‌گیری می‌شود.

۲-۷-۱-۱- روش اول: ریشه میانگین مربعات خطا ($RMSE$)

$$RMSE = \frac{\sum_{i=1}^n D_i^2}{n} \quad \text{رابطه (۴)}$$

مقدار خطا توسط این روش بوسیله رابطه (۴) محاسبه می‌گردد که در آن n برابر با تعداد نقاط انتخابی و D_i برابر با فاصله نقطه هدف i ام (یکی از نقاطی که توسط کارشناس نقشه بردار به عنوان نقاط شکست انتخاب شده‌اند) تا نزدیک‌ترین نقطه انتخاب شده توسط الگوریتم‌های پیشنهادی این پژوهش است. در واقع این معیار، مقدار خطا در انتخاب نقاط درست را برآورد می‌کند [۲۲].

۲-۷-۱-۲- روش دوم: متوسط خطا توسط پلیگون اسلایور (ME)

با توجه به اینکه تنها عنصر تشکیل دهنده عوارضی خطی، نقطه نمی‌باشد و طول خطوط یا همان یال‌ها نیز در تشکیل یک عارضه خطی سهیم هستند، در روش دوم محاسبه خطا سعی شده است که اختلاف میان محدوده تعیین شده توسط کارشناس نقشه بردار (عارضه اصلی) و محدوده به دست آمده از ترکیب برداشت با تلفن همراه هوشمند و ساده‌سازی توسط رویکردهای پیشنهادی (خط سیر ساده شده)، در تمام طول عارضه خطی محاسبه گردد. بدین منظور مساحت پلیگون اسلایور [۲۳] میان دو عارضه مذکور، محاسبه می‌گردد و نهایتاً به منظور نرمال سازی، عدد نهایی بر طول خط عارضه اصلی تقسیم می‌گردد. مقدار متوسط خطای میان عارضه اصلی و خطوط ساده شده، از طریق رابطه (۵) محاسبه می‌گردد، که در آن L برابر است با طول کل عارضه اصلی و S_i ‌ها نیز مساحت پلیگون‌های اسلایور هستند.

¹ Root Mean Squares Error

² Sliver Polygon

۲-۷-۱-۲- معیار درصد شباهت انحنای متوسط (MC) مقدار انحنای متوسط از تقسیم مجموع زاویه‌های هندسه‌ی عارضه، بر مجموع طول تمامی لبه‌های تشکیل دهنده آن به دست می‌آید. درصد شباهت انحنای متوسط، از رابطه (۷) محاسبه می‌شود [۲۶].

$$MC = 100 \times |MC_2 - MC_0| / MC_0 \quad \text{رابطه (۷)}$$

که انحنای متوسط هندسه‌ی اولیه (MC_0) و ساده شده (MC_2)، هر یک از تقسیم مجموع زوایا بر مجموع طول یال‌هایشان به دست می‌آید.

۲-۷-۲-۲- معیار درصد شباهت تغییرات تنیدی زاویه‌ها^۱ (AC)

همان‌طور که در رابطه (۸) ذکر شده است، این شاخص برابر است با نسبت مجموع زوایای هندسه‌ی ساده شده (Ang_2) به مجموع زوایای هندسه اولیه (Ang_0) ضرب در ۱۰۰ [۲۷].

$$AC = 100 \times Ang_2 / Ang_0 \quad \text{رابطه (۸)}$$

۲-۷-۳- مقایسه الگوریتم‌های پیشنهادی تحقیق، نسبت به روش‌های متداول در انتخاب برترین نقاط

در استفاده از الگوریتم‌های ساده‌سازی هندسی متداول، که پارامتر ورودی آن‌ها اغلب شامل یک یا چند معیار هندسی به عنوان حد آستانه است، اگر هدف کاربر، رسیدن به حجم ساده‌سازی خاصی باشد، ناچار است مقدار پارامتر را توسط آزمون و خطا آنقدر تغییر دهد تا برای رسیدن به حجم ساده‌سازی دلخواه، به مقدار بهینه حد آستانه معیار هندسی برسد. در رویکردی هوشمندانه‌تر، می‌توان به صورت یک مسئله بهینه‌سازی عمل نمود. بدین ترتیب، حد آستانه معیار هندسی مورد استفاده، طوری بهینه می‌گردد که اختلاف میان تعداد نقاط خروجی و تعداد نقاط مورد نظر، حداقل گردد.

$$ME = \frac{\sum_{i=1}^3 S_i}{L} \quad \text{رابطه (۵)}$$

۲-۷-۳-۱- درصد شباهت مساحت (A)

همچنین از آنجا که عرصه یک ملک به صورت چندضلعی می‌باشد و دارای مساحت است، اختلاف مساحت هندسه نهایی ارائه شده توسط کارشناس نقشه بردار و آنچه توسط الگوریتم‌های مبتنی بر DP و VW ساده شده است نیز می‌تواند معیار مناسبی جهت مقایسه کارایی الگوریتم‌های پیشنهادی باشد [۲۴]. بنابراین درصد شباهت مساحت یک عارضه (A_1) نسبت به عارضه دیگر (A_0) را می‌توان توسط رابطه (۶) محاسبه نمود.

$$A = 1 - \frac{|A_0 - A_1|}{A_0} \quad \text{رابطه (۶)}$$

در رابطه (۶)، A_0 مساحت هندسه اولیه و A_1 مساحت هندسه ساده شده است.

۲-۷-۲- بررسی کیفیت ساده‌سازی هندسی الگوریتم‌های پیشنهادی تحقیق

پس از بررسی کارایی الگوریتم‌های پیشنهادی تحقیق حاضر در انتخاب برترین n نقطه، کیفیت عملکرد آن‌ها در ساده‌سازی خط سیرهای جمع‌آوری شده با یکدیگر مقایسه می‌گردد. هر چه شباهت میان عارضه اصلی و عارضه ساده شده بیشتر باشد، کیفیت ساده‌سازی بهتر است [۲۵] و به طور مشابه، کم بودن اختلاف میان یک عارضه و نسخه ساده شده آن نیز می‌تواند تعریفی از با کیفیت بودن ساده‌سازی باشد و در نتیجه معمولاً معیارهایی که استفاده می‌شوند یا از منظری خاص، شباهت را می‌سنجند (مانند معیار درصد شباهت انحنای متوسط و معیار درصد شباهت تغییرات تنیدی زاویه‌ها) و یا اختلاف و فاصله را مورد ارزیابی قرار می‌دهند (مانند ME)، که در ادامه برخی دیگر از آن‌ها مورد بررسی قرار می‌گیرد.

¹ Mean Curvature

² Angular Change

افزار QGIS استفاده گردید که نتایج آن‌ها در ادامه بررسی خواهند شد.

۳- نتایج و بحث

نتایج به دست آمده از ساده‌سازی الگوریتم‌های پیشنهادی این تحقیق در سه فاز مختلف بررسی خواهند شد؛

۳-۱- نتایج بررسی عملکرد الگوریتم‌های پیشنهادی در انتخاب برترین نقاط نسبت به کارشناس

مطابق شکل (۶)، کارشناس نقشه بردار از ۱۴ نقطه‌ای که برداشت کرده بود تعداد ۱۰ نقطه را به عنوان شکستگی نهایی انتخاب کرده بود که تعداد ۸ نقطه آن با آنچه که توسط الگوریتم‌های پیشنهادی مبتنی بر DP و VW مطابقت دارد؛ به عبارت دیگر ۸۰ درصد نقاط انتخابی بر اساس هر دو الگوریتم DP و VW صحیح بوده است و منظور از صحت در این بخش، انطباق با نظر کارشناس می‌باشد.

درصد صحت انتخاب نقاط توسط دو الگوریتم مبتنی بر DP و VW برابر است. اما با توجه به اینکه مجموعه نقاط انتخابی آن‌ها یکسان نبوده است، نتایج مقایسه توانایی آن‌ها در حفظ مساحت هندسه اولیه درج شده در جدول (۱) نیز می‌تواند حائز اهمیت باشد.

همان‌طور که از جدول (۱) پیداست، مساحت هندسه ساده شده توسط کارشناس نقشه بردار، به مساحت هندسه اولیه (برداشت شده توسط گیرنده ماهواره‌ای چندفرکانسه) شباهت بیشتری دارد و این درحالی است که الگوریتم مبتنی بر VW نیز عملکرد نسبتاً مشابهی داشته است و در مقایسه با الگوریتم مبتنی بر DP، بهتر عمل کرده است. بنا بر آنچه که پیش از این گفته شد، الگوریتم VW بر اساس مساحت به ساده‌سازی می‌پردازد و می‌توان نتیجه گرفت که در مواردی مانند ساده‌سازی عرصه املاک که در آن مساحت عرصه بسیار حائز اهمیت است و تأثیر مستقیم بر ارزش ملک دارد، الگوریتم ساده‌سازی مبتنی بر VW می‌تواند کارایی بیشتری داشته باشد.

بدین منظور، با توجه به رابطه (۹)، می‌توان قدرمطلق تفاضل تعداد نقطه مورد نظر (n) و تعداد نقطه حاصل از خلاصه‌سازی (N_{out}) با در نظر گرفتن مقدار T بعنوان حد آستانه معیار هندسی را به عنوان تابع هدف (f) در نظر گرفت و مقدار T را طوری بهینه نمود که مقدار $f(T)$ کمینه گردد.

$$f(T) = |n - N_{out}| \quad \text{رابطه (۹)}$$

برای اینکه نتایج روش مذکور، قابل مقایسه با نتایج الگوریتم‌های پیشنهادی تحقیق باشد، به اندازه‌ی مدت زمانی که الگوریتم‌های پیشنهادی مبتنی بر DP و مبتنی بر VW برای خلاصه‌سازی هندسی به منظور کاهش حجم دلخواه نیاز دارند، به الگوریتم بهینه‌سازی (در این پژوهش از الگوریتم ژنتیک استفاده شده است) زمان داده می‌شود تا مقدار حد آستانه را طوری بهینه کند (T_{opt}) که تابع هدف کمینه شود. آنگاه خطای نسبی الگوریتم بهینه‌سازی ($E_{f(T)}$) در انتخاب برترین نقاط توسط رابطه (۱۰) بررسی می‌شود.

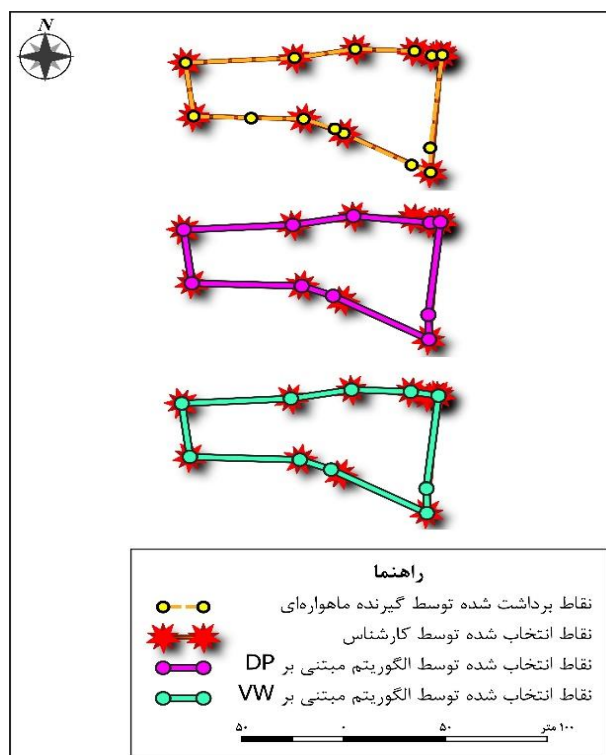
$$E_{f(T)} = f(T_{opt})/n \quad \text{رابطه (۱۰)}$$

برای به دست آوردن میزان خطای نسبی ($E_{f(T)}$)، صرفاً مقدار تابع هدف به ازای حد آستانه بهینه (T_{opt})، بر تعداد نقاط مورد نظر (n) تقسیم شده است. همچنین برای تبدیل آن به درصد نیز، می‌توان آن را در عدد صد ضرب نمود.

۲-۷-۲-۱- الگوریتم ژنتیک^۱

الگوریتم ژنتیک یک الگوریتم بهینه‌سازی فرا ابتکاری است که از تکامل ژنتیکی به عنوان یک شیوه حل مسئله استفاده می‌کند و از طریق آن می‌توان مقدار بهینه یک یا چند متغیر را نسبت به هدف مورد نظر، برای کمینه یا بیشینه‌سازی تابع هدف تقریب زد [۲۸]. به منظور پیاده‌سازی روش‌های پیشنهادی تحقیق و همچنین ارزیابی آن‌ها توسط رویکردها و معیارهای تشریح شده، از زبان برنامه نویسی پایتون در محیط نرم

^۱ Genetic algorithm



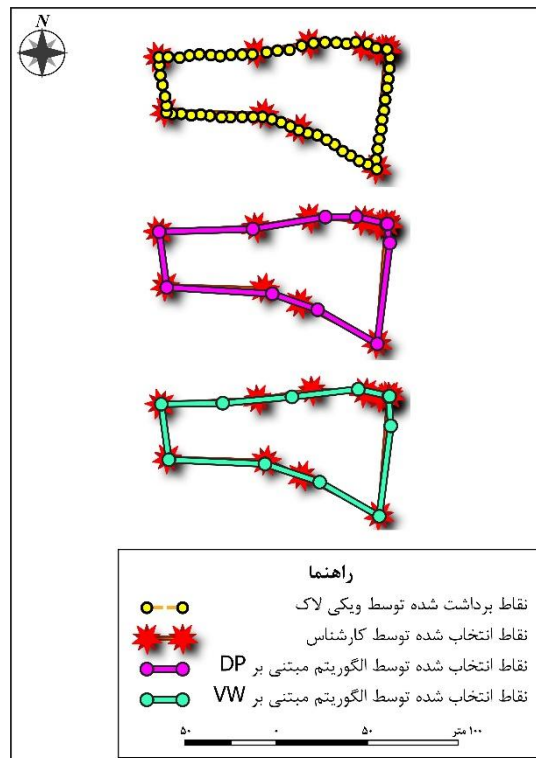
شکل ۶: نتایج انتخاب ده نقطه از چهارده نقطه اولیه

جدول ۱: نتایج درصد شباهت مساحت، نسبت به مساحت اولیه

نوع هندسی	مساحت (مترمربع)	درصد شباهت مساحت نسبت به هندسه اولیه
اولیه برداشت شده توسط گیرنده ماهواره‌ای	۵۱۸۷٫۰۸	۱
ساده شده بر مبنای نظر کارشناس	۵۲۰۳٫۲۹۲	۹۹٫۶۹
ساده شده بر مبنای الگوریتم مبتنی بر DP	۵۱۲۸٫۲۶۴	۹۸٫۸۶
ساده شده بر مبنای الگوریتم مبتنی بر VW	۵۱۶۹٫۸۰۵	۹۹٫۶۷

نقطه انتخاب شدند و اینبار توسط معیارهای $RMSE$ و ME و A عملکرد آن‌ها محک زده شد که نتایج آن در جداول (۲) و (۳) قابل مشاهده است.

جهت تکمیل بررسی انتخاب برترین n نقطه توسط الگوریتم‌های پیشنهادی، در بررسی بعدی، از میان ۶۳ نقطه برداشت شده در قالب خط سیر توسط نرم افزار ویکی‌لوک نمایش داده شده در شکل (۷)، تعداد ۱۰



شکل ۷: نتایج انتخاب ۱۰ نقطه از نقاط ویکی لاک

جدول ۲: نتایج درصد شباهت مساحت هندسه برداشت شده توسط ویکی لاک و ساده شده آن، نسبت به مساحت اولیه

نوع هندسی	مساحت (مترمربع)	درصد شباهت مساحت نسبت به هندسه اولیه
اولیه برداشت شده توسط گیرنده ماهواره‌ای	۵۱۸۷/۰۸	۱
اولیه برداشت شده توسط ویکی لاک	۵۲۰۳/۲۹۲	۹۷/۱۸
ساده شده بر مبنای الگوریتم مبتنی بر DP	۵۱۲۸/۲۶۴	۹۷/۱۳
ساده شده بر مبنای الگوریتم مبتنی بر VW	۵۱۶۹/۸۰۵	۹۸/۶۵

جدول ۳: مقایسه نتایج الگوریتم‌های پیشنهادی بر اساس شاخص‌های ME و $RMSE$

الگوریتم ساده‌سازی هندسی	$RMSE (m)$	$ME (m)$
الگوریتم مبتنی بر DP	۵/۶۰۷	۰/۹۵۹
الگوریتم مبتنی بر VW	۹/۰۷۳	۰/۹۵۱

چندانی ندارد. این در حالی است که درصد شباهت هندسه ساده شده توسط الگوریتم مبتنی بر VW حدود ۱/۵ درصد بهبود پیدا کرده است. اما بر اساس شاخص $EMSE$ ، الگوریتم ساده‌سازی مبتنی بر DP عملکرد بهتری را نسبت به الگوریتم ساده‌سازی مبتنی بر VW

مطابق نتایج جدول (۲)، درصد شباهت مساحت هندسه برداشت شده توسط نرم افزار ویکی لاک نسبت به مساحت هندسه برداشت شده توسط گیرنده ماهواره‌ای چندفرکانسه ۹۷/۱۸ درصد می‌باشد که پس از ساده‌سازی توسط الگوریتم مبتنی بر DP، تفاوت

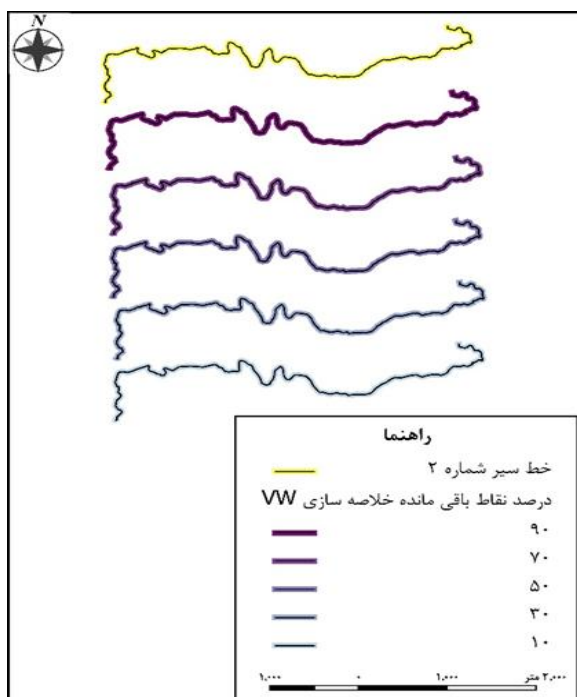
لازم به ذکر است که به طور کلی در برداشت‌های مربوط به صدور اسناد که خطای کمتر از ۱۰ سانت، برای موقعیت مطلق هریک از نقاط محدوده مورد نظر است، نمی‌توان از روش پیشنهادی استفاده نمود. اما هنگامی که هدف، برآورد مساحت عرصه می‌باشد، می‌توان با دقت ۹۸/۶۸ درصد ترکیب ویکی‌لوک و الگوریتم پیشنهادی تحقیق را جایگزین برداشت با گیرنده ماهواره‌ای نمود.

همچنین در برداشت برخی زمین‌های زراعی که حدود آن‌ها بدون حصار کشی و صرفاً با نهرها و عوارض طبیعی زمین مانند خط القعرها و خط الراس‌ها مشخص می‌شود، در عمل برداشت با دقت زیر ۱۰ سانت، حتی با گیرنده ماهواره‌ای ناممکن می‌شود و در چنین مواردی خطای رویکرد پیشنهادی این تحقیق، می‌تواند برای انتخاب نقاط، مورد قبول باشد.

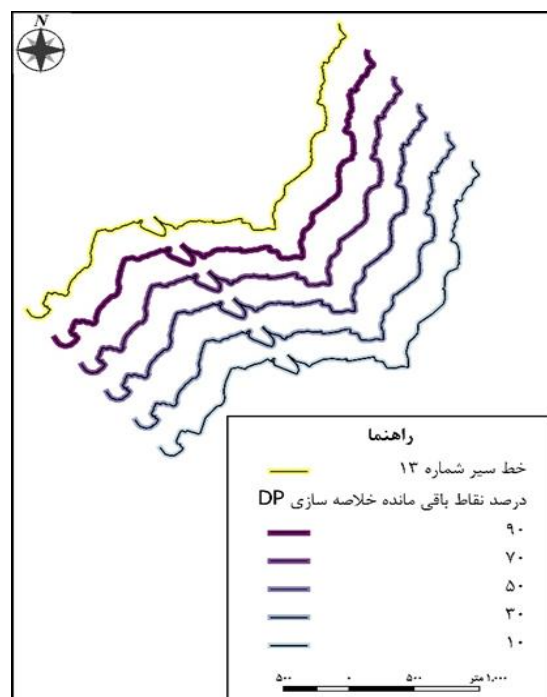
۳-۲- نتایج بررسی کیفیت ساده‌سازی هندسی الگوریتم‌های پیشنهادی تحقیق

در فاز بعدی ۱۴ خط سیر معرفی شده در این تحقیق مورد ارزیابی قرار گرفت. ساده‌سازی نقاط این خط سیرها در ۵ حالت مختلف (یعنی با نگه داشتن ۹۰٪، ۷۰٪، ۵۰٪، ۳۰٪ و ۱۰٪ نقاط آن‌ها) به کمک روش‌های مبتنی بر DP و VW انجام شد که به طور نمونه، بخشی از خروجی‌های به دست آمده برای خط سیر شماره ۱۳ و ۲ به ترتیب با الگوریتم‌های DP و VW در شکل (۸) و شکل (۹) نشان داده شده است. سپس برای ارزیابی نتایج از شاخص‌های ME ، MC و AC استفاده گردید. میانگین ارزیابی نتایج به دست آمده از سه شاخص مذکور، برای ۱۴ خط سیر مورد استفاده در این تحقیق، در درصدهای ساده‌سازی مذکور، در جدول (۴) قابل مشاهده می‌باشد.

از خود نشان داده است. $RMSE$ ریشه میانگین مربعات خطاست و منظور از خطا فاصله هر نقطه از عارضه ساده شده به نزدیکترین نقطه از عارضه اولیه است که توسط کارشناس انتخاب شده است. هرچه میزان $RMSE$ کمتر باشد، نشان دهنده این نکته است که الگوریتم در انتخاب نقاط بهتر عمل کرده است. چراکه $RMSE$ بر روی نقطه تمرکز دارد و فاصله کمتر، حاکی از نزدیکی نقاط انتخابی به نقاط ساده شده توسط الگوریتم بوده است. به بیان دیگر $RMSE$ ، مقدار خطا در انتخاب درست نقاط اصلی را می‌سنجد که بر اساس نتایج، الگوریتم مبتنی بر DP خطای کمتری در انتخاب نقاط داشته است. محیط قطعه زمین برداشت شده حدود ۳۴۶ متر است که با توجه به اینکه نرم افزار ویکی‌لوک از ۶۳ نقطه برای ثبت آن استفاده کرده است، به طور میانگین نقاط متوالی حدود ۴/۵ متر با هم فاصله دارند. بنابراین برای اینکه گفته شود با پذیرفتن چه مقدار خطا می‌توان از داده‌های ویکی‌لوک ساده شده با الگوریتم‌های پیشنهادی بجای برداشت توسط گیرنده‌های ماهواره‌ای استفاده کرد، شاخص $RMSE$ مناسب نیست و صرفاً می‌تواند برای مقایسه عملکرد دو الگوریتم استفاده شود. معیار بعدی که ME می‌باشد، تقسیم جمع مساحت‌های پلی‌گون‌های اسلایور بر طول بوده است که میانگینی از خطا را می‌دهد. این معیار نشان می‌دهد که، با پذیرفتن چه میزانی از خطا می‌توان از داده‌های ویکی‌لوک (که با سرعت بالا و رایگان به دست می‌آیند) به جای داده‌های گیرنده ماهواره‌ای چندفرکانسه (که هزینه و زمان بالایی برای جمع‌آوری آن‌ها نیاز است) استفاده کرد. مطابق با نتایج نشان داده شده در جدول (۳)، با پذیرفتن خطای در حدود ۹۵ سانتی‌متر که برای الگوریتم‌های پیشنهاد شده تقریباً یکسان است، می‌توان از داده‌های ویکی‌لوک به جای داده‌های گیرنده ماهواره‌ای چندفرکانسه استفاده کرد که به راحتی به کمک تلفن‌های همراه هوشمندی که در دست عموم مردم قرار دارد، قابل جمع‌آوری است.



شکل ۹: ساده‌سازی خط سیر شماره ۲ توسط الگوریتم مبتنی بر VW



شکل ۸: ساده‌سازی خط سیر شماره ۱۳ توسط الگوریتم مبتنی بر DP

جدول ۴: مقایسه نتایج ارزیابی ساده‌سازی خطوط سیر توسط الگوریتم‌های پیشنهادی با درصدهای ساده‌سازی مختلف

شاخص	الگوریتم	۱۰	۳۰	۵۰	۷۰	۹۰
ME(m)	DP	۱,۹۵	۰,۳۹	۰,۱۳	۰,۰۴	۰,۰۰۳
	VW	۱,۷۶	۰,۳۵	۰,۱۲	۰,۰۳	۰,۰۰۳
MC(%)	DP	۲۷,۲۳	۶۰,۳۹	۸۱,۲۷	۹۳,۳۵	۹۲,۲۰
	VW	۲۴,۴۸	۵۶,۸۷	۸۰,۰۲	۹۲,۲۰	۹۹,۲۷
AC(%)	DP	۲۵,۸۲	۵۹,۸۰	۸۱,۰۴	۹۳,۳۱	۹۲,۰۸
	VW	۲۲,۹۹	۵۶,۱۰	۷۹,۷۱	۹۲,۰۸	۹۹,۲۷

۳-۳- نتایج مقایسه الگوریتم‌های پیشنهادی تحقیق نسبت به روش‌های متداول در انتخاب برترین نقاط

به منظور بررسی خطای نسبی الگوریتم‌های پیشنهادی مبتنی بر DP و مبتنی بر VW تحقیق، نسبت به شکل متداول آن‌ها، توسط هریک از الگوریتم‌ها در مدت زمانی یکسان، به ساده‌سازی پرداخته شد. این عملیات به ازای هر خط سیر، به ترتیب با کاهش ۱۰، ۳۰، ۵۰، ۷۰ و ۹۰ درصدی تعداد نقاط تکرار پذیرفت.

نظر به نتایج نمایش داده شده در جدول (۴)، به طور کل، بر اساس شاخص ME، الگوریتم مبتنی بر VW در ساده‌سازی خطوط سیر ثبت شده توسط ویکی‌لوک، به طور میانگین حدود پنج سانتی‌متر خطای کمتری نسبت به الگوریتم مبتنی بر DP دارد. بر اساس شاخص‌های MC و AC نیز، الگوریتم پیشنهادی مبتنی بر DP به ترتیب به طور میانگین ۰,۳۲ درصد و ۰,۷۸ درصد نسبت به الگوریتم مبتنی بر VW عملکرد بهتری داشته است.

شوند و بدین صورت می‌توان توسط رابطه (۱۰) برای آن‌ها خطای نسبی محاسبه کرد. در جدول (۵) نشان داده شده است که با کاهش ۱۰، ۳۰، ۵۰، ۷۰ و ۹۰ درصدی نقاط، خطای نسبی الگوریتم‌های DP و VW که مقدار معیار هندسی آن‌ها (T) توسط الگوریتم ژنتیک بهینه‌سازی می‌شود، به طور میانگین برای ۱۴ خط سیر جمع آوری شده، چند درصد خواهد بود.

منطق الگوریتم‌های پیشنهادی این پژوهش به نحوی است که همیشه تعداد نقاط در خروجی الگوریتم (N_{out})، برابر با تعداد نقاط مدنظر (n) می‌شود و بدین ترتیب تفاضل آن‌ها همیشه برابر با صفر خواهد بود و در نتیجه، خطای نسبی برابر با صفر محاسبه خواهد شد. اما در شکل متداول الگوریتم‌ها، مقدار حد آستانه، توسط کاربر و یا در روندی هوشمندانه و خودکار، می‌تواند توسط یک الگوریتم بهینه‌سازی مانند الگوریتم ژنتیک، تعیین گردد که باعث می‌شود نتایج تقریبی

جدول ۵: مقایسه نتایج ارزیابی ساده‌سازی خطوط سیر توسط الگوریتم‌های پیشنهادی با درصد‌های ساده‌سازی مختلف

الگوریتم	درصد کاهش نقاط	۱۰	۳۰	۵۰	۷۰	۹۰
DP	درصد خطای نسبی	۸٫۹۳	۳۱٫۲۱	۶۸٫۱۴	۶۱٫۰۷	۲۱٫۷۱
	مدت زمان پردازش (ثانیه)	۸٫۱۰	۸٫۱۷	۸٫۱۸	۸٫۶۸	۸٫۵۱
VW	درصد خطای نسبی	۹٫۵۰	۳۵٫۶۴	۷۵٫۷۱	۵۶٫۴۳	۲۳٫۸۶
	مدت زمان پردازش (ثانیه)	۰٫۸۱	۳٫۱۱	۴٫۸۳	۰٫۲۰	۰٫۴۲

نتایج نشان می‌دهند که می‌توان نقاط عارضه خطی را با استفاده از ویژگی هندسی مورد تأکید در الگوریتم ساده‌سازی هندسی مورد نظر، رتبه بندی نمود و سپس از بهترین‌های آن‌ها به تعداد دلخواه انتخاب کرد. مزیت رویکرد پیشنهادی تحقیق حاضر نسبت به روال متداول الگوریتم‌های خلاصه‌سازی که اکثراً از حدآستانه‌های هندسی مانند فاصله و مساحت به عنوان معیار قابل تنظیم استفاده می‌کنند، کنترل بر میزان حجم ساده‌سازی هندسی است. در روال متداول، برای رسیدن به حجم خلاصه‌سازی مورد نظر در هنگام استفاده از حدآستانه هندسی، نیاز به آزمون و خطا است و حتی در حالت هوشمندانه‌تر می‌توان از الگوریتم‌های بهینه‌سازی استفاده نمود تا فهمید که با تنظیم حدآستانه هندسی روی چه مقداری، می‌توان به

همان‌طور که از جدول (۵) مشخص است، به طور میانگین، به ترتیب خطای نسبی الگوریتم‌های DP و VW بهینه شده با الگوریتم ژنتیک، ۳۸٫۲۱۲ و ۴۰٫۲۲۸ درصد می‌باشد، که نسبت به الگوریتم‌های پیشنهادی این پژوهش عملکرد بسیار ضعیف‌تری را از خود نشان داده‌اند.

همچنین با توجه به مدت زمان پردازش، می‌توان نتیجه گرفت که به طور میانگین، الگوریتم مبتنی بر VW به میزان ۴٫۴۴ برابر سریع‌تر از الگوریتم مبتنی بر DP می‌تواند به نتیجه مورد نظر برسد.

۴- نتیجه گیری و جمع بندی

به منظور ارائه یک رویکرد کلی با هدف کنترل بر میزان حجم ساده‌سازی هندسی، روند پیشنهادی به طور نمونه بر روی الگوریتم‌های DP و VW پیاده‌سازی شد و

هندسی، حفظ مساحت از اهمیت بالایی برخوردار است، استفاده از الگوریتم مبتنی بر VW می‌تواند به نتایج بهتری منجر شود و زمانی که اندازه طول‌ها و زوایا و موقعیت مطلق نقاط شکستگی مهم‌تر می‌باشد، بهتر است از الگوریتم مبتنی بر DP استفاده نمود.

نظر به نتایج نمایش داده شده در جدول (۲)، درصد شباهت مساحت هندسه ساده شده توسط الگوریتم مبتنی بر VW ، حتی از هندسه اولیه برداشت شده نیز بیشتر است. بدین معنی که نه تنها از تعداد نقاط (که ازدیاد آن علاوه بر پیچیدگی هندسی منجر به افزایش فضای مورد نیاز برای ذخیره‌سازی می‌شود) کاسته می‌شود، بلکه درصد شباهت آن به هندسه اصلی عارضه افزایش می‌یابد. این افزایش، نشان‌دهنده کارایی بالای این الگوریتم در انتخاب برترین نقاط است.

الگوریتم‌های ساده‌سازی هندسی فراوانی توسط محققین مختلف ارائه شده‌اند که علاوه بر مساحت و فاصله، مشخصات هندسی دیگری از قبیل زاویه و مورفولوژی را به منظور ساده‌سازی مورد استفاده قرار داده‌اند. پیشنهاد داده می‌شود با در نظر گرفتن روند کلی ارائه شده در این تحقیق، با هدف کنترل بر حجم ساده‌سازی هندسی، الگوریتم‌های پرکاربرد دیگر نیز مجدداً پیاده‌سازی شوند تا بتوان در پایگاه‌های داده مکانی شامل عوارض خطی مورد استفاده قرار گیرند.

حجم ساده‌سازی مورد نظر رسید و این خود نیازمند تکرار اجرای الگوریتم به ازای حدآستانه‌های متفاوت است. در این صورت نه تنها رسیدن به نتیجه، مدت زمان بیشتری به طول می‌انجامد، بلکه میزان ساده‌سازی حاصل، تقریبی از میزان ساده‌سازی مورد نظر خواهد بود. اما در رویکرد پیشنهادی این پژوهش، نیازی به تکرار نیست و پس از اجرای الگوریتم، حجم ساده‌سازی مطابق با خواسته کاربر خواهد بود که به طور میانگین به بهبود $۳۸/۲۱۲$ و $۴۰/۲۲۸$ درصدی خطای نسبی، برای الگوریتم‌های مبتنی بر DP و مبتنی بر VW منجر شده است.

همان‌طور که از پیش گفته شد، انتخاب نقاط باید به گونه‌ای باشد که هندسه ساده شده نسبت به هندسه اولیه دارای بیشترین شباهت و به بیان دیگر دارای کمترین اختلاف باشد. این اختلاف و شباهت توسط شاخص‌های مختلف مورد بررسی قرار گرفت. به طور کلی، نتایج حاکی از آن است که بر اساس شاخص‌های ME و A که مبتنی بر مساحت هستند، الگوریتم مبتنی بر VW در مجموع دارای عملکرد بهتری نسبت به الگوریتم مبتنی بر DP داشته است و به همین ترتیب بر اساس شاخص‌های $RMSE$ ، MC و AC که مبتنی بر فاصله و زاویه هستند، الگوریتم مبتنی بر DP عموماً بهتر از الگوریتم مبتنی بر VW عمل کرده است. بنابراین می‌توان نتیجه گرفت، زمانی که در ساده‌سازی

مراجع

- [1] X. Lin, S. Ma, J. Jiang, Y. Hou, and T. Wo, "Error bounded line simplification algorithms for trajectory compression: An experimental evaluation", *ACM Transactions on Database Systems (TODS)*, vol. 46, no. 3, pp. 1-44, 2021.
- [2] J. L. Morrison, "Map Generalization: Theory, Practice, and Economics", in *Second International Symposium on Computer-Assisted Cartography*, 1975.
- [3] K. Misue, "Conditions of Preserving Mental Images by Contour Deformation", in *Proceedings of the 15th International Symposium on Visual Information Communication and Interaction*, 2022, pp. 1-5.
- [4] J. Jaafar, "Line generalization: least square with double tolerance", *WIT Transactions on Information and Communication Technologies*, vol. 26, 2002.
- [5] B. Liu, X. Liu, D. Li, Y. Shi, G. Fernandez, and Y. Wang, "A vector line simplification algorithm based on the Douglas-Peucker algorithm, monotonic chains and dichotomy", *ISPRS International Journal*

- of Geo-Information, vol. 9, no. 4, p. 251, 2020.
- [6] A. Pinandito and C. P. Wulandari, "Integrating douglas-peucker line simplification into routeboxer algorithm on a map-based Android application", in *Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology*, 2020, pp. 213–219.
- [7] Y. Ren, J. Tang, and S. Wu, "Geometric properties preserved line simplification algorithm based on fractal", in *2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 3019–3022.
- [8] V. H. Clayton, "Cartographic generalization: a review of feature simplification and systematic point algorithms", 1985.
- [9] R. B. McMaster, "Automated line generalization", *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 24, no. 2, pp. 74–111, 1987.
- [10] J. A. Young, "Effects of automated cartographic generalization on linear map features", *Virginia Tech*, 1991.
- [11] D. Kotsur and V. Tereshchenko, "Optimization heuristics for computing the Voronoi skeleton", in *International Conference on Computational Science*, 2019, pp. 96–111.
- [12] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature", *Cartographica: the international journal for geographic information and geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.
- [13] T. Samsonov and O. Yakimova, "Regression modeling of reduction in spatial accuracy and detail for multiple geometric line simplification procedures", *International Journal of Cartography*, vol. 6, no. 1, pp. 47–70, 2020.
- [14] R. B. McMaster, "A mathematical evaluation of simplification algorithms (in computer cartography)", *Scanning Electron Microsc Meet at*, pp. 267–276, 1983.
- [15] B. J. Kronenfeld and J. Deng, "Between the Lines: Measuring Areal Displacement in Line Simplification", *Advances in Cartography and GIScience of the ICA*, vol. 1, p. NA-NA, 2019.
- [16] G. F. Jenks, "Geographic logic in line generalization", *Cartographica: The International Journal for Geographic Information and Geovisualization*, vol. 26, no. 1, pp. 27–42, 1989.
- [17] Y. Shen, T. Ai, and Y. He, "A new approach to line simplification based on image processing: A case study of water area boundaries", *ISPRS International Journal of Geo-Information*, vol. 7, no. 2, p. 41, 2018.
- [18] P. Zhao, Q. Zhao, C. Zhang, G. Su, Q. Zhang, and W. Rao, "CLEAN: Frequent pattern-based trajectory compression and computation on road networks", *China Communications*, vol. 17, no. 5, pp. 119–136, 2020.
- [19] J. Du, F. Wu, J. Yin, C. Liu, and X. Gong, "Polyline simplification based on the artificial neural network with constraints of generalization knowledge", *Cartography and Geographic Information Science*, vol. 49, no. 4, pp. 313–337, 2022.
- [20] M. Visvalingam and J. D. Whyatt, "Line generalization by repeated elimination of points", in *Landmarks in Mapping*, Routledge, 2017, pp. 144–155.
- [21] M. Mizuta, "Algebraic curve fitting for multidimensional data with exact squares distance", in *1996 IEEE International Conference on Systems, Man and Cybernetics. Information Intelligence and Systems (Cat. No. 96CH35929)*, 1996, vol. 1, pp. 516–521.
- [22] B. Sheng, S. Wenzhong, F. Wenzheng, C. Pengxin, N. Mingyan, and X. Haodong, "A

- tight coupling mapping method to integrate the ESKF, g2o, and point cloud alignment'', *The Journal of Supercomputing*, vol. 78, no. 2, pp. 1903–1922, 2022.
- [23] H. Liu et al., 'Influences of different surveying and mapping methods on fractal characteristics of gully-head shoulder lines'', *Physical Geography*, vol. 37, no. 6, pp. 387–408, 2016.
- [24] Y. Filippovska, V. Walter, and D. Fritsch, 'Quality evaluation of generalization algorithms'', *ISPRS Commission II, WG II*, vol. 7, 2008.
- [25] M. Garland and Y. Zhou, 'Quadric-based simplification in any dimension'', *ACM Transactions on Graphics (TOG)*, vol. 24, no. 2, pp. 209–239, 2005.
- [26] Y. Zhang, H. Liu, D. Chen, and P. Xu, 'Stylized line rendering for three-dimensional models'', in *2018 International Conference on Information and Computer Technologies (ICICT)*, 2018, pp. 52–60.
- [27] W. Shi and C. Cheung, 'Performance evaluation of line simplification algorithms for vector generalization'', *The Cartographic Journal*, vol. 43, no. 1, pp. 27–44, 2006.
- [28] A. Lambora, K. Gupta, and K. Chopra, 'Genetic algorithm-A literature review'', in *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*, 2019, pp. 380–384.



A new approach in implementing geometric simplification algorithms of linear features with the aim of controlling the degree of simplification

Amir Gholami ¹, Abolfazl Ghanbari ^{2*}, Neda Kaffash Charandabi ³

1- Ms.c student of RS & GIS in Department of RS & GIS, College of Environmental sciences, University of Tabriz

2- Full Professor in Department of RS & GIS, College of Environmental sciences, University of Tabriz

3- Assistant professor in Department of Geomatics, Technical College of Marand, University of Tabriz

Abstract

Geographers and cartographers have been interested in the geometric simplification of linear features since the early 1900s. As time went by and data creation rose, it became much more significant. Most of the algorithms described by different researchers work on the basis of the threshold limit and have no special control on the amount of geometric simplification. Thus, in this study, a technique was proposed for determining the percentage or number of reductions in the beginning points of the linear features without using trial and error. This methodology was used to the Douglas-Poker (DP) and Visvalingam-Whyatt (VW) algorithms to identify the optimal nodes for an arbitrary quantity (n), and the results were assessed using three distinct datasets. It demonstrated that, while cadastering agricultural land, each of the suggested algorithms can accurately choose 80% of the spots chosen by an expert surveyor, with a 98.68% accuracy rate in computing the area and a 95 cm error rate in determining the location.

The results of checking the quality of simplification by the proposed algorithms based on DP and VW, according to three different indicators, show that the average error by sliver polygons, the average curvature similarity percentage, and the angle sharpness similarity percentage for 10%, 30%, 50%, 70%, and 90% simplification on the fourteen recorded trajectories, was 0.50, 72.30, and 71.85 for the DP-based algorithm and 0.45, 70.34, and 69.84 for the VW-based approach.

It has also been demonstrated that, even after optimizing the threshold value of DP and VW algorithms using the genetic optimization algorithm, their relative error in selecting the best points for a given number, as an average for 14 collected trajectories, is 38.212% and 40.228% higher than the ones proposed by the research in the same processing time.

Key words: Geometric Simplification, Linear Feature, Trajectory.