

## بهبود الگوریتم شبیه‌سازی زمین‌آماری چند نقطه‌ای فیلترسیم از طریق طراحی فیلترهای سازگار با الگوهای موجود در تصاویر آموزشی و شرطی‌سازی نظارت شده

منصوره شریف زاده لاری<sup>۱</sup>، نادر فتحیان پور\*<sup>۲</sup>، رسول امیرفتاحی<sup>۲</sup>، سعید صدوری<sup>۴</sup>

۱- دانشجوی دکتری، آزمایشگاه تحقیقاتی پردازش سیگنال دیجیتال، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی اصفهان

۲- دانشیار، دانشکده مهندسی معدن، دانشگاه صنعتی اصفهان

۳- دانشیار، آزمایشگاه تحقیقاتی پردازش سیگنال دیجیتال، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی اصفهان

۴- دانشیار، آزمایشگاه تحقیقاتی پردازش سیگنال دیجیتال، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی اصفهان

(دریافت: بهمن ۱۳۹۴، پذیرش: اسفند ۱۳۹۶)

### چکیده

شبیه‌سازی چند نقطه‌ای یکی از ابزارهای زمین‌آماری است که در سال‌های اخیر مورد توجه بسیاری از محققان علوم زمین نفتی قرار گرفته و از آن برای توصیف ساختارهای فضایی فرایندهای زمین‌شناسی حاکم بر مخازن نفتی استفاده شده است. روش زمین‌آمار چند نقطه‌ای با لحاظ آماره‌های چند نقطه‌ای به جای آماره‌ی دو نقطه‌ای مرسوم در روش‌های مبتنی بر واریوگرام، قادر است الگوهای فضایی اشکال و ساختارهای با هر هندسه پیچیده‌ای را باز یابی نماید. باز تولید ساختارهای چند نقطه‌ای موجود در فرایندهای زمین‌شناسی، از طریق به کارگیری تصویر آموزشی انجام می‌شود که شامل ساختارهای کلی است که زمین‌شناسان به عنوان الگوهای مفهومی حاکم بر فرایندهای زمین‌شناسی زیر سطحی ارائه می‌دهند. یکی از الگوریتم‌های کاربردی و مهم زمین‌آماری چند نقطه‌ای بر پایه الگو، فیلترسیم است که از آن برای مدل کردن متغیرهای پیوسته و گسسته استفاده می‌شود. در روش کلاسیک فیلترسیم ابتدا همه الگوهای شناسائی شده در تصویر آموزشی با استفاده از تعدادی فیلتر با مشخصات ثابت و غیر مرتبط با تصویر آموزشی به فضای جدید امتیاز فیلتری منتقل و سپس شبیه‌سازی شرطی انجام می‌شود. در این مقاله با استفاده از روش آنالیز مولفه اصلی، فیلترها را با توجه به ویژگی‌های تصویر آموزشی و بصورت وفقی با آن طراحی کرده و بعد از طبقه‌بندی فضای امتیاز فیلتری جدید، الگوی مناسب با توجه به داده‌های سخت موجود در پنجره الگو انتخاب شده و در محل تصویر شبیه‌سازی شده چسبانده می‌شود. در الگوریتم فیلترسیم از ضرائب ثابت جهت وزن دادن به داده‌ها در مراحل شبیه‌سازی شرطی استفاده شده و با توجه به نتایج نامناسب مربوط به استفاده از این ضرائب ثابت، در این مقاله یک الگوریتم شرطی‌سازی نظارت شده با تعریف روابط مناسب ارائه شده و سهم داده‌های مختلف موجود در پنجره الگو در فرایند شبیه‌سازی، به طور بهینه با استفاده از این روابط اعمال می‌شود. مقایسه نتایج به دست آمده از این روش پیشنهادی با روش فیلترسیم کلاسیک نشان دهنده بهبود قابل ملاحظه‌ای در باز یابی پیوستگی‌های مورد انتظار در نتایج شبیه‌سازی است.

### کلمات کلیدی

زمین‌آمار چند نقطه‌ای، آنالیز مولفه اصلی، فیلترسیم

## ۱- مقدمه و سابقه موضوع تحقیق

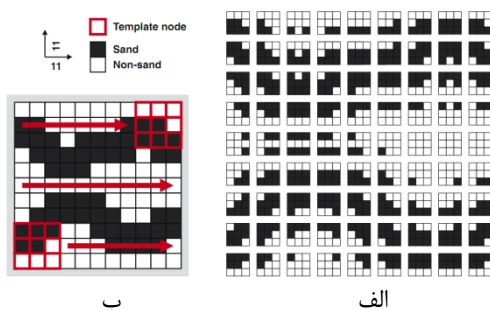
۲-

## آمار چند نقطه‌ای

امروزه به کارگیری روش‌های زمین‌آماري و به ویژه گونه‌های چند نقطه‌ای آن در شبیه‌سازی مخازن پیچیده نفتی امری اجتناب‌ناپذیر می‌نماید. تکنیک شبیه‌سازی چندنقطه‌ای<sup>۱</sup> در سال ۱۹۹۳ ارائه شد. این تکنیک محاسن الگوریتم‌های بر پایه واریوگرام، بر پایه شیء و بر پایه موضوع را داشت [۱]. اولین الگوریتم کاربردی این تکنیک، SNESIM<sup>۲</sup> بود. در این روش تمام توزیع احتمال شرطی لازم در یک ساختار داده‌ای بنام درخت جستجو ذخیره می‌شد [۲]. این روش امروزه در برخی کاربردها [۳] استفاده می‌شود. اما در موارد ساختارهای پیچیده هندسی و چند رخساره‌ای زمان محاسبات بسیار بالا می‌رفت. به منظور رفع مشکل برنامه‌نویسی پیچیده و زمان شبیه‌سازی‌ها، روش‌های بر پایه الگو همچون فیلترسیم معرفی شدند [۴]. الگوریتم فیلترسیم توانایی کار بر روی تصاویر آموزشی با متغیرهای پیوسته و گسسته (رخساره‌ای) را دارند. در این روش فیلترهایی معرفی شد تا الگوهای با ابعاد بالا را به یک امتیاز فیلتری کوچک خلاصه کند و سپس الگوها را به یک تعداد محدود کلاس طبقه‌بندی کردند. فیلترها برای سرعت بخشیدن به کار شبیه‌سازی الگوها تعریف شدند. در سال ۲۰۰۸ این الگوریتم به صورت کاربردی و تجاری در نرم‌افزار SGeMS<sup>۳</sup> پیاده‌سازی شد [۵]. در کاری دیگر روشی ارائه شده که سرعت الگوریتم فیلترسیم را در کاربردهای سه بعدی افزایش می‌دهد [۶]. یکی از کاستی‌های الگوریتم فیلترسیم این است که جهت کاهش بُعد، از فیلترهای ثابتی استفاده می‌کند و این فیلترها با تغییر تصویر آموزشی تغییر نمی‌کنند. در این مقاله با استفاده از روش تحلیل مولفه اصلی<sup>۳</sup> فیلترهایی طراحی شده است که بیشترین سازگاری را با الگوهای موجود در تصویر آموزشی دارند. برای مقایسه روش اجرا شده با الگوریتم فیلترسیم، تعداد فیلترهای طراحی شده ۶ فیلتر هستند که البته این تعداد قابل تغییر هستند. سپس برای شرطی کردن شبیه‌سازی با داده‌های مشاهده‌ای سخت همچون چاه، الگوریتم جدیدی ارائه شده است. در این الگوریتم، بر خلاف روش ارائه شده در الگوریتم فیلترسیم، که ضرائب مربوط به داده‌های مختلف در کل الگوریتم ثابت بودند، این ضرائب با یک سری روابط به دست می‌آیند.

گواردیانو و سری واستاوا در سال ۱۹۹۳، آمار چند نقطه‌ای را برای مدلسازی ناهمگنی‌های پیچیده زمین‌شناسی با قدرت تفکیک بالا در مخازن نفتی مطرح کردند. این روش، مخازن را با استفاده از الگوهای استخراج شده از یک تصویر آموزشی و شرطی شده با داده‌های چاه، شبیه‌سازی می‌کند [۷]. آمار چند نقطه‌ای در بسیاری از کاربردها از جمله شبیه‌سازی مخازن نفتی [۸]، توصیف آماری داده‌های فضایی [۹] و بازسازی داده‌های صوتی [۱۰] استفاده شده است.

اساس روش آمار چندنقطه‌ای بر مبنای ابزاری به نام تصویر آموزشی است که نمایشی عددی از الگوها و ساختارهایی است که انتظار می‌رود در مخزن مورد نظر که باید شبیه‌سازی شود، وجود داشته باشند. آماره‌های چندنقطه‌ای می‌توانند به صورت احتمال شرطی پدیدار شوند [۱۱]. یکی از مهم‌ترین الگوریتم‌های الگوپایه مورد استفاده جهت شبیه‌سازی استاتیکی مخزن که تجاری‌سازی هم شده است، الگوریتم شبیه‌سازی فیلترسیم است. در الگوریتم کلاسیک فیلترسیم از تعداد اندکی فیلتر خطی (۶ فیلتر بری مقاطع دو بُعدی و ۹ فیلتر برای مدل‌های سه بعدی) برای دسته‌بندی الگوهای آموزشی در یک فضای امتیاز فیلتری کاهش بعد یافته استفاده می‌کند [۱۱]. به این ترتیب که الگوهای مشابه، گروه‌بندی شده و در یک دسته ذخیره می‌شوند. نماینده<sup>۴</sup> (الگوی اصلی) این دسته، میانگینی از الگوهای قرار گرفته در آن است. در شکل ۱ الگوهای آموزشی به دست آمده از یک تصویر آموزشی نمونه نشان داده شده است.



شکل ۱: یک تصویر آموزشی نمونه (ب) و الگوهای آموزشی با توجه به پنجره مربعی ۳×۳ (الف) [۱۲]

در این الگوریتم، الگوی جستجو به شکل مکعب مستطیل (در فضای سه بُعدی) و یا مستطیل (در فضای دو بُعدی) است.

در یک دسته مشخص می شود. در این مقاله برای طبقه بندی فضای امتیاز فیلتری از الگوریتم K-means استفاده شده است. سپس یک پنجره به طور تصادفی از داده مشاهده ای انتخاب شده و با تمام میانگین دسته ها مقایسه می شود. معیار مقایسه در الگوریتم فیلترسیم می تواند نرم ۱، فاصله اقلیدسی و یا هر معیار فاصله دیگری باشد. نزدیک ترین میانگین دسته به داده مشاهده ای مشخص شده و از دسته مربوط یک الگو تصادفی انتخاب شده و در داده مشاهده ای درج می شود. قسمت داخلی این پنجره با یک طول مشخص، در این تصویر کاملاً ثابت شده و در شبیه سازی های بعد تغییر نمی کند.

### ۳- تحلیل مولفه اصلی (PCA)

تحلیل مولفه اصلی تبدیلی در فضای برداری است که بیشتر برای کاهش ابعاد مجموعه داده ها مورد استفاده قرار می گیرد [۱۴]. این تحلیل، شامل تجزیه مقادیر ویژه ماتریس کواریانس چند متغیره است. در این روش ابتدا بردار میانگین  $\mu$  و ماتریس کواریانس  $\Sigma$  با توجه به تمام داده ها محاسبه می شود. در ادامه مقادیر ویژه و بردارهای ویژه ماتریس کواریانس محاسبه شده و با توجه به مقادیر ویژه به صورت نزولی مرتب می شوند. سپس تعداد  $k$  بردار ویژه متناظر با بزرگترین مقادیر ویژه انتخاب شده و ماتریس  $A$  را می سازند که ستون های آن، بردارهای ویژه به دست آمده هستند. در این صورت می توان تمام داده های  $n$  بعدی را طبق رابطه ۳ با ماتریس  $A$  به فضای  $k$  بعدی منتقل نمود.

$$x' = A^t(x - \mu) \quad (3)$$

می توان نشان داد که تحلیل مولفه اصلی خطای مربع را کمینه می کند [۱۴].

### ۴- انتخاب پنجره مناسب برای اجرای الگوریتم شبیه سازی چند مقیاسه

بازیابی ساختارهای با مقیاس های مختلف در مدل تولید شده نهایی به شدت تحت تاثیر انتخاب شکل و اندازه پنجره جستجو است [۱]. بزرگ بودن اندازه پنجره باعث می شود اطلاعات نقاط دور از هم در تصویر آموزشی امکان شناسایی الگوهای ساختاری بزرگ مقیاس را بدهد. از طرفی ابعاد کوچک پنجره مانع باز تولید الگوهای با اندازه بزرگ می شود.

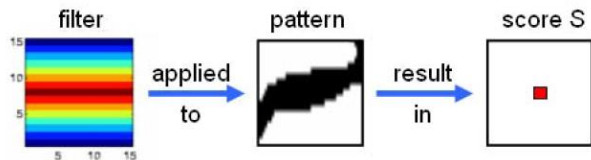
یک فیلتر در واقع یک دنباله ای از وزن های مرتبط با نقاط الگوی جستجو است که به صورت رابطه ۱ تعریف می شود [۱۳].

$$\{f(h_j); j = 1 \dots J\} \quad (1)$$

در این رابطه  $f(h_j)$  مقدار فیلتر یا وزن مرتبط با نقاط متناظر از الگوی جستجو است. این فیلتر به هر کدام از الگوهای آموزشی که از تصویر آموزشی استخراج شده با توجه به رابطه ۲، اعمال می شود.

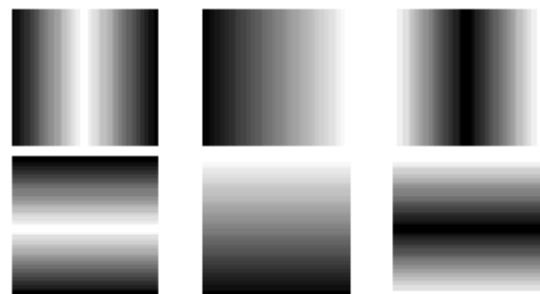
$$S_T(u) = \sum_{j=1}^n f(h_j) \cdot pat(u + h_j) \quad (2)$$

خروجی فیلتر به صورت مقدار میانگین وزن دار  $S_T(u)$  که مقدار امتیاز فیلتری هم نامیده می شود ذخیره می شود. در شکل ۲ فرایند ایجاد یک مقدار امتیاز فیلتری از طریق اعمال یک فیلتر دو بعدی خاص نشان داده شده است.



شکل ۲: به دست آمدن یک مقدار امتیاز از یک الگوی آموزشی [۹]

در الگوریتم فیلترسیم، سه نوع فیلتر میانگین، گرادیان و انحنای هر کدام از جهت های  $X, Y$  و  $Z$  در نظر گرفته شده است. به عنوان مثال برای یک تصویر آموزشی دو بعدی، فیلترها به صورت شکل ۳ هستند. این فیلترها برای هر دو نوع داده پیوسته و رخساره ای و برای هر نوع تصویر آموزشی استفاده می شود. بر این اساس، برای حالت دوبعدی، تعداد ۶ فیلتر و برای حالت سه بعدی، نه فیلتر وجود خواهد داشت.



شکل ۳: فیلترهای مربوط به الگوریتم فیلترسیم با پنجره ۳۲×۳۲

بعد از مشخص شدن فیلترها، تمام الگوهای به دست آمده با یک ابعاد پنجره مشخص در فیلترها ضرب شده و به فضای امتیاز فیلتری برده می شوند [۴]. پس از طبقه بندی الگوها در فضای امتیاز فیلتری، میانگین الگوهای قرار گرفته

الگوریتم‌های MPS، مقادیر فرد هستند بنابراین در این تحقیق ابعاد پنجره‌ها به صورت رابطه ۵ بکار گرفته شده است.

$$T = \{(3 \times 3), (5 \times 5), \dots, (n' \times n')\} \quad (5)$$

که در آن  $n' = 0.4 \times \max(N_x \times N_y)$  و  $N_x$  و  $N_y$  ابعاد تصویر آموزشی هستند.

سپس آنتروپی متوسط (ME) مربوط به همه الگوها با ابعاد مشخص، محاسبه می‌شود. پنجره با ابعاد مشخص  $\omega \times \omega$  با  $T^\omega$  نمایش داده می‌شود. رابطه محاسبه آنتروپی متوسط برای  $T^\omega$  به صورت رابطه ۶ است.

$$ME(\omega) = \frac{1}{n_{T^\omega}} \sum_{k=1}^{n_{T^\omega}} \text{entropy}(pat_{T^\omega}^k) \quad (6)$$

در این رابطه  $n_{T^\omega}$  نمایانگر تعداد الگوهای است که با ابعاد پنجره مشخص  $\omega \times \omega$  از تصویر آموزشی استخراج می‌شود. با توجه به ابعاد انتخاب شده در هر مرحله، تمام الگوهای موجود در تصویر آموزشی استخراج می‌شود. مقادیر آنتروپی متوسط به دست آمده ( $ME(\omega)$ ) با توجه به ابعاد تصویر  $\omega$  رسم می‌شوند. تغییرات آنتروپی متوسط باید در ابتدا به شدت افزایش یابد و وقتی به ابعاد پنجره بهینه رسید شروع به هموار شدن می‌نماید. نمودار آنتروپی متوسط برای تصویر آموزشی شکل ۴-الف در شکل ۴-ب رسم شده است. در این شکل، محور افقی ابعاد پنجره و محور عمودی آنتروپی متوسط را نشان می‌دهد. با توجه به این نمودار، ابعاد مناسب برای پنجره مربوط به این تصویر آموزشی ۲۵ در ۲۵ پیکسل تخمین زده می‌شود. تعداد الگوهای خارج شده از تصویر آموزشی با ابعاد به دست آمده را می‌توان با رابطه ۷ محاسبه کرد.

$$N = (TI_x - s) \times (TI_y - s) \quad (7)$$

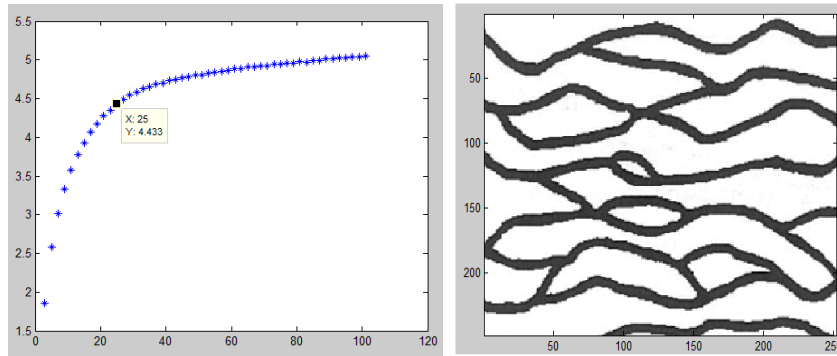
که در این رابطه،  $TI_x$  ابعاد تصویر آموزشی در جهت  $x$ ،  $TI_y$  ابعاد تصویر آموزشی در جهت  $y$  و  $s$  ابعاد انتخاب شده برای پنجره است. عدد ۲۵ برای ابعاد پنجره در این تصویر آموزشی در واقع بازوی منحنی بوده و نقطه‌ای است که منحنی از شیب زیاد به شیب کم تغییر حالت می‌دهد. در تمام بررسی‌های انجام شده در این مقاله، برای انتخاب ابعاد مناسب برای پنجره، از این روش استفاده شده است.

اندازه پنجره باید متناسب با ساختارهای با اندازه کوچک در قدرت تفکیک تصویر آموزشی و الگوهای واقعی طراحی شود. اگر اندازه پنجره با توجه به الگوهای واقعی خیلی کوچک انتخاب شود یا به عبارتی، مشخصات واقعی برای انتخاب پنجره کافی نباشد، آمارگان برای هر پنجره به شدت تغییر می‌کند. بنابراین ابعاد پنجره باید تا حدی کوچک انتخاب شود که بتواند ساختارهای کوچک را بازیابی نماید و در عین حال باید به اندازه کافی نیز بزرگ باشد تا مشخصات و الگوهای واقعی که در تصویر آموزشی اتفاق می‌افتند را بازتولید کند. در این مقاله برای پیدا کردن ابعاد بهینه پنجره، از آنتروپی یک بعدی استفاده شده است. آنتروپی، یک اندازه‌گیری آماری از میزان تصادفی بودن است که می‌تواند برای توصیف بافت تصویر آموزشی استفاده شود. آنتروپی یک الگو با ابعاد  $n_x \times n_y \times n_z$  می‌تواند طبق رابطه ۴ محاسبه شود [۱۵].

$$H = \sum_{i=1}^K p_i \log(p_i) \quad (4)$$

به طوری که  $k$  تعداد نتایج ممکن متغیر تصادفی و  $p_i$  نشانگر تابع چگالی احتمال (معادل هیستوگرام) آن است. با افزایش ابعاد پنجره در یک تصویر آموزشی ایستا، دو رفتار متفاوت در منحنی تغییرات آنتروپی مشاهده می‌شود. در مرحله اول، آنتروپی به شدت افزایش می‌یابد. در مرحله بعد، وقتی که ابعاد پنجره در حال افزایش و بیشتر شدن از ابعاد پنجره بهینه است، که مشخصات ایستایی تصویر آموزشی را نشان می‌دهد، آنتروپی با یک آهنگ خیلی ملایم افزایش می‌یابد. در واقع با افزایش ابعاد پنجره در یک تصویر آموزشی ایستا، تعداد پیکسل‌های موجود در پنجره افزایش می‌یابد و این افزایش موجب افزایش آنتروپی می‌شود. با حاصل شدن ابعاد خاصی از پنجره، آنتروپی شروع به ثابت شدن می‌کند. این مطلب نشان دهنده این است که الگوهای موجود در این پنجره در حال تکرار شدن است و افزایش ابعاد پنجره، تغییری در آمارگان پنجره ایجاد نمی‌کند. بنابراین ثابت شدن آنتروپی به معنای حاصل شدن ابعاد مناسب پنجره است.

برای انتخاب پنجره بهینه، تصویر آموزشی به وسیله یک پنجره با ابعاد مختلف مرور و مقدار آنتروپی پنجره مربوطه محاسبه می‌شود. از آنجایی که معمولاً ابعاد پنجره‌ها در



ب

الف

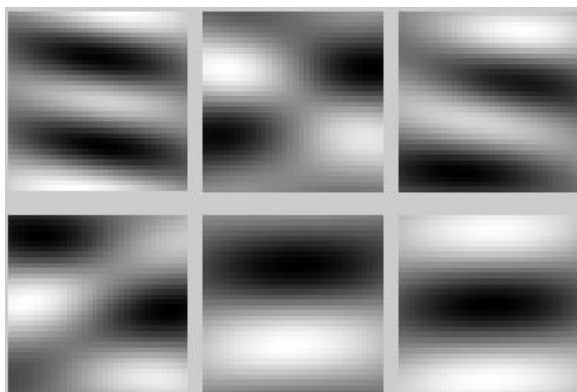
شکل ۴: (الف) تصویر آموزشی دو رخساره ای (باینری) و (ب) منحنی آنتروپی مربوطه نسبت به ابعاد پنجره، جهت انتخاب ابعاد پنجره مناسب

#### ۵- طراحی فیلترهای سازگار با تصویر آموزشی در الگوریتم اصلاح شده

$K$  عدد بردار ویژه متناظر با بزرگترین مقادیر ویژه را انتخاب می کنیم. فیلترهای افقی مورد نظر، در واقع این بردارهای ویژه است. این ۶ فیلتر در واقع ۶ بردار ویژه متناظر با بزرگترین مقادیر ویژه ماتریس کواریانس است و تمام این ۶ فیلتر از هم مستقل بوده و متعامدند. ویژگی های ورودی الگوریتم PCA در اینجا تمام الگوهای استخراج شده از تصویر آموزشی با ابعاد مشخص شده در مرحله قبل است.

برای مقایسه روش ارائه شده با الگوریتم فیلترسیم، با توجه به تعداد فیلترها در این روش، در تمام این مقاله، در تصاویر آموزشی دو بعدی  $k = 6$  انتخاب شده است. البته در مراحل بعد می توان این مقادیر را تغییر داد و نتایج را با یک دیگر مقایسه کرد.

مثلا اگر تصویر آموزشی به صورت شکل ۴-الف باشد فیلترهای مربوطه به فرم شکل ۵ به دست می آیند.



شکل ۵: فیلترهای طراحی شده به روش PCA برای تصویر آموزشی شکل ۴-الف

در الگوریتم فیلترسیم تعدادی فیلتر پیش فرض (۶ فیلتر در فضای دو بعدی و ۹ فیلتر در فضای سه بعدی) به کار گرفته شده است که ارتباطی با ساختار تصویر آموزشی نداشته و تنها جهت کاهش بعد محاسباتی استفاده می شوند. در این مقاله با استفاده از روش PCA و با توجه به نوع تصویر آموزشی، فیلترهای مناسب و سازگار با تصویر آموزشی استخراج می شوند. مراحل انجام کار به صورت زیر است. ابتدا با توجه به تصویر آموزشی یک پنجره جستجوی مناسب انتخاب می شود. پنجره جستجو، برای تصاویر آموزشی سه بعدی به صورت مکعب مستطیل و برای تصاویر آموزشی دو بعدی، مربع است. ابعاد مناسب برای پنجره با توجه به روش گفته شده در قسمت ۴ این مقاله محاسبه می شود.

با توجه به پنجره جستجو، تمام الگوهای موجود در تصویر آموزشی استخراج می شود. بدین منظور پنجره جستجو را پیکسل به پیکسل بر روی تصویر آموزشی جلو برده و الگوهای به دست آمده را در کتابخانه الگوها ذخیره می کنیم. یک نمونه از کتابخانه الگوها برای یک تصویر آموزشی نمونه در شکل ۲ نشان داده شده است.

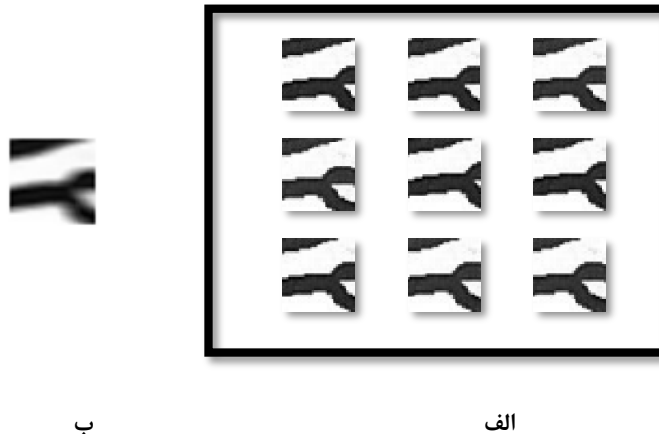
با توجه به الگوهای به دست آمده، ماتریس کواریانس، بردار میانگین الگوها، مقادیر ویژه و بردارهای ویژه ماتریس کواریانس را محاسبه کرده و با توجه به مقادیر ویژه، به صورت نزولی مرتب می شود.

## ۶- خوشه‌بندی الگوها در فضای امتیاز فیلتری

خوشه تا حد ممکن بوده و معیار جدایش که میزان دوری تا حد ممکن خوشه‌ها از هم کنترل می‌نماید، می‌باشند. شاخص کیفیت از طریق جمع وزن دار جدایش کلی بین خوشه‌ها و میانگین پراکندگی خوشه‌ها تعریف می‌شود. در این پژوهش بهترین نتایج خوشه‌بندی با استفاده از شاخص گپ (gap) و با استفاده از نرم‌افزار متلب (دستور evalclusters) حاصل شد [۱۷]. در این روش تعداد بهینه خوشه‌ها در حداکثر میزان شاخص گپ اتفاق می‌افتد. با این روش ساختار واقعی خوشه‌بندی داده‌های مختلف به خوبی بازیابی می‌شود.

طراحی فیلترهای وفقی شده با تصویر آموزشی، نتایج بهتری نسبت به استفاده از فیلترهای ثابت دارد و مخصوصاً در مواردی که الگوهای موجود در تصویر آموزشی دارای پیچیدگی بیشتری باشند استفاده از این روش کارا تر است. بعد از اجرای الگوریتم k-means در روش اصلاح شده، الگوهای دارای همبستگی بیشتر در یک دسته قرار خواهند گرفت. شکل ۶- الف الگوهای قرار گرفته در یک دسته را به عنوان نمونه نشان داده است. در این دسته فقط ۹ الگو قرار گرفته که میانگین این الگوها نیز در شکل ۶- ب نشان داده شده است.

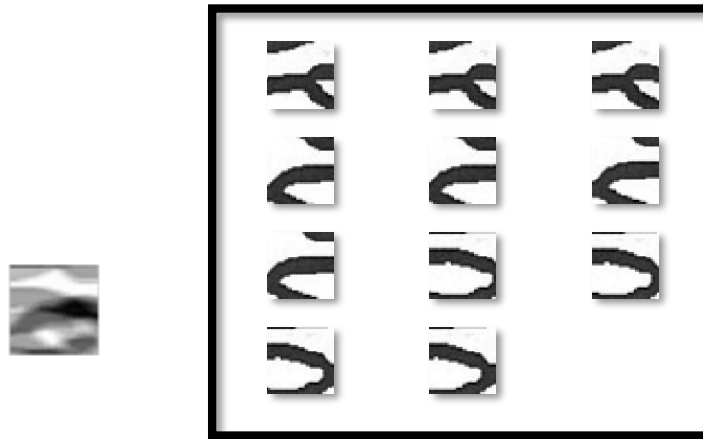
در این تحقیق برای خوشه‌بندی الگوها در فضای امتیاز فیلتری از الگوریتم k-means که یک الگوریتم کلاسیک و شناخته شده خوشه‌بندی است، استفاده شده است. الگوریتم خوشه‌بندی k-means از نوع توده‌ای بوده و معمولاً خوشه‌ها را به صورت همسانگرد توزیع می‌کند. عملکرد این الگوریتم بر اساس فواصل بین کلاسی و درون کلاسی استوار است. تعداد خوشه‌ها بسته به نوع تصویر آموزشی متفاوت است و در پژوهش‌های گذشته از تعداد ثابت و روش‌های ترسیم نمودار به صورت سعی و خطا و بسته به پیچیدگی الگوهای موجود در تصویر آموزشی استفاده شده است. در این پژوهش علاوه بر روش‌های ترسیم نمودار از روش‌های شاخص کیفیت خوشه بندی جهت تعیین تعداد بهینه خوشه‌ها نیز استفاده شده است. روش‌های شاخص کیفیت خوشه‌بندی استاندارد و کلاسیک ارائه شده برای خوشه‌بندی‌های غیر پوششی سخت (Crisp Classifier) که بر اساس فشردگی (Compactness) و جدایش (Separation) تعریف می‌شوند، ابزار مناسبی برای تعیین تعداد بهینه خوشه‌ها هستند [۱۸]. مهم‌ترین معیارهای به کار گرفته شده در این شاخص‌ها شامل معیار فشردگی که مؤید مشابهت اعضای هر



شکل ۶: (الف) الگوهای قرار گرفته در یک دسته به عنوان نمونه در روش فیلترسیم اصلاح شده و (ب) میانگین الگوهای قرار گرفته در دسته الف

که ۱۱ الگو قرار گرفته است، الگوها به صورت شکل ۷- الف و میانگین آنها به صورت شکل ۷- ب است.

اما در روش فیلترسیم کلاسیک، در برخی موارد، چند الگوی ناسازگار در یک دسته قرار می‌گیرد. مثلاً در یک دسته



الف

ب

شکل ۷: (الف) الگوهای قرار گرفته در یک دسته به عنوان نمونه در روش فیلترسیم کلاسیک (ب) و میانگین الگوهای قرار گرفته در دسته الف

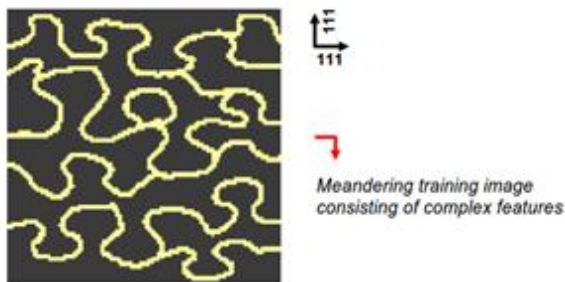
با توجه به اینکه میانگین دسته ها بسیار مهم بوده و در نهایت تمام تصاویر با میانگین دسته ها مقایسه می شوند، عدم دقت این میانگین دسته ها باعث عدم دقت در اجرای این الگوریتم می شود که این مطلب همانگونه در اجرای الگوریتم در شکل ۱۲ نشان داده شده، در الگوریتم پیشنهادی به خوبی اصلاح شده است. قابل ذکر است که در این مقاله تمام محاسبات مربوط به فاصله بین الگوها و انتخاب الگوی مناسب، در فضای امتیاز فیلتری انجام شده و این قضیه باعث انجام سریع الگوریتم ها می شود.

#### ۷- اجرای الگوریتم پیشنهادی و مقایسه نتایج

این قسمت از پنجره در مراحل بعدی شبیه سازی تغییر نکرده و به عنوان داده سخت به حساب می آید.

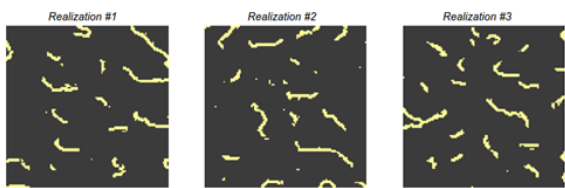
جهت اطمینان از اجرای بهتر الگوریتم پیشنهادی حتی با کم کردن تعداد کلاس های در نظر گرفته شده برای اجرای الگوریتم، تعداد کلاس ها در روش فیلترسیم اصلی ۲۰۰۰ و در روش فیلترسیم اصلاح شده ۱۰۰۰ در نظر گرفته می شود. کم کردن تعداد کلاس ها سرعت اجرای الگوریتم را افزایش می دهد. همان گونه که مشاهده می شود، با اینکه تعداد کلاس ها در روش فیلترسیم اصلاح شده نصف روش فیلترسیم اصلی است، نتایج به دست آمده در روش اصلاح شده از نظر پیوستگی بسیار بهتر است.

در این قسمت نتایج حاصل از دو الگوریتم فیلترسیم کلاسیک و فیلترسیم اصلاح شده با فیلترها و فقی نشان داده شده و نتایج مقایسه می شوند. ابتدا حالت های تحقق یافته برای هر تصویر آموزشی در هر دو روش نشان داده شده و سپس با معیارهای ارزیابی مناسب، نتایج دو روش با هم مقایسه شده است.



شکل ۸: یک نمونه تصویر آموزشی [۱]

#### a. حالت های تحقق یافته

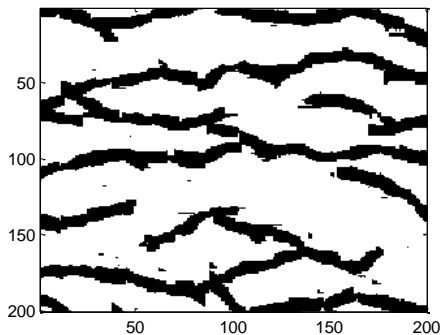
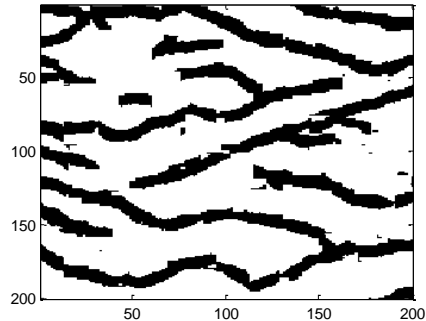


شکل ۹: نتیجه اجرای الگوریتم در روش فیلترسیم کلاسیک با ۲۰۰۰ کلاس برای طبقه بندی [۱]

در شکل ۸ یک تصویر آموزشی و در شکل ۹ و ۱۰ به ترتیب، نتایج حاصل از روش فیلترسیم کلاسیک و فیلترسیم اصلاح شده با فیلترهای و فقی، نشان داده شده است. در این الگوریتم، ابعاد پنجره مربعی ۱۵×۱۵ و پنجره داخلی ۹×۹ است. پنجره داخلی با ابعاد کوچک تر انتخاب می شود و بعد از چسباندن پنجره اصلی در محل تصویر شبیه سازی شده،

$\{q_i\}$  شبیه است [۱۶]. دیورژانس کالیک لیبلر به صورت رابطه ۸ محاسبه می‌شود.

$$D_{KL}(p||q) = \sum_i p_i \log \frac{p_i}{q_i} \quad (۸)$$



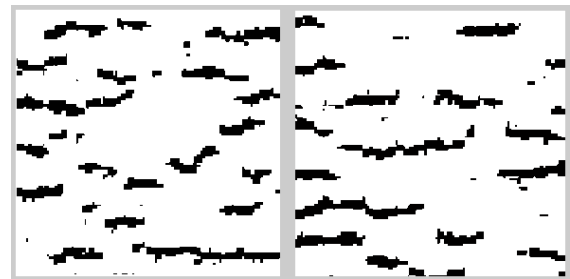
شکل ۱۰: نتیجه اجرای الگوریتم در روش فیلترسیم اصلاح شده با فیلترهای طراحی شده با استفاده از الگوریتم PCA با ۱۰۰۰ کلاس برای طبقه بندی

شکل ۱۱: نتایج حاصل از اجرای غیر شرطی الگوریتم شبیه‌سازی فیلترسیم با استفاده از تصویر آموزشی شکل ۴-الف و شکل ۱۲، شبیه‌سازی غیرشرطی با استفاده از روش ارائه شده در این مقاله با همان تصویر آموزشی را نشان می‌دهد. همان گونه که در این شکل‌ها مشاهده می‌شود، پیوستگی در روش پیشنهادی، بسیار بهتر از الگوریتم فیلترسیم کلاسیک است. در این الگوریتم‌ها، ابعاد پنجره مربعی ۹×۹، پنجره داخلی ۵×۵ و تعداد کلاس‌ها در هر دو روش، ۱۰۰ است.



شکل ۱۲: نتایج مربوط به روش پیشنهادی با تعداد ۱۰۰ کلاس برای تصویر آموزشی شکل ۴-الف

شکل ۱۱: نتایج اجرای الگوریتم فیلترسیم اصلی بر روی تصویر آموزشی شکل ۴-الف [۱]



شکل ۱۱: نتایج اجرای الگوریتم فیلترسیم اصلی بر روی تصویر آموزشی شکل ۴-الف [۱]

## b. ارزیابی و مقایسه نتایج

در این مقاله از دو معیار برای ارزیابی و مقایسه نتایج به دست آمده از پیاده‌سازی غیر شرطی روش پیشنهادی و روش فیلترسیم استفاده شده است.

### ۷-۲-۱ معیار دیورژانس جنسن-شانون<sup>۵</sup>

$$D_{JS}(p||q) = \frac{1}{2}D_{KL}(p||m) + \frac{1}{2}D_{KL}(q||m) \quad (۹)$$

به منظور ارزیابی نتایج به دست آمده، از معیار دیورژانس جنسن-شانون استفاده شده است. این روش، یک معیار رایج برای اندازه‌گیری میزان شباهت دو توزیع تصادفی است و بر اساس دیورژانس کالیک لیبلر<sup>۶</sup> است [۱]. دیورژانس کالیک لیبلر یک معیار اندازه‌گیری در آمار است که به صورت بسیار دقیق و جزئی تعیین می‌کند که یک توزیع احتمال  $p = \{p_i\}$  تا چه اندازه به یک توزیع احتمال مدل (یا مرجع)  $q =$



جدول ۱: دیورژانس جنسن-شانون بین هیستوگرام چند نقطه ای تصویر آموزشی با توجه به هر حالت تحقق یافته

$D_{JS}(hist_{T1}  hist_{Realization})$	
روش اصلاح شده پیشنهادی	روش فیلترسیم کلاسیک
۰/۰۲۲۹	۰/۰۷۰۵ [۱]

مشاهده می شود که حالت های تحقق یافته به دست آمده با استفاده از روش پیشنهادی دارای  $D_{JS}$  کمتر نسبت به روش فیلترسیم کلاسیک هستند. همچنین این نتیجه نشان می دهد که روش پیشنهادی نسبت به فیلترسیم کلاسیک عملکرد بهتری دارد.

### ۷-۲-۲- مقایسه درصد اختلاف آنتروپی در روش های مختلف بر حسب تصویر آموزشی

با مقایسه آنتروپی نتایج به دست آمده از روش پیشنهادی و روش کلاسیک فیلترسیم مشاهده می شود که آنتروپی نتایج به دست آمده با روش پیشنهادی به آنتروپی تصویر آموزشی نزدیکتر است. نتایج کامل مربوط به تصویر آموزشی شکل ۵ در جدول ۲ نشان داده شده است.

جدول ۲: درصد اختلاف آنتروپی در حالت های تحقق یافته با روش های مختلف برای تصویر آموزشی شکل ۵

درصد اختلاف آنتروپی	
روش پیشنهادی	روش فیلترسیم اصلی
۱/۵۶%	۲۵/۲۸ %

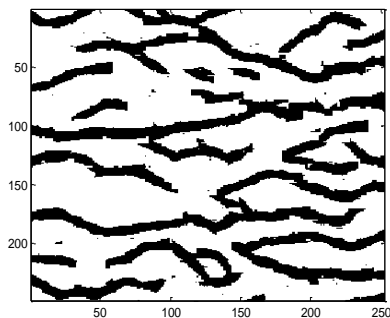
### ۸- شرطی سازی اطلاعات

یکی از مزایای روش آمار چند نقطه ای نسبت به روش های برپایه شی، توانایی شرطی سازی حالت های تحقق یافته با توجه به داده سخت و داده نرم است. این داده ها، اطلاعات کمکی هستند که اگر در کنار تصاویر آموزشی قرار گیرند، عدم قطعیت را در حالت های تحقق یافته ممکن

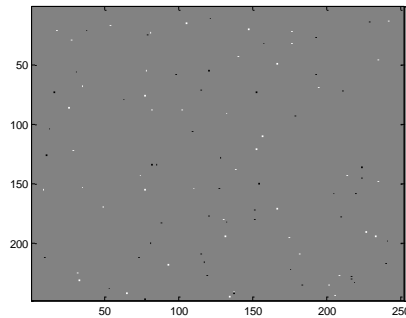
کاهش می دهند [۱]. برای مقایسه روش های مختلف در این قسمت، حالت های تحقق یافته با داده سخت شرطی سازی می شوند. در روش فیلترسیم کلاسیک، شرطی سازی با دادن وزن های متفاوت به پیکسل ها در موقعیت های مختلف، انجام می شود. سپس نزدیک ترین دسته با کمینه کردن تابع شباهت انتخاب می شود [۱۳]. تابع شباهت به صورت رابطه ۱۰ تعریف می شود [۱۳].

$$D_l(u_0) = \sum_{d=1}^3 \omega(d) Ave|data_d(u_0 + u) - prot^{(l)}(u)| \quad (10)$$

در این رابطه،  $d$  نوع داده ( $d = 1$ ) برای داده سخت  $Y$ ،  $d = 2$  برای نقاط شبیه سازی شده که در پنجره داخلی قرار دارند و  $d = 3$  نقاط شبیه سازی شده در قسمت خارجی پنجره داخلی) و  $\omega(d)$  وزن های مربوط به سه نوع داده هستند و برای تاکید بیشتر روی داده سخت و داده های شبیه سازی شده قبلی که در داخل پنجره قرار دارند ضرائب با توجه به رابطه  $\omega(1) > \omega(2) > \omega(3)$  انتخاب می شود.  $data_d(u_0 + u)$  داده ها در تمام نقاط همسایگی  $u_0$  و  $prot^{(l)}(u)$  میانگین هر دسته در فضای امتیاز فیلتری است. نتایج مربوط به این الگوریتم با مقادیر  $\omega(1) = 0.5$ ،  $\omega(2) = 0.3$  و  $\omega(3) = 0.2$  در ادامه که مقادیر پیش فرض در الگوریتم فیلترسیم در نرم افزار SGEMS است [۴]، در ادامه نشان داده شده است. شکل ۱۳-الف یک تصویر با ۱۰۰ داده سخت را نشان می دهد. نقاط مربوط به داده سخت بصورت تصادفی در یک تصویر با ابعاد تصویر آموزشی قرار داده شده است. شکل ۱۳-ب نتیجه اجرای الگوریتم پیشنهادی با روش شرطی سازی ژانگ با ضرائب ثابت را نشان می دهد. همانگونه که در این تصویر نشان داده شده است، تقریباً ده درصد از نقاط داده سخت به صورت منفرد قرار دارند و درست شبیه سازی نشده اند.



ب



الف

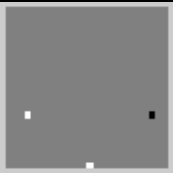


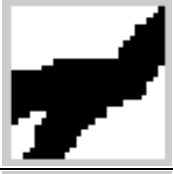


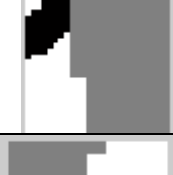
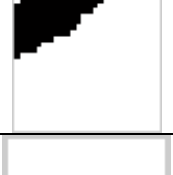

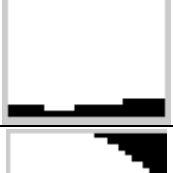
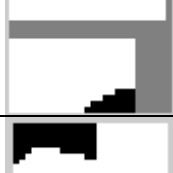
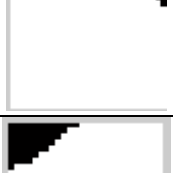


شکل ۱۳: (الف) داده های مشاهده ای  $u_0$  با ۱۰۰ داده سخت و (ب) نتیجه اجرای الگوریتم پیشنهادی با ضرائب ثابت بر روی تصویر الف

a. روش ارائه شده در مقاله برای شرطی سازی داده سخت

مثال، اگر مقادیر پیش فرض نرم افزار SGEMS که به صورت  $\omega(1) = 0.5$ ،  $\omega(2) = 0.3$  و  $\omega(3) = 0.2$  هستند در نظر گرفته شود، نتایج به صورت جدول ۳ هستند.

زنگ، مقادیر  $\omega$  را ثابت فرض کرده بود. یعنی در تمام مراحل اجرای الگوریتم مقادیر  $\omega$  ثابت هستند. به عنوان

جدول ۳: مشخصات مربوط به الگوی انتخاب شده در هر مرحله برای یک داده مشاهده‌ای

الگوی انتخاب شده از داده مشاهده‌ای	تعداد پیکسل داده سخت در الگوی انتخاب شده	درصد وجود نقاط داخلی پنجره‌ها در شبیه‌سازی‌های قبلی در پنجره انتخاب شده از داده مشاهده‌ای	درصد وجود نقاط خارجی پنجره‌ها در شبیه‌سازی‌های قبلی در پنجره انتخاب شده از داده مشاهده‌ای	الگوی انتخاب شده برای چسباندن در داده مشاهده‌ای
	۳	۰	۰	
	۰	٪۷	٪۱۹	
	۰	٪۰	٪۱۸	
	۱	٪۱۱	٪۲۴	
	۰	٪۲۲	٪۲۷	
	۳	٪۳۰	٪۴۶	
	۲	٪۶۲	٪۳۶	

همان گونه که از جدول ۳ فوق مشاهده می‌شود، اگر درصد حضور پیکسل‌های شبیه‌سازی شده قبلی در الگوی

انتخاب شده از داده‌های مشاهده‌ای بیشتر از ۳۰ درصد باشد، ضرائب ثابت ژانگ برای این پیکسل‌ها مناسب نبوده و این

$\Omega_{out}$  و نواحی دارای پیکسل های مجهول  $\bar{\Omega}$  تقسیم کرد، در این صورت روابط نسبی زیر را تعریف می کنیم.

$$R_{h\_d}(T) = \frac{|\Omega_{h\_d}|}{|\Psi_T|} \times 100 \quad (11)$$

$$R_{freeze}(T) = \frac{|\Omega_{frized}|}{|\Psi_T|} \times 100 \quad (12)$$

$$R_{out}(T) = \frac{|\Omega_{out}|}{|\Psi_T|} \times 100 \quad (13)$$

$$\Psi_T = \Omega_{h\_d} + \Omega_{frized} + \Omega_{out} + \bar{\Omega} \quad (14)$$

که در این روابط  $|\cdot|$  تعداد پیکسل ها را نشان می دهد و  $T$ ، پنجره مربوط به داده های مشاهده ای در هر مرحله است. در این صورت باید ضرائب  $\omega_1(T)$ ،  $\omega_2(T)$  و  $\omega_3(T)$  از روی مقادیر  $R_{h\_d}(T)$ ،  $R_{freeze}(T)$  و  $R_{out}(T)$  به صورت روابط (۱۵) تا (۱۷) تعیین شود.

$$\omega_2(T) = -\frac{3}{7}(R_{freeze}(T) - 1) \quad (15)$$

$$\omega_3(T) = -\frac{2}{7}(R_{out}(T) - 1) \quad (16)$$

$$\omega_1(T) = 1 - \omega_2(T) - \omega_3(T) \quad (17)$$

نتایج اعمال این روابط در الگوریتم ها، به صورت منحنی E-type نشان داده شده است. از E-type برای چک کردن کیفیت شرطی سازی استفاده می شود. از ۱۰۰ نقطه شرطی داده سخت برای مقایسه روش ها استفاده می شود. منحنی E-type مربوط به روش فیلترسیم و باینری شده آن به روش اتسو در شکل ۱۵ و منحنی E-type مربوط به روش پیشنهادی و باینری شده آن با استفاده از روش اتسو در شکل ۱۶ نشان داده شده است. در الگوریتم اتسو، ابتدا یک حد آستانه محلی انتخاب می شود. این حد آستانه در واقع مقدار چگالی نرمال در بازه  $[0,1]$  است و عملیات باینری کردن با این حد آستانه انجام می شود. در هر دو روش از ۱۰۰ دسته برای طبقه بندی الگوها استفاده شده است.

همان گونه که در شکل ۱۵ مشخص است در شبیه سازی ساختارها و داده های با تغییرات تدریجی و پیوسته هر چه وزن بیشتری به داده های سخت داده شود نتیجه به سمت اریب شدن به سمت داده های سخت پیش رفته و از پیوستگی کم کرده و مدل های چشم گاوی ایجاد می کند (مانند مدل رنگی با تغییرات پیوسته شکل ۱۵-ب). اما برای مدل های باینری با داده های سخت باینری یا مدل های ناپیوسته مقادیر مطلق داده های سخت از نظر بصری در به نقشه در آوردن ساختارها نقش ندارند (شکل ۱۵-ب).

ضرائب یعنی  $\omega(2)$  و  $\omega(3)$  باید به شدت کم شوند. در واقع با افزایش تعداد پیکسل های شبیه سازی شده، چون تعداد پیکسل های داده سخت محدود است، این افزایش، باعث کم شدن تاثیر داده سخت در انتخاب الگوی نهایی می شود. بنابراین با کاهش ضرائب مربوط به پیکسل های شبیه سازی شده، در واقع تاثیر داده سخت بیشتر می شود.

به عنوان مثال برای شکل ردیف ۵ جدول اگر ضرائب به صورت،  $\omega(1) = 0.5$ ،  $\omega(2) = 0.04$  و  $\omega(3) = 0.03$  انتخاب شوند الگوی انتخاب شده از پایگاه اطلاعاتی به صورت شکل ۱۴-الف خواهد بود که بسیار مناسب تر از الگوی انتخاب شده با ضرائب ثابت ژانگ است. همچنین برای شکل ردیف ۶ جدول اگر  $\omega(1) = 0.5$ ،  $\omega(2) = 0.04$  و  $\omega(3) = 0.03$  انتخاب شوند الگوی انتخاب شده به صورت شکل ۱۴-ب خواهد بود که در هر صورت، نتیجه نسبت به الگوی انتخاب شده در حالت ضرائب ثابت ژانگ بهتر است. بنابراین با تغییر مقادیر  $\omega$  در طول اجرای الگوریتم، الگوی مناسب تری از پایگاه اطلاعاتی انتخاب شده و نتایج با پیوستگی بهتری به دست خواهند آمد.



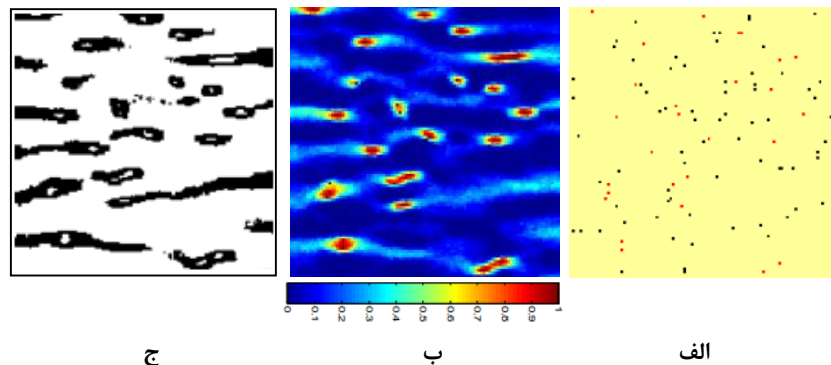
(ب)



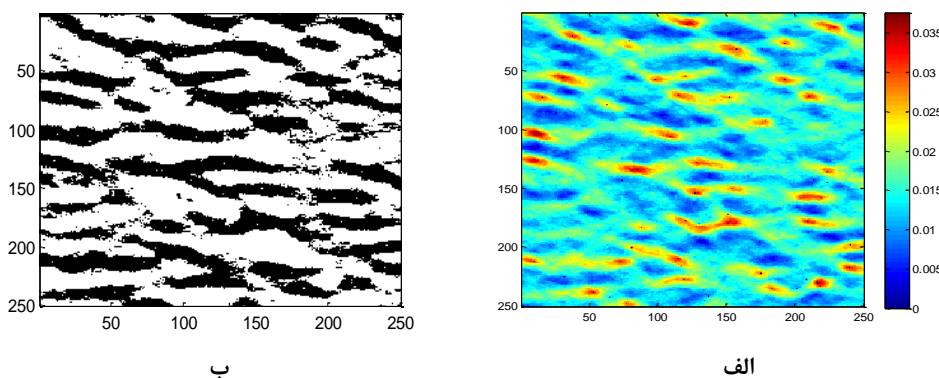
(الف)

شکل ۱۴: الف) الگوی انتخاب شده از پایگاه اطلاعاتی مربوط به تصویر داده های مشاهده ای ردیف ۵ جدول با مقادیر ضرائب مقادیر  $\omega(1) = 0.5$ ،  $\omega(2) = 0.04$  و  $\omega(3) = 0.03$  و (ب) الگوی انتخاب شده از دیتا بیس مربوط به تصویر داده های مشاهده ای ردیف ۷ جدول با مقادیر ضرائب مقادیر  $\omega(1) = 0.5$ ،  $\omega(2) = 0.04$  و  $\omega(3) = 0.03$

بر این اساس، می توان مقادیر  $\omega$  را به طور وفقی و بسته به درصد حضور داده های موجود در پنجره انتخاب کرد. اگر فرض کنیم در  $n$  امین تکرار از الگوریتم شبیه سازی، شبکه شبیه سازی<sup>۹</sup> را بتوان به ۴ ناحیه شامل نواحی دارای پیکسل های معلوم متعلق به داده سخت،  $\Omega_{h\_d}$  نواحی دارای پیکسل های معلوم شبیه سازی شده قبلی که در پنجره داخلی قرار دارند،  $\Omega_{freeze}$  نواحی دارای پیکسل های معلوم شبیه سازی شده که در خارج پنجره داخلی قرار دارند،



شکل ۱۵: (الف) تصویر حاوی ۱۰۰ داده سخت، (ب) منحنی E-type حاصل از ۱۰۰ حالت تحقق یافته و ۱۰۰ داده سخت با روش فیلترسیم کلاسیک (ج) باینری شده تصویر E-type با استفاده از الگوریتم اتسو [۱]



شکل ۱۶: (الف) منحنی E-type حاصل از ۱۰۰ حالت تحقق یافته و ۱۰۰ داده سخت نشان داده شده در شکل ۱۵-الف با روش‌های پیشنهادی و (ب) باینری شده تصویر E-type با استفاده از الگوریتم اتسو

### ۹- تحلیل نتایج و پیشنهادات

پیاده‌سازی آن در نرم‌افزار تجاری شده SGEMS جهت بهبود عملکرد الگوریتم فیلترسیم کلاسیک است. تغییر الگوریتم فیلترسیم به فرم ارائه شده در این مقاله، تعداد پارامترهای ورودی مربوطه در فیلترسیم کلاسیک را تغییر نمی‌دهد و پارامترهای ورودی توسط کاربر همان پارامترهای مربوط به الگوریتم اصلی فیلترسیم کلاسیک هستند.

### مراجع

[1] Honarkhah, M, (2011) Stochastic Simulation of Patterns Using Distance-Based Pattern Modeling, Caers, J., Department of Energy Resources Engineering, Stanford University.

[2] Strebelle, S. (2002). Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics, *Mathematical Geology*, 34(1), 1-21.

[3] Mariethoz G, Renard P, Caers J (2010) Bayesian inverse problem and optimization with iterative spatial resampling. *Water Resour Res* 46.

[4] Zhang, T. (2006). Filter-based Training Pattern classification for spatial Pattern simulation, Switzer, P., The Department of Geological and Environmental Sciences, Stanford University.

همان طور که در بخش‌های قبلی این تحقیق نشان داده شد طراحی فیلترهای جدید سازگار با تصویر آموزشی با روش پیشنهاد شده مبتنی بر مؤلفه‌های اصلی در این مقاله، عملکرد الگوریتم فیلترسیم کلاسیک را در بازیابی ساختارهای غیرخطی و پیوستگی آنها در حالات تحقق یافته غیرشرطی بدون افزایش بار محاسباتی، بهبود بخشید. همچنین به منظور قاعده‌مند نمودن شرطی‌سازی داده‌های سخت از تعریف رابطه‌هایی که سهم هریک از سه نوع داده موجود در پنجره الگوی شبیه‌سازی مشخص می‌نمایند، استفاده شد. نتایج اعمال روش جدید شرطی‌سازی نظارت شده حاکی از عملکرد بهتر این روش در مقایسه با روش فیلترسیم کلاسیک است. ضمناً در مواردی که تصویر آموزشی پیچیدگی زیادی داشته باشد می‌توان به جای استفاده از ۶ فیلتر، از تعداد بیشتری فیلتر استفاده کرده و عملکرد بهتری را در بازیابی ساختارهای کوچک مقیاس به دست آورد. از جمله مزیت‌های روش پیشنهادی قابلیت

- [12] Burc, G., Caers, J. (2004), A Multiple-scale, Pattern-based Approach to Sequential Simulation, Springer, [Quantitative Geology and Geostatistics](#), 14(1), 255-264.
- [13] Zhang, T., Switzer, P., Journel, A., (2006), Filter-Based Classification of Training Image Pattern for Spatial Simulation, *Mathematical Geology*, 38(1), 63-80.
- [14] Duda, R., Hart, P., Seork, D., (2001). Pattern Classification, the Electrical Engineering Department at San Jose State University, San Jose, California.
- [15] Shannon, C.E., (1948) A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(1) 379-423.
- [16] Cover TM., Thomas JA., (1991). Elements of information theory. Wiley, New York.
- [17] Tibshirani, R., G. Walther, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B*. Vol. 63, Part 2, 2001, pp. 411-423.
- [18] Maimon O. and Rokach L., (Ed). 2010, Data Mining and Knowledge Discovery Handbook, Second Edition, Springer
- [5] Wua, J., Boucherb, A., Zhang, T. (2008), A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM, *Computers & Geosciences*, 34(1), 1863-1876,
- [6] Wua, J., Zhang, T., Journel, A. (2008), Fast FILTERSIM Simulation with Score-based Distance, *Math Geosci*, 40(1), 773-788.
- [7] Mariethoz, G., Caers, J. (2015), Multiple-point geostatistics, John Wiley & Sons, Ltd, USA.
- [8] Carvalho PRM, Costa JFCL, Rasera LG, Varella LES (2016) Geostatistical facies simulation with geometric patterns of a petroleum reservoir. *Stoch Environ Res Risk Assess*:1-18. doi: 10.1007/s00477-016-1243-5
- [9] Zhang T, Du Y, Li B, Zhang A (2017) Stochastic reconstruction of spatial data using LLE and MPS. *Stoch Environ Res Risk Assess*:1-14.
- [10] Zhang T, Du Y, Huang T, Yang J, Lu F, Li X (2016) Reconstruction of porous media using ISOMAP-based MPS. *Stoch Environ Res Risk Assess*. 30(1):395-412.
- [11] Wu, J. (2007), 4D Seismic and Multiple-Point Pattern Data Integration Using Geostatistics, Journel, A.G. Department of Energy Resources Engineering, Stanford University.

<sup>7</sup> hard data

<sup>8</sup> Data event

<sup>9</sup> simulation grid

<sup>1</sup> Multiple-point (MP)

<sup>2</sup> Single Normal Equation Simulation

<sup>3</sup> Principle Component Analysis (PCA)

<sup>4</sup> Prototype

<sup>5</sup> Jensen Shannon (JS)

<sup>6</sup> Kullback-Leibler