



تقریب مدل تراوایی مخزن با استفاده از روش های تعلیم لغت نامه ILS-DLA و کدگذاری تنک LARS

محمد حسینی^۱ و محمدعلی ریاحی^{۲*}

۱- دانشجوی دکتری، مؤسسه ژئوفیزیک، دانشگاه تهران

۲- استاد، مؤسسه ژئوفیزیک، دانشگاه تهران

دریافت مقاله: ۱۳۹۶/۱۲/۱۶؛ پذیرش مقاله: ۱۳۹۷/۰۹/۱۸

* نویسنده مسئول مکاتبات: mariahi@ut.ac.ir

چکیده

در این مقاله از روش های کدگذاری تنک و تعلیم لغت نامه به منظور ایجاد مدل تقریب تنک از یک مدل اصلی استفاده شده است. ایجاد مدل تقریب تنک از یک مدل اصلی شامل دو مرحله است؛ کدگذاری تنک و تعلیم لغت نامه. در کدگذاری تنک در هر تکرار با انتخاب مناسب ترین متغیر و اضافه نمودن آن به مجموعه متغیرهای بکار برده شده، مدل رگرسیون بهینه می شود. لغت نامه در واقع ماتریسی است که بر اساس آن، مدل ها کدگذاری می شوند. برای ایجاد ماتریس لغت نامه مجموعه ای از مدل های قابل قبول از یک خاصیت را به عنوان مجموعه تعلیم در نظر می گیرند و مراحل کدگذاری تنک و تعلیم لغت نامه به صورت متناوب روی این مجموعه تعلیم انجام می گیرد. در هر تناوب لغت نامه بهینه تر می شود. مجموعه تعلیم را می توان با روش های زمین آمار چند نقطه ای و بر اساس بهترین مدل های موجود از آن خاصیت ایجاد کرد. در این مقاله، از روش زمین آمار چند نقطه ای DisPat برای ایجاد تحقیق های مورد نیاز در مجموعه آموزشی استفاده شده است. روش ILS-DLA به منظور تعلیم لغت نامه و روش LARS به منظور کدگذاری تنک به کار برده شده اند. روش ILS-DLA فرآیند وارون سازی در تعلیم لغت نامه را بر اساس ساختار داخلی لغت نامه و با بلوک کردن لغت نامه به زیر-ماتریس های کوچکتر انجام می دهد. اساس روش LARS همان رگرسیون خطی است ولی سعی می کند به جای استفاده از همه متغیرها، تعداد مناسبی از آنها را انتخاب و در رگرسیون استفاده کند. بر اساس آزمایش های انجام شده مدل تقریب تنک نهایی به احتمال ۹۰ درصد نسبت به ۹۰ درصد از مدل های موجود در مجموعه آموزشی به مدل اصلی نزدیک تر است. از مدل تقریب تنک می توان در مدل سازی و شناسایی خواص مخزن به عنوان نماینده ای از مدل اصلی و به عنوان حد پایین در انتخاب مجموعه مدل های نزدیک به مدل اصلی استفاده کرد.

واژگان کلیدی

تقریب تنک
تعلیم لغت نامه
کدگذاری تنک
زمین آمار چند نقطه ای
شناسایی خواص مخزن
وارون سازی لرزه ای
شبیه سازی مخزن

۱- مقدمه

در این مقاله سعی شده است پتانسیل استفاده از روش های تقریب تنک در شناسایی و مدل سازی مخزن مورد بررسی قرار گیرد. استفاده از روش های تقریب تنک برای پردازش تصویر طی دهه اخیر ایجاد و گسترش داده شده است (Aharon et al., 2006a and 2006b; Horev et al., 2012; Mairal et al., 2009; Rubinstein et al., 2008, 2010a, 2010b, and 2013; Skretting and Engan, 2010; Skretting and Husøy, 2003; Skretting et al., 1999). در پردازش تصویر، از تقریب تنک به منظور فشردن سازی تصویر استفاده می گردد. بدین ترتیب که با داشتن یک ماتریس مرجع (یا همان لغت نامه) می توان یک ترکیب خطی خاص با ضرایب وزنی مشخص با هدف ارائه فشردن یک تصویر پیدا کرد و به جای ارائه و انتقال تصویر، می توان تنها از آدرس ستون های (یا همان اتم ها) استفاده شده از ماتریس مورد نظر و ضرایب متناظر با آنها استفاده کرد.

در علوم زمین، تنها یک مدل واقعی برای یک خصوصیت خاص از زمین وجود دارد؛ هرچند ممکن است تعداد زیادی مدل از آن خاصیت زمین همزمان صادق باشند. به این مفهوم در علوم زمین و مدل سازی مخزن، عدم قطعیت گفته می شود (Mariethoz and Caers, 2015). در علوم زمین و مهندسی نفت آنچه بیشتر مطرح است، یافتن و ایجاد مدل های بیشتر از یک خصوصیت زمین و انتخاب بهترین مدل از میان تعداد محدود مدل های موجود است. انتخاب بهترین مدل بر اساس آزمایش های بیشتر و مشاهده نتایج صورت می پذیرد. هیچ تضمینی وجود ندارد که مدل انتخاب شده از میان یک مجموعه مدل، همان مدل واقعی باشد؛ یا بهترین مدلی باشد که در عمل قابل دستیابی است. بدیهی است که هر چه فضای نمونه ای بزرگتر باشد، مدل نهایی انتخاب شده ترجمانی نزدیک تر به مدل واقعی زمین خواهد بود.

۱-۱- تکرر مدل ها با استفاده از زمین آمار چند نقطه ای

در شناسایی و مدل سازی خواص مخزن، ایجاد یک مجموعه بزرگ از فضای نمونه ای بدون تعریف خاصیت تصادفی در مدل ها با روش های زمین آمار چند نقطه ای عملاً غیر ممکن است. این مساله به خوبی مشخص است که روش های زمین آمار دو نقطه ای (Deutsch and Journel, 1998; Strebelle, 2000) نمی توانند به صورت مناسبی مدل های تصادفی از مخزن ایجاد کنند. روش های زمین آمار دو نقطه ای برای مدل کردن تغییرات بین دو نقطه از مخزن بر اساس فاصله مکانی از واریوگرام استفاده می کنند. برای فواصل دور، خصوصاً اگر تغییرات خواص زمین ناگهانی یا ناپیوسته باشد، واریوگرام قادر به مدل سازی مناسب آن خاصیت نخواهد بود.

به عنوان یک راه پیشرفته تر، می توان از یک مدل مستقل از مکان به عنوان منبع اطلاعات و تغییرات آماری در روش زمین آمار چند نقطه ای استفاده کرد و مدل های ایستایی قابل اعتمادی برای مخزن مورد مطالعه آموزش داد و ایجاد کرد. به این مدل مستقل از مکان تصویر آموزشی

گفته می شود و خاصیت استقلال از مکان آن باعث می شود تا بتوان مدل های مخزنی تصادفی مختلفی ایجاد نمود (Haldorsen and Damsleth, 1990; Mariethoz and Caers, 2015). اگر اطلاعات اولیه دقیق و گسترده ای درباره مخزن در دسترس باشد، تصاویر آموزشی واقعی تر و مدل های تصادفی قابل اعتمادتری تولید خواهند شد. به طور کلی بهترین روش ها برای مدل سازی مخزن روش های زمین آمار چند نقطه ای الگو-اساس می باشند (Honarkhah, 2011). الگوریتم DisPat به دلیل توانایی در ادغام داده ها بر اساس سیستم بینایی انسان و ارائه الگوریتم های جدید برای مدل سازی تصاویر ناپیوسته یک الگوریتم قوی محسوب می شود.

۱-۲- تقریب تنک به عنوان یک ابزار ریاضی برای دستیابی

به مدل فشردن تنک

روش های زمین آمار چند نقطه ای بر اساس یک تصویر آموزشی مجموعه ای از مدل ها ایجاد می کند، که هر مدل ایجاد شده به نوبه خود ترجمانی از یک مدل واقعی نامشخص است. مساله مورد پرسش این است که چگونه می توان بر اساس مجموعه ای از مدل های ایجاد شده به وسیله الگوریتم MPS، تنها یک مدل به عنوان ارائه ای از مدل حقیقی بدست آورد. ابزار ریاضی بدست آوردن چنین مدلی، الگوریتم تقریب تنک است. اصطلاح تقریب بر این مساله تاکید می کند که بدست آوردن دقیق مدل حقیقی غیر ممکن یا در بیشتر مواقع غیر ضروری است.

مجموعه بزرگی از تصاویر مدل های تصادفی مخزن ایجاد شده با روش زمین آمار چند نقطه ای را می توان به عنوان مجموعه ای از تصاویر آموزشی برای تعلیم لغت نامه و به منظور فشردن سازی تصویر به کار برد (Bryt and Elad, 2008; Elad and Aharon, 2006; Mairal et al., 2008; Skretting and Engan, 2011a and 2011b). به طور کلی دو نوع لغت نامه وجود دارد، که می تواند به منظور فشردن سازی یک دسته از تصاویر مدل بکار روند: لغت نامه های ضمنی و لغت نامه های صریح. لغت نامه های ضمنی لغت نامه هایی هستند که عناصر آنها ثابت و معمولاً با معادلات تحلیلی ارائه شده است. لغت نامه های صریح از این لحاظ با لغت نامه های ضمنی متفاوت هستند؛ که به وسیله مجموعه خاصی از تصاویر آموزشی تعلیم و انطباق داده شده است. بنابراین، لغت نامه های ضمنی، عمومیت بیشتری برای حل مسائل دارند؛ ولی کم اثرتر هستند. هرچند لغت نامه های صریح برای مجموعه خاصی از تصاویر تعلیم داده شده اند و روی همان مجموعه یا مجموعه های آموزشی مشابه موثرترند. اساساً فرآیند تقریب تنک با لغت نامه های صریح انجام می شود؛ که روی مجموعه های خاصی از نمونه های آموزشی (تصویر یا سیگنال) تعلیم داده و منطبق می شود. تقریب تنک در واقع یک فرآیند تعویض متوالی بین دو مرحله کدگذاری تنک و بروزرسانی لغت نامه است. در کدگذاری تنک، بهترین بردار (یا مدل) از مجموعه ای بزرگ از بردارها (یا مدل ها) انتخاب می شود؛ به طوری که یک مساله بهینه سازی تحت یک محدودیت تنک کمینه می شود.

۸۹٫۵۸٪ آزمایش‌ها در محدوده ۹۰٪-۹۰٪ (۱۰٪ بالایی)، و در ۹۵٫۸٪ آزمایش‌ها در محدوده ۸۵٪-۸۵٪ (۱۵٪ بالایی) قرار می‌گیرد. این نتایج وقتی قابل جالب‌تر خواهد بود که یادآوری شود مدل‌های ایجاد شده با روش‌های زمین‌آمار چند نقطه‌ای به صورت کاملاً تصادفی ایجاد شده و با داده‌های سخت یا نرم مشروط نشده‌اند؛ ضمن این که پارامترهای دخیل در فرآیندهای تقریب تنک را می‌توان به گونه‌ای بهینه‌سازی کرد که پاسخ‌های مناسب تری نیز بدست آید. مشاهدات ارائه شده در این مقاله راه را برای کاربردهای بیشتر تقریب تنک در مدل‌سازی و تعیین خواص سیستم، هم در زمینه شبیه‌سازی مخزن و هم در سایر زمینه‌های مرتبط با علوم زمین، باز می‌کند.

۱-۴- ساختار مقاله

در این مقاله نخست بلوک‌های اساسی فلوچارت مدل‌سازی و شناسایی مخزن فشرده تنک ارائه می‌شود. این بلوک‌های اساسی شامل روش زمین‌آمار چند نقطه‌ای DisPat و الگوریتم تقریب تنک است. الگوریتم تقریب تنک شامل مراحل کدگذاری تنک و تعلیم لغت نامه است. برای مرحله کدگذاری تنک، الگوریتم LARS معرفی شده است و برای مرحله به‌روزرسانی لغت نامه، الگوریتم ILS-DLA ارائه شده است. فلوچارت فشرده‌سازی تنک روی یک مورد بررسی شده و نتایج آن در پایان ارائه شده است. مقاله با برخی نتایج پایان می‌یابد.

۲- روش‌ها

هدف اصلی در این مقاله بررسی روش‌های تقریب تنک و استفاده از آنها در زمینه مدل‌سازی و شناسایی مخزن است. تقریب پراکنده خود شامل روش تعلیم لغت‌نامه و روش کدگذاری است. به طور طبیعی تعلیم لغت-نامه بر اساس مجموعه‌ی آموزشی انجام می‌شود. از این رو، در راستای اعمال روش‌های تقریب پراکنده به منظور مدل‌سازی و شناسایی مخزن، ایجاد مجموعه‌ای آموزشی از مدل‌های مختلف مربوط به خاصیت مورد نظر، امری گریزناپذیر است. بر این اساس، روش DisPat به منظور ایجاد تعداد زیادی مدل از خاصیت مورد نظر (تراوایی)، بکار رفته است. پس از ایجاد این مجموعه آموزشی، روش تعلیم لغت نامه ILS-DLA روی این مجموعه اعمال می‌شود. روش تعلیم لغت نامه در واقع یک روش بهینه‌سازی است. هدف از تعلیم لغت نامه بدست آوردن ماتریسی است که بتوان یک تصویر را بر اساس ترکیبی خطی از ستون‌های آن نوشت. این مساله که کدام ستون‌های ماتریس لغت نامه و با چه وزنی در ترکیب خطی حضور خواهند داشت، به وسیله روش‌های کدگذاری LARS مشخص می‌شود. تعلیم لغت نامه خود یک فرآیند تکرار شونده است؛ که طی آن ابتدا یک ماتریس لغت نامه اولیه در نظر گرفته می‌شود؛ سپس یک تصویر تصادفی از مجموعه آموزشی انتخاب می‌شود و بر اساس آن با روش کدگذاری LARS، ترکیب خطی مربوط به تصویر مورد نظر

۱-۳- مطالعه موردی

این تحقیق بر پایه داده‌های واقعی از یک مخزن نفتی در جنوب غرب ایران استوار است و از داده‌های نمودارهای درون چاهی و داده‌های لرزه‌ای به عنوان مبنای تحقیق استفاده شده است. بر این اساس، هدف اولیه این تحقیق استخراج اطلاعات لازم و قابل اعتماد از داده‌های نمودارهای درون چاهی و لرزه‌ای بوده است. از آنجا که روش‌های تقریب تنک بر اساس یک مجموعه آموزشی استوار است؛ هدف ثانویه تحقیق ایجاد مدل‌های متعدد از یک خاصیت مخزنی، تراوایی، بوده است. هدف نهایی این تحقیق، ایجاد تقریبی از مدل خاصیت مورد نظر از مخزن، تراوایی، است. در این مقاله، فلوچارت مدل‌سازی و شناسایی مخزن فشرده تنک با استفاده از روش تعلیم لغت‌نامه ILS-DLA به همراه روش کدگذاری تنک LARS توضیح داده شده و روی یک مخزن نفتی در جنوب غربی ایران اعمال شده است. در مخزن آسماری این میدان به منظور تامین فشار و حفظ ضریب بازیافت، گاز تزریق می‌شود. از لحاظ زمین‌شناسی، مخزن آسماری در این میدان به دلیل وجود میان لایه‌های ماسه و کربنات، احتمال وجود شبکه شکستگی که مخزن را در طول راس آن به هم وصل می‌نماید و حضور سیستم رسوبی دلتایی در ناحیه غرب مخزن، به عنوان یک مخزن خیلی پیچیده تعریف شده است. مدل‌سازی چنین مخزن پیچیده‌ای با روش‌های زمین‌آمار دو نقطه‌ای به طور کامل و به خوبی قابل انجام نیست و به روش‌های تصادفی زمین‌آمار چند نقطه‌ای نیازمند است.

داده‌های لرزه‌ای، نمودارهای درون چاهی، و داده‌های مخزنی برای ایجاد مدل‌های تصویری نیمه صنعتی مخزن مورد استفاده قرار گرفته‌اند و روش زمین‌آمار چند نقطه‌ای DisPat به منظور ایجاد تصاویر مدل‌های چندگانه و تولید مجموعه بزرگی از تصاویر مدل‌های تصادفی مخزن برای دو مورد سیستم شکستگی و دلتایی بکار برده شده است. فرآیند تزریق گاز روی دو مدل دو بعدی استخراج شده از روی بخش‌های مربوط به سیستم شکستگی و بخش دلتایی اعمال شده است. مجموعه‌ای از مدل‌های تصادفی بر اساس تصاویر آموزشی برای هر دو بخش شکستگی با فرض ایستایی و بخش دلتایی با فرض نالیستایی تولید شده است. برای آموزش لغت نامه روی هر کدام از مجموعه‌های آموزشی از روش ILS-DLA استفاده و لغت نامه آموزش داده شده نهایی در یک فرآیند فشرده‌سازی تصویر تنک با استفاده از روش LARS بکار برده شده است. راه مناسب برای مقایسه کیفیت مدل فشرده تنک با مدل‌های تصادفی ایجاد شده از طریق روش زمین‌آمار چند نقطه‌ای این است که برای هر مدل فرآیند شبیه‌سازی مخزن انجام شود و پروفیل‌های تولید و فشار برای هر مدل با مدل اصلی مقایسه گردد.

نتایج آزمایش‌ها نشانگر آن است که مدل تصاویر فشرده تنک، تقریباً در همه موارد و با آزمودن تمام روش‌ها، نسبت به بیشتر مدل‌های تصادفی ایجاد شده از طریق زمین‌آمار چند نقطه‌ای بهتر است؛ به طوری که در

بدست می آید. اطلاعات ترکیب خطی بدست آمده، شامل شماره ستون-های استفاده شده در ترکیب خطی و وزن های متناظر با آنها، مجددا در فرآیند کدگذاری تنگ LARS بکار برده می شوند؛ تا ماتریس لغت نامه جدید و مناسب تری بدست آید. در مرحله بعد، از ماتریس لغت نامه جدید استفاده می شود و مراحل بیان شده برای یک تصویر تصادفی جدید از مجموعه آموزشی مجددا اجرا می شود. این روند تکرار شونده متناوب به قدری ادامه می یابد که معیار در نظر گرفته شده برای توقف تکرار حاصل آید. حاصل این فرآیند ماتریسی است که بر اساس مجموعه آموزشی خاصی از مدل های تعلیم داده شده است. این ماتریس لغت نامه تعلیم داده شده سپس به منظور کدگذاری هر مدل مورد نظر بکار برده می شود.

در این تحقیق، دو مدل تراوایی مربوط به یکی از مخازن نفتی جنوب غرب ایران در نظر گرفته شد و با روش DisPat دو مجموعه آموزشی متناظر با این دو مدل تراوایی ایجاد شدند. بر اساس این مجموعه های آموزشی، دو ماتریس لغت نامه تعلیم داده شدند و بر اساس این ماتریس ها، دو مدل تراوایی معادل تقریب زده شدند. کیفیت مدل های تراوایی تقریب زده شده از دو منظر بررسی شدند: از منظر ساختار فیزیکی و از منظر خاصیت انتقال سیال. مقایسه کیفیت ساختار فیزیکی مدل تقریب زده شده نسبت به مدل اصلی از طریق مطالعه و بررسی حاصل تفریق این دو مدل انجام می شود. مقایسه کیفیت خاصیت انتقال سیال مدل تقریب زده شده از طریق مقایسه پروفیل های فشار و تولید برای مدل تقریب زده شده نسبت به پروفیل های فشار و تولید مدل اصلی انجام می پذیرد. پروفیل های فشار و تولید برای مدل ها با استفاده از روش شبیه سازی مخزن بدست می آید.

۲-۱- الگوریتم زمین آمار چند نقطه ای DisPat

روش زمین آمار چند نقطه ای DisPat برای اولین بار توسط Honarkhah (2011) ارائه شد. در این الگوریتم، مانند سایر الگوریتم های الگو-اساس زمین آمار چند نقطه ای، تصویر آموزشی توسط یک قالب طراحی شده اسکن می شود و الگوها از تصویر آموزشی استخراج می گردند. سپس الگوها بر اساس توابع فاصله و با استفاده از روش های کرنل کلاسه بندی می شوند. فاصله بین دو الگو در فضای متریک با یک تابع فاصله اندازه گیری می شود. هر دو نقطه نزدیک به هم در فضای متریک مربوط به دو الگوی مشابه موجود در پایگاه داده الگو است. از الگوریتم خوشه بندی kernel k-means به منظور کلاسه بندی الگوها در پایگاه داده استفاده شده است.

پس از کلاسه بندی الگو، به صورت متوالی روی شبکه تحقیق فرآیند شبیه سازی صورت می گیرد. نزدیک ترین خوشه به رویداد داده مشخص می شود و یک الگو از آن خوشه به صورت تصادفی بر مکان آن

گره جای گذاری می شود. به منظور بهبود خواص مکانی تحقیق ها در فواصل دور، الگوریتم DisPat از دو روش چند-شبکه و چند-تفکیک بهره می برد. در روش چند-شبکه، اندازه گرید تغییر نمی کند؛ ولی اندازه قالب از یک تراز به تراز بعدی تغییر می کند. از سوی دیگر، در روش چند-تفکیک، اندازه قالب ثابت است و اندازه گرید از یک تراز به تراز بعدی متغیر است. فرآیند شبیه سازی از تفکیک درشت تر آغاز می شود و تا ریزترین تفکیک (تفکیک نهایی) ادامه می یابد.

الگوریتم DisPat دو چارچوب ادغام داده سخت برای دو روش چند-شبکه و چند-تفکیک در نظر می گیرد. در روش چند-شبکه، الگوریتم DisPat هر داده سخت را به نزدیک ترین گره ممکن اختصاص می دهد و اگر پیش از این به نزدیک ترین گره داده ای اختصاص داده شده باشد، داده سخت به نزدیک ترین گره بعدی اختصاص خواهد یافت؛ که این داده سخت را احاطه کرده است. داده های سختی که در این دستورالعمل جا می مانند؛ از فرآیند شبیه سازی حذف می شوند. در روش چند-تفکیک، اندازه شبکه ها در هر مرحله نسبت به مرحله دیگر متفاوت است و مکان گره ها از یک تراز به تراز دیگر متفاوت است و هم ارزی یک-به-یکی وجود ندارد. در این حال، از میانمایی سه بعدی برای انتقال داده سخت از یک تراز به تراز بعدی استفاده شده است.

در DisPat، ادغام داده نرم از طریق محاسبه فاصله انجام می گیرد. در این الگوریتم، دو مجموعه پایگاه داده برای هر دو تصویر آموزشی اصلی و تصویر آموزشی نرم شکل می گیرد. الگوها کلاسه بندی شده و نمونه های اولیه به هر کلاسه اختصاص داده می شوند. برای هر رویداد داده که روی یک مکان گره قرار دارد، فاصله تا نمونه های اولیه در هر دو مجموعه پایگاه های داده محاسبه می شود. این دو فاصله محاسبه شده ترکیب می شوند؛ تا یک تابع جدید فاصله را ایجاد کنند؛ که در آن به فاصله نرم (فاصله رویداد داده تا نمونه های اولیه مربوط به داده نرم) وزنی اختصاص داده شده است.

تصاویر آموزشی می توانند ایستا یا ناپایستا باشند. در یک تصویر ایستا، الگوها تکرار می شوند و خواص آماری ثابت است؛ در حالی که در یک تصویر ناپایستا، برای هیچ مکانی نمی توان اطلاعات از پیش تعیین شده ای در نظر گرفت. Honarkhah (2011) به منظور ایجاد تحقیق های ناپایستا سه الگوریتم ارائه کرده است؛ که تنها پیش نیاز آن وجود تصویر آموزشی است. در روش شباهت مکانی^۱، الگوریتم DisPat الگوها را در یک پایگاه داده و مکان الگوها را در یک پایگاه داده دیگر ذخیره سازی می کند. روش SSM، فرآیند یافتن شبیه ترین الگو را روی تمام شبکه انجام می دهد. شبیه ترین الگو به رویداد داده مورد نظر، الگویی است که تابع فاصله متشکل از ترکیب فاصله وزن داده شده الگو و فاصله وزن داده شده مکانی را کمینه می کند؛ به طوری که مجموع این دو وزن برابر یک می شود. روش شعاع همسایگی^۲ اساسا شبیه روش SSM است؛ با این تفاوت که در فرآیند شبیه سازی، یافتن شبیه ترین الگو درون یک دایره

1 Spatial Similarity Method (SSM)

2 Neighborhood-Radius Method (NRM)

حتما راه حلی برای $Gm = d$ وجود دارد. این راه حل یکتاست و تنک-ترین راه حلی است که می‌توان با روش‌های دستیابی^۴ بدان دست یافت. در اینجا، ماتریس $d \in \mathbb{R}^{n \times N}$ ($n \ll N$) مجموعه‌ای از N سیگنال است؛ که ستون‌های آن را تشکیل می‌دهد و m ، راه حل فشرده شده است؛ که آنتروپی آن قطعا نسبت به آنتروپی سیگنال d کمتر است (Aharon et al., 2006a and 2006b; Horev et al., 2012;) (Rubinstein et al., 2010b).

انتقال تنک در زمینه فشرده سازی، استخراج ویژگی، قاعده مند سازی در مسائل وارون، حذف نویز، فشرده سازی بازه دینامیک در تصاویر، جداسازی نوشته‌ها از محتوای تصاویر، رنگ آمیزی، تصویرسازی صورت و غیره کاربرد های زیادی دارد.

مشکل اصلی ارائه تنک پرداختن به مساله نرم l^0 است. روش دستیابی اساسی^۵ (Chen et al., 1999) این مساله را به نرم l^1 تبدیل می‌کند؛ که یک مساله بهینه سازی است. بدین معنی که معادلات (۱) و (۲) با جایگزینی l^1 نرم به جای l^0 نرم، محدب سازی^۶ می‌شوند. بنابراین هدف روش دستیابی اساسی پیدا کردن راه حلی برای مساله ارائه تنک است؛ به طوری که تعداد ضرایب غیر صفر کمینه گردد. از سوی دیگر، Lasso (Tibshirani, 1996) تعداد ضرایب غیر صفر را کمینه نمی‌کند؛ بلکه آنقدر ضرایب غیر صفر را به صفر تبدیل می‌کند، تا تعداد ضرایب غیر صفر به حد مشخصی برسد. Lasso در واقع یک مساله کدگذاری تنک با محدودیت اعمال شده روی پراکندگی است. Lasso در عمل تغییر یافته الگوریتم LARS است.

۲-۳- الگوریتم های انتخاب مدل

در این مقاله به منظور انتخاب بردار (مدل) های مناسب، از الگوریتم LARS استفاده شده است. روش LARS توسط Efron et al. (2004) ارائه شد. روش LARS به صورت متوالی مدل رگرسیون را ایجاد می‌کند و در هر توالی، بهترین متغیر ممکن به مجموعه متغیرهای فعال اضافه می‌شود. این روش با تمام متغیرهای موجود در مجموعه فعال کار نمی‌کند و از این میان تنها تعداد مناسبی از متغیرها را انتخاب می‌کند و در رگرسیون بکار می‌برد.

ورودی ها باید به مقدار متوسط صفر متمرکزسازی و به نرم یک استانداردسازی شوند. الگوریتم با ضرایب $\hat{m}_1, \hat{m}_2, \dots, \hat{m}_k = 0$ شروع بکار می‌کند و باقیمانده برابر $r = d - \bar{d}I$ قرار داده می‌شود؛ بگونه‌ای که $\bar{d} = \sum_{i=1}^n d_i / N$ الگوریتم، کار را با پیدا کردن متغیری که بیشترین هم-ارزی^۷ را با پاسخ دارد، مثلا $\hat{m}_i$ آغاز می‌کند. مقدار $\hat{m}_i$ (که صفر است) به صورت پیوسته به سمت مقدار ضریب حداقل

تعریف شده همسایگی آن رویداد انجام می‌پذیرد. روش قطعه بندی خودکار^۱ بر اساس استخراج مشخصه های مختلف خواص موجود در تصویر و کلاسه بندی تصویر آموزشی بر اساس مشخصه های استخراج شده از تصویر با روش کلاسه بندی k-means بنا شده است. بر اساس این ویژگی های ورودی، الگوریتم k-means تصویر را به k زیر ناحیه تقسیم می‌کند.

۲-۲- تقریب تنک از طریق تعلیم لغت نامه

یک سیگنال مانند $d \in \mathbb{R}^n$ به عنوان ترکیب دقیق خطی از ستون های ماتریس $G \in \mathbb{R}^{n \times k}$ به صورت $Gm = d$ ارائه می‌گردد؛ به گونه ای که $m \in \mathbb{R}^k$ باشد. تنک ترین حل برای این مساله به صورت $\min_m \|m\|_0$ نشان داده می‌شود؛ به گونه ای که $\|.\|_0$ تعداد عناصر غیر صفر را در این راه حل بیان می‌کند. دلخواه ترین پاسخ برای این سیستم معادلات پاسخی است که کمترین تعداد ضرایب غیر صفر را داشته باشد.

در یک مساله پردازش تصویر، ماتریس G لغت نامه و ستون های آن اتم خوانده می‌شوند. این اتم ها، بلوک های سازنده لغت نامه و نرمال شده هستند. بردار d را می‌توان به صورت ترکیب خطی از اتم های لغت نامه، $\{G_j\}_{j=1}^k$ بیان کرد. Wedgelets, Contourlets, Curvelets, Steerable Wavelets, Bandlets, Short-Time Fourier Transforms مثال هایی از ماتریس های لغت نامه هستند (et Starck al., 2010). ارائه بردار d بر حسب اتم های لغت نامه می‌تواند دقیق یا تقریبی باشد؛ ارائه دقیق بردار d به صورت زیر نوشته می‌شود:

$$\min_m \|m\|_0 \quad \text{subject to} \quad Gm = d \quad (1)$$

راه حل مساله دقیق ارائه شده در رابطه (۱) به صورت $Q(G^{-1}d)$ است و این پاسخ تنها در صورتی وجود دارد؛ که G قابل وارون سازی باشد. در مسائل واقعی، پاسخ دقیق معمولا دست یافتنی یا دلخواه نیست و پاسخ تقریبی کفایت می‌کند. یافتن پاسخ تقریبی برای مساله $Gm \approx d$ که دارای معیار بیشترین تنکی است، تقریب تنک گفته می‌شود و به صورت زیر ارائه می‌شود:

$$\min_m \|m\|_0 \quad \text{subject to} \quad \|Gm - d\|_2 \leq \epsilon \quad (2)$$

در این مورد، ماتریس G یک لغت نامه بیش از حد کامل^۲ است. به کل این فرآیند روش فشرده سازی گفته می‌شود. در موارد معروف کدگذاری تبدیلی^۳، از لغت نامه های DCT و تبدیل موجک به منظور فشرده سازی داده‌ها استفاده می‌شود. فرآیند فشرده سازی، افزونگی داده را پیدا و حذف می‌کند. برای مجموعه ای متشکل از N سیگنال، اگر ستون های ماتریس $G \in \mathbb{R}^{n \times k}$ نرمال شده باشند و $m \in \mathbb{R}^{k \times N}$ ($n < k \ll N$) با داشتن s_0 عنصر غیر صفر در هر ستون به اندازه کافی تنک باشد، آنگاه

5 basis pursuit
6 convexitate
7 correlation

1 Automatic Segmentation Method (ASM)
2 over complete
3 transform coding
4 pursuit methods

ایده الگوریتم تعلیم لغت نامه حداقل مربعات مکرر (ILS-DLA) توسط Engan et al. (2007) ارائه شد. روش ILS-DLA مساله تقریب تنگ را با در نظر گرفتن ساختار درونی لغت نامه حل می کند. ماتریس لغت نامه به تعداد مشخصی از ماتریس های کوچک تقسیم بندی می شود؛ که این بلوک های داخلی ممکن است به صورت همپوشان یا ناهمپوشان کنار یکدیگر قرار گرفته باشند. در ILS-DLA، لغت نامه ساختار بندی شده است و برای هر بلوک درون لغت نامه، داده های متناظر و ضرایب آنها در فرآیند وارون سازی دخیل هستند؛ تا بخشی از لغت نامه که درون آن بلوک است را به روزرسانی کند. این بلوک به روزرسانی شده از لغت نامه، ثابت است و در نتیجه در هر مرحله از وارون سازی، فرآیند به روزرسانی لغت نامه به به روزرسانی بخش بسیار کوچکی از لغت نامه متناظر با آن بلوک کاهش می یابد. دستورالعمل الگوریتم ILS-DLA در ادامه آمده است.

مربعات متناظر با آن، $\bar{m}_i$ ، کشانده می شود؛ مقدار باقیمانده به صورت پیوسته کاهش می یابد. این کشاندن پیوسته تا جایی ادامه می یابد که هم-ارزی یک متغیر جدید مانند $\bar{m}_j$ با باقیمانده با هم-ارزی متغیر فعلی مانند $\bar{m}_i$ برابر شود. این کشاندن پیوسته مقادیر برای دو ضریب $\bar{m}_i$ و $\bar{m}_j$ به سمت مقادیر ضرایب حداقل مربعات متناظر با آنها آنقدر ادامه می یابد تا این که هم-ارزی یک متغیر سوم مانند $\bar{m}_l$ با باقیمانده فعلی برابر هم-ارزی ترکیب متغیرهای $\bar{m}_i$ و $\bar{m}_j$ با باقیمانده فعلی گردد. بعد از f مرحله، تعداد f عدد از متغیرها غیر صفر و مابقی هنوز صفر هستند. این فرآیند آنقدر ادامه می یابد تا همه متغیرهای لازم، مانند k ، در مدل قرار گیرند و پاسخ کامل حداقل مربعات بعد از $\min(k, n-1)$ مرحله بدست آید.

۲-۴- الگوریتم ILS-DLA

الگوریتم ILS-DLA، برداشت شده از (Engan et al., 2007 و Skretting and Engan, 2011b)

Having the set of data, $d_{aug} = \{d_j\}_{j=1}^L$

Defining an initial dictionary, G^0

Set $itr = 0$

The sparse coding step:

Finding the coefficient matrix, $m_{aug}^{(i)}$, by solving for the optimization problem as

$$\underset{m_{aug}}{\operatorname{argmin}} \|d_{aug} - G^{(i)} m_{aug}\|_F^2 \quad \text{subject to } \|m_l\|_0 = s_0 \quad \forall l \quad (3)$$

using the LARS method

The dictionary update step:

Having the coefficient matrix, $m_{aug}^{(i)}$, obtained from the previous step, the dictionary is updated by solving for the optimization problem

$$\underset{G}{\operatorname{argmin}} \|d_{aug} - G m_{aug}^{(i)}\|_F^2 \quad (4)$$

using an inversion step as

$$G^{(i+1)} = d_{aug} (m_{aug}^{(i)})^T \left[m_{aug}^{(i)} (m_{aug}^{(i)})^T \right]^{-1} \quad (5)$$

The dictionary columns should to be normalized

Set $itr = itr + 1$

به قطعه های کوچکتر تقسیم بندی می شوند و سپس مقدار متوسط ثابت^۲ برای هر قطعه از مقادیر مربوط به آن قطعه کم می شود. سپس مقادیر غیر صفر آن قطعه ها با یک تابع تدریجی کننده یکنواخت^۳ درجه بندی^۴ می شوند. علاوه بر این، کدگذاری های آنتروپی، AC، و هافمن نیز روی تصاویر انجام می شود. این قطعه ها سپس به منظور تعلیم لغت نامه

۲-۵- فشرده سازی تصویر مدل

می توان الگوریتم ILS-DLA را به منظور ایجاد تصاویری که به صورت تنگ تقریب زده شده اند، در یک رویه فشرده سازی تصویر بکار برد. فرآیند فشرده سازی تصویر شامل دو مرحله الگوریتم تعلیم ILS-DLA آفلاین و فرآیند کدگذاری تصویر آنلاین است. در این رویه، نخست تصاویر

3 uniform quantizer

4 quantize

1 patch

2 mean DC value

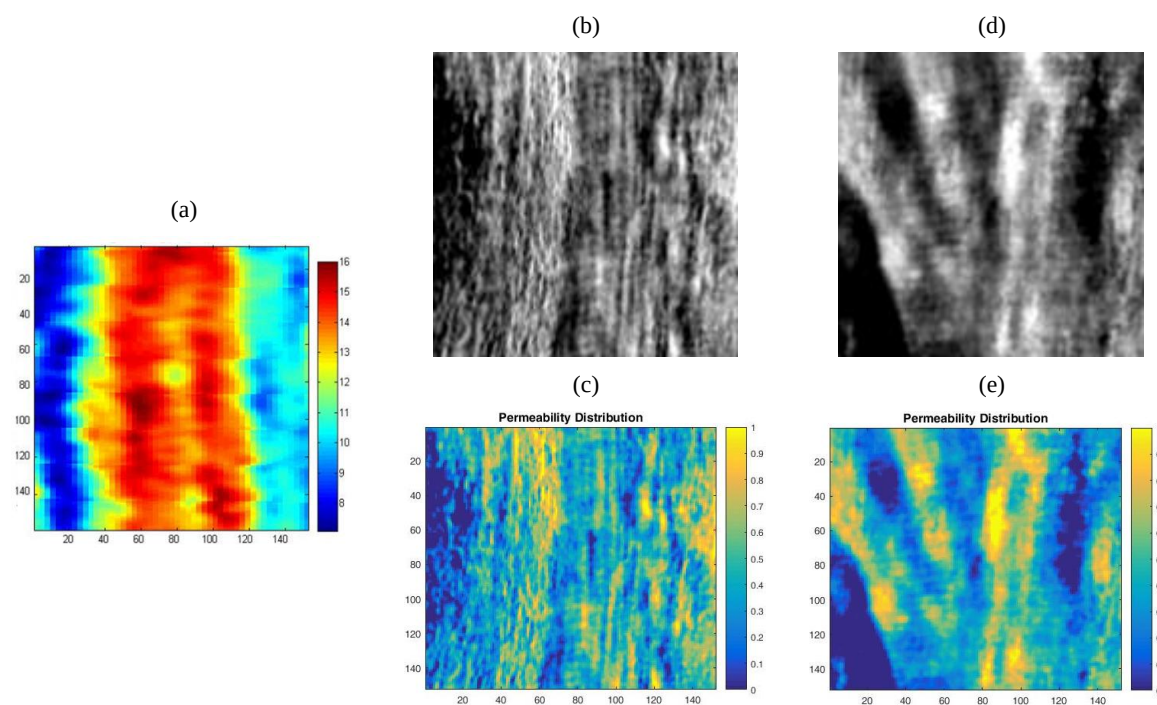
نامه مختلفی به منظور تعلیم لغت نامه های محتوا-خاص بر روی مجموعه‌های آموزشی متشکل از تحقیق های ایجاد شده با الگوریتم‌های زمین آمار چند نقطه‌ای، بکار برده شدند. روش های کدگذاری تنک مختلفی به منظور بازسازی مدل تصویر فشرده تنک بکار برده شدند. نسبت کیفیت مدل تصویر فشرده بازسازی شده تنک در مقایسه با تحقیقات ایجاد شده با روش‌های زمین آمار چند نقطه ای موجود در مجموعه آموزشی، مطالعه شده و آمار آن ارائه شده است. ملاک‌های ارزیابی مدل تصویر بازسازی شده فشرده تنک، شباهت ساختار فیزیکی مدل‌ها و خواص جریان سیال بوده است.

مدل های انتخاب شده به عنوان تصاویر آموزشی در فرآیند زمین آمار چند نقطه ای باید بر اساس واقعیت مخزن انتخاب و با هر نوع داده مخزن که در دسترس است، سازگار باشند. بر این اساس، یک مقطع تخلخل (شکل ۱a) از پروفیل افقی تخلخل استخراج شده از مکعب سه بعدی داده لرزه ای انتخاب شد. به علاوه، دو مدل تراوایی بر اساس ویژگی‌های مشاهده شده در پروفیل بر اساس افق^۳ تجزیه طیفی انتخاب شدند (شکل‌های ۱b و ۱d). مدل‌های نشان داده شده در شکل ۱ از ستیغ مخزن مورد مطالعه، در راستای غالب چین‌خوردگی‌های زاگرس (شمال-غرب-جنوب شرق)، و با ابعاد 760 × 760 مترمربع گرفته شده؛ به گونه‌ای که راستای واگرایی پنجه‌های دلتا به سمت جنوب شرق است.

با استفاده از الگوریتم ILS-DLA بکار برده می شوند. تبدیل کوسینوس ناپیوسته (DCT) به عنوان ماتریس لغت نامه ثابت بکار برده می شود. به عنوان سایر تبدیل های جایگزین می توان به روش های LOT و LET اشاره کرد. ماتریس نهایی تعلیم داده شده ماتریسی مانند S خواهد بود؛ این ماتریس همان لغت نامه ای است، که به صورت خاص روی تصاویر موجود در مجموعه آموزشی با استفاده از الگوریتم ILS-DLA تعلیم داده شده است. سپس ضرایب ماتریس S درجه بندی و بازتشکیل^۱ می شود. حاصل ضرب لغت نامه ثابت و ضمنی (ϕ)، مانند DCT و لغت نامه تنک صریح (S) منجر به ایجاد ماتریس ورودی-انطباقی $G = \phi S$ می شود. الگوریتم LARS نیز به عنوان روش کدگذاری تنک بکار برده می شود.

۳- نتایج

نتایج این مطالعه چندگانه است. مکعب داده لرزه ای سه بعدی با روش قطعی^۲ وارون سازی شده و مکعب های سه بعدی مقاومت صوتی، تخلخل، و اشباع آب بدست آمدند. مکعب سه بعدی تجزیه طیفی نیز ایجاد و مطالعه شد. دو مدل تراوایی بر اساس واقعیت از مکعب تجزیه طیفی استخراج شد و بر این اساس، تحقیق های بسیاری بر اساس روش های مختلف زمین آمار چند نقطه ای ایجاد شدند. الگوریتم های تعلیم لغت



شکل ۱: (a) توزیع تخلخل برای مدل شبیه‌سازی مخزن؛ (b) تصویر آموزشی شبکه شکستگی ایستا؛ (c) توزیع تراوایی متناظر با (a) ورودی به مدل شبیه‌سازی مخزن؛ (d) تصویر آموزشی دلتای نایستا؛ (e) توزیع تراوایی متناظر با (d) ورودی به مدل شبیه‌سازی مخزن
بر این اساس، روش‌های زمین آمار چند نقطه ای برای ایجاد تعداد زیادی از مدل های تصادفی بکار گرفته شدند. تعداد اندکی از تحقیقات

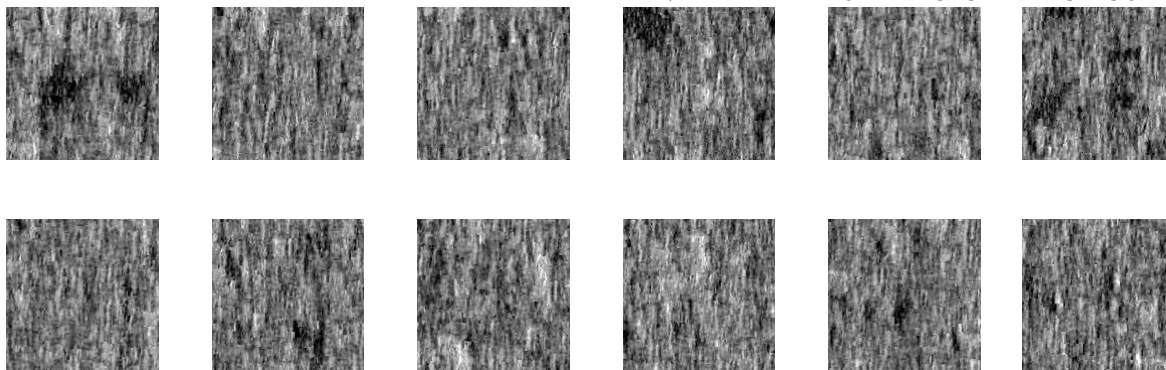
شده از طریق روش ILS-DLA (Engan et al., 2007; Skretting and Efron et al., 2004; Hastie et al., 2009) و روش LARS (Engan, 2011b) روی مجموعه آموزشی متشکل از تحقیقات سیستم شکستگی ایستا تولید شده با الگوریتم زمین آمار چند نقطه ای در شکل a5 نشان داده شده است.

به منظور بررسی کیفیت تصویر بازسازی شده، به جز کلیت ساختار فیزیکی تصاویر بازسازی شده، خواص جریان سیال برای تصاویر بازسازی شده نیز به عنوان یک معیار دیگر در نظر گرفته شد. به منظور محاسبه تغییرات فشار و نرخ تولید در چاه مشاهده ای بر حسب زمان و ترسیم پروفیل آنها، از کدهای متلب MRST (Lie, 2015) استفاده شد. پروفیل های فشار و تولید مربوط به مدل تصویر فشرده تنگ سیستم شکستگی در مقابل پروفیل های فشار و تولید مدل اصلی رسم شده است (شکل های a6 و b6). پروفیل تولید سیال تزریق شده در چاه مشاهده ای بر حسب زمان در حال افزایش و پروفیل فشار آن در حال کاهش است.

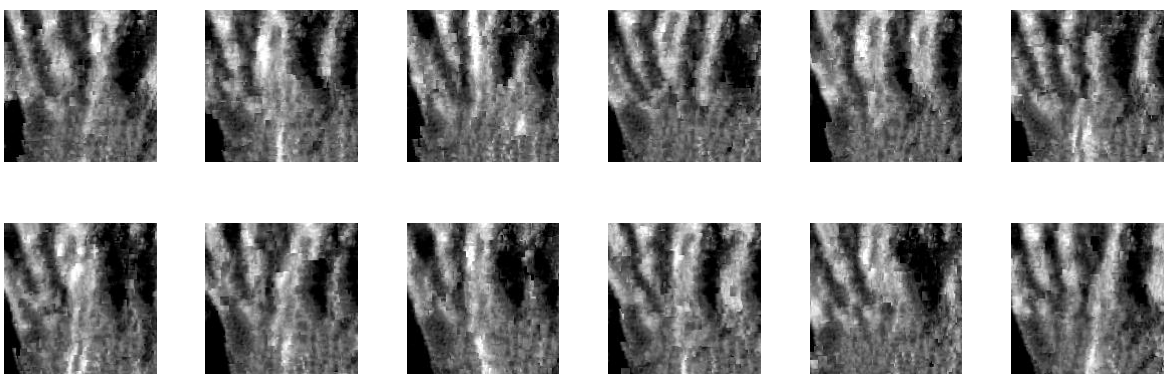
انجام شده از روش DisPat به صورت تصادفی از مجموعه ای متشکل از سه هزار فضای نمونه ای انتخاب شده اند؛ که می توان تصویر آنها را در شکل های ۲ و ۳، به ترتیب بر اساس تصاویر آموزشی ارائه دهنده سیستم شکستگی (شکل b1) و سیستم دلتایی (شکل d1)، مشاهده کرد. جهت و ابعاد مدل های نشان داده شده در شکل های ۲ و ۳ هم راستا و هم اندازه با مدل های تصویر شده در شکل ۱ هستند.

مجموعه مدل های تولید شده از طریق روش DisPat به عنوان مجموعه های آموزشی در فرآیند تعلیم لغت نامه از طریق روش ILS-DLA با LARS بکار برده شدند. در فرآیند تبدیل تنگ، ارائه شده در معادله (۱)، هر تصویر به عنوان یک بردار ستونی در m مشارکت می کند. ستون های لغت نامه تعلیم داده شده را اتم می نامند؛ که اساس های ارائه تنگ سیگنال ها و تصاویر را تشکیل می دهند. شکل های a4 و b4 تعداد ۵۶ اتم از این لغت نامه های تعلیم داده شده روی مجموعه های آموزشی متشکل از تحقیقات زمین آمار چند نقطه ای به ترتیب برای سیستم شکستگی ایستایی و سیستم دلتایی ناپایستا، را نشان می دهد.

مدل تصویر فشرده تنگ بازسازی شده توسط لغت نامه تعلیم داده



شکل ۲: چند تحقیق تولید شده با روش زمین آمار چند نقطه ای DisPat بر اساس تصویر آموزشی شبکه شکستگی ایستا



شکل ۳: چند تحقیق تولید شده با روش زمین آمار چند نقطه ای DisPat بر اساس تصویر آموزشی سیستم دلتایی ناپایستا

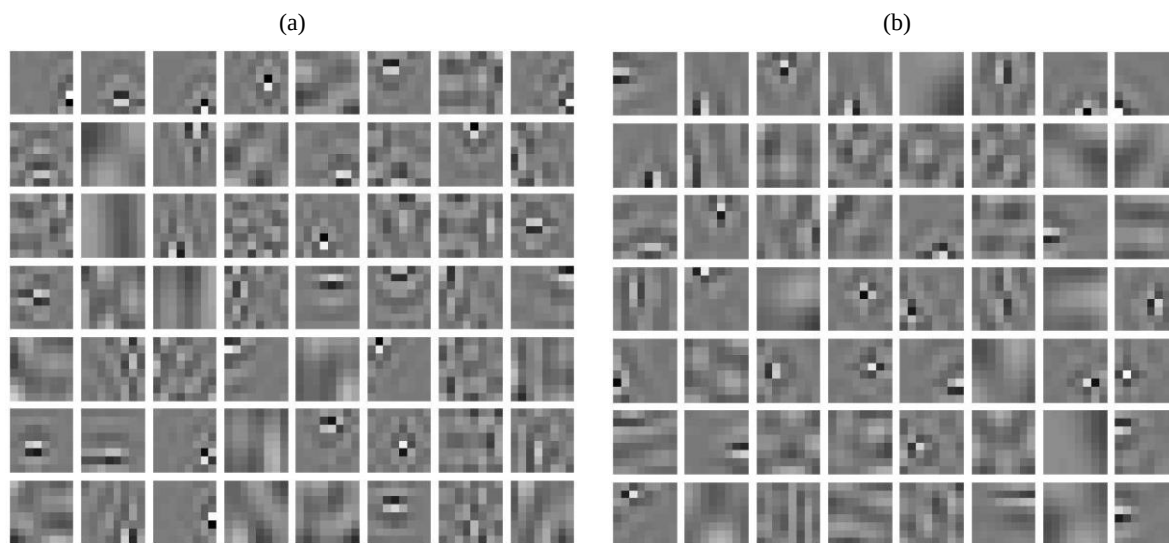
تولید مدل اصلی رسم شده اند (شکل های c6 و d6). برای هر دو پروفیل فشار و تولید، خطاها محاسبه شده اند و به عنوان معیاری برای بررسی کیفیت مطلوب تطابق تاریخچه مخزن بکار برده شده اند. گردش کار توضیح داده شده در این مقاله ۴۸ بار برای تنظیم های

پروفیل های تولید برای مدل بازسازی شده سیستم شکستگی ایستا (شکل a6) و مدل بازسازی شده سیستم دلتایی ناپایستا (شکل b6) با پروفیل تولید تصویر اصلی تقریباً همخوانی کامل دارد. همچنین پروفیل های فشار مدل بازسازی شده برای سیستم دلتایی مقابل پروفیل فشار و

واقعی (مدل‌های شکل b1 و d1) در مقابله با خطای تطابق تولید یا فشار برای تک‌تک اعضای فضای نمونه‌ای (متشکل از ۳۰۰۰ تحقیق موجود در مجموعه آزمایشی مربوط به آن آزمایش خاص) محاسبه شده است. به عنوان مثال اگر پروفیل تولید مدل تصویر فشرده تنک بازسازی شده در یک آزمایش ۹۸٪ بهتر بودن را نشان دهد، بدین معنی است که خطای پروفیل تولید مربوط به آن تصویر بازسازی شده از خطای پروفیل تولید مربوط به ۹۸٪ از تحقیق‌های موجود در آن مجموعه آموزشی کمتر است. با در نظر داشتن همه ۴۸ آزمایش نتایج بسیار تهييج کننده است. جالب است اشاره شود که از ۴۸ آزمایشی انجام شده ۵۴٫۱۷٪ از کل مدل‌های فشرده تنک در محدوده ۹۵-۹۵٪ خوب بودن نسبت به فضای نمونه‌ای قرار می‌گیرند (توجه شود که ۹۵٪ اول متناظر با احتمال خوب بودن از لحاظ تولید و ۹۵٪ دوم متناظر با احتمال خوب بودن از لحاظ فشار است، یا برعکس)، ۸۹٫۵۸٪ از مدل‌های فشرده تنک در محدوده ۹۰-۹۰٪، و ۹۵٫۸۳٪ در محدوده ۸۵-۸۵٪ قرار می‌گیرند.

مختلف از آزمایش‌ها انجام شده است. برای هر آزمایش سه بخش متفاوت در نظر گرفته شده است که عبارتند از تولید تحقیق‌های زمین‌آمار چند نقطه‌ای، تعلیم لغت نامه، و کدگذاری تنک. از لحاظ روش، آزمایش‌ها از طراحی "یک در هر بار" پیروی می‌کنند بدین معنی که هر آزمایش نسبت به آزمایش‌های دیگر بر اساس روش‌های بکار برده شده در حداقل یکی از بخش‌های ذکر شده متفاوت است. برای دو نوع تصاویر آموزشی ایستا و نایستا (شکل‌های b1 و d1)، مجموعاً ۴۸ آزمایش بر اساس طراحی "یک در هر بار" برقرار شده است.

نتایج تمامی ۴۸ آزمایش در شکل ۷ آمده است. هر نقطه متناظر با یک آزمایش است. محور افقی نشان دهنده بهتر بودن مدل نهایی فشرده تنک (بازسازی شده بر اساس تنظیم‌های خاص مساله در آن آزمایش) از لحاظ تولید است. به همین ترتیب محور عمودی نشان دهنده بهتر بودن مدل از لحاظ فشار است. شاخص بهتر بودن برای هر محور بر اساس خطای تطابق تولید یا فشار برای مدل بازسازی شده در مقایسه با مدل



شکل ۴: (a) چند اتم انتخاب شده از لغت نامه تعلیم داده شده بر اساس روش ILS-DLA برای تحقیق‌های ایجاد شده بر اساس تصویر آموزشی شبکه شکستگی؛ (b) همان، برای تحقیق‌های ایجاد شده بر اساس تصویر آموزشی سیستم دلتایی

۴- بحث

در این مقاله الگوریتم مدل‌سازی و شناسایی مخزن فشرده تنک ارائه شد؛ که قادر است تصاویر مدل را چندگانه سازی کند و ترجمه نزدیکی از مدل اصلی ارائه دهد. مدل فشرده تنک نهایی با مدل واقعی برابر نیست و همچنین نمی‌توان ادعا کرد که از مدل‌های موجود در میان یک مجموعه بزرگ آموزشی بهتر است. بر اساس تنظیمات به کار برده شده برای آزمایش‌ها و اعمال شده بر روی مجموعه تحقیق‌های تولید شده بر

با در نظر داشتن تنظیمات آزمایش‌های بحث شده در این مقاله، نتایج نشان دهنده بهتر بودن مدل تصویر فشرده تنک نسبت به ۹۷٫۰۸٪ از مجموعه تحقیق‌های زمین‌آمار چند نقطه‌ای ایجاد شده بر اساس تصویر شکستگی ایستا است. همین مقایسه برای تصویر سیستم دلتایی نایستا نشان دهنده بهتر بودن نسبت به ۹۴٫۴۱٪ از تمام تحقیق‌های ایجاد شده از روش‌های زمین‌آمار چند نقطه‌ای است.

نفی شناخته شده است و دو نوع تخلخل شکستگی و زمینه برای این مخزن در فرایند شبیه سازی مخزن در نظر گرفته شده است. این باور وجود دارد؛ که در ناحیه ستیغ مخزن، تبادل فشار از طریق شبکه شکستگی باز، انجام می پذیرد.

در زمان انجام این مطالعه ۲۰ چاه در گستره مخزن حفاری شده بود. توزیع مناسب چاه ها بانک اطلاعات مناسبی از کل مخزن فراهم کرده است. گاز از طریق چاه شماره ۸ در سال ۲۰۰۵ و از طریق چاه شماره ۳ در سال ۲۰۰۷ به منظور حفظ فشار در حال تزریق به مخزن است. برداشت داده های لرزه ای سه بعدی تک مولفه ای طی سال های ۲۰۰۹ و ۲۰۱۰ انجام گرفت.

۲-۴- عدم قطعیت و خاصیت تصادفی بودن

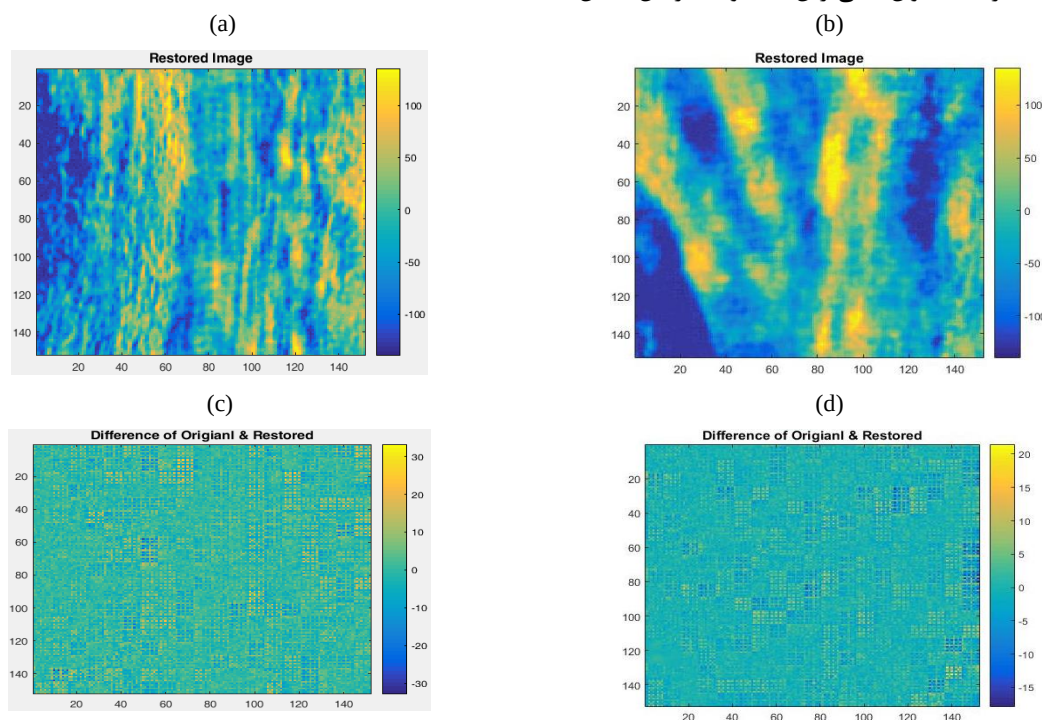
مکعب سه بعدی داده لرزه ای برای مخزن مورد مطالعه با استفاده از نرم افزار تجاری پترل تفسیر شد و داده های چاه و داده های پروفیل لرزه ای به عنوان داده های ورودی برای انجام این کار مورد استفاده قرار گرفت. هیچ گسل موثری در محدوده مخزن نفت تفسیر نشد. تفسیر دقیق گسل ها خصوصا وقتی میدان تحت تزریق گاز باشد، بسیار حائز اهمیت است. با داشتن خواص لرزه ای به عنوان متغیر ها، توزیع تخلخل در مخزن را می توان با تبدیل های خطی یا غیر خطی و با محدود سازی به داده های تخلخل واقع در مکان چاه، بدست آورد. مکعب سه بعدی توزیع تخلخل با تبدیل چند خصوصیتی بدست آمد. برای این منظور، از نرم افزار تجاری Hampson-Russell استفاده شد.

اساس تصاویر آموزشی (شکل های a1 و b1) با استفاده از روش های مختلف تقریب تنک، می توان ادعا کرد که تصویر مدل فشرده تنک نهایی به احتمال ۸۹٫۵۸٪ نسبت به ۹۰٪ تحقیقات های ایجاد شده حاصل از زمین آمار چند نقطه ای و به احتمال ۹۵٫۸۳٪ نسبت به ۸۵٪ از تحقیقات عملکرد بهتری دارد.

گردش کار مدل سازی و شناسایی مخزن فشرده تنک، روش های مدل سازی زمین آمار چند نقطه ای را با تعلیم لغت نامه، کدگذاری تنک و فشرده سازی تصویر ترکیب می کند و با عمل روی داده های ورودی، تصاویر مدل فشرده تنک را ایجاد می کند. کنترل کیفیت مدل های بدست آمده از طریق انجام فرایند شبیه سازی مخزن برای همه تحقیقات زمین آمار چند نقطه ای به همراه مدل های فشرده تنک نهایی انجام می شود و پروفیل های تولید و افت فشار به عنوان معیاری از خوب بودن مدل ها بکار می رود. داده های ورودی برای این گردش کار مکعب های سه بعدی وارون سازی شده لرزه ای (مقاومت صوتی، تخلخل، و اشباع آب)، مکعب سه بعدی تجزیه کیفی داده لرزه ای، داده های تفسیر لرزه ای، داده های تولید، و نمودارهای چاه هستند. این گردش کار روی تصاویر مدل نیمه صنعتی مربوط به پروژه تزریق گاز در مخزن آسماری یکی از میدان های نفتی ایران، به خاطر اهمیت موضوع تزریق گاز و به خاطر نشان دادن تاثیر گردش کار فشرده تنک درباره مخازن پیچیده، بکار گرفته شده است.

۱-۴- مخزن مورد مطالعه، پیچیدگی ها، و عدم قطعیت ها

مخزن مورد مطالعه در این تحقیق یک میدان نفتی واقع در جنوب غربی ایران است. مخزن آسماری به عنوان اصلی ترین لایه تولید در این میدان

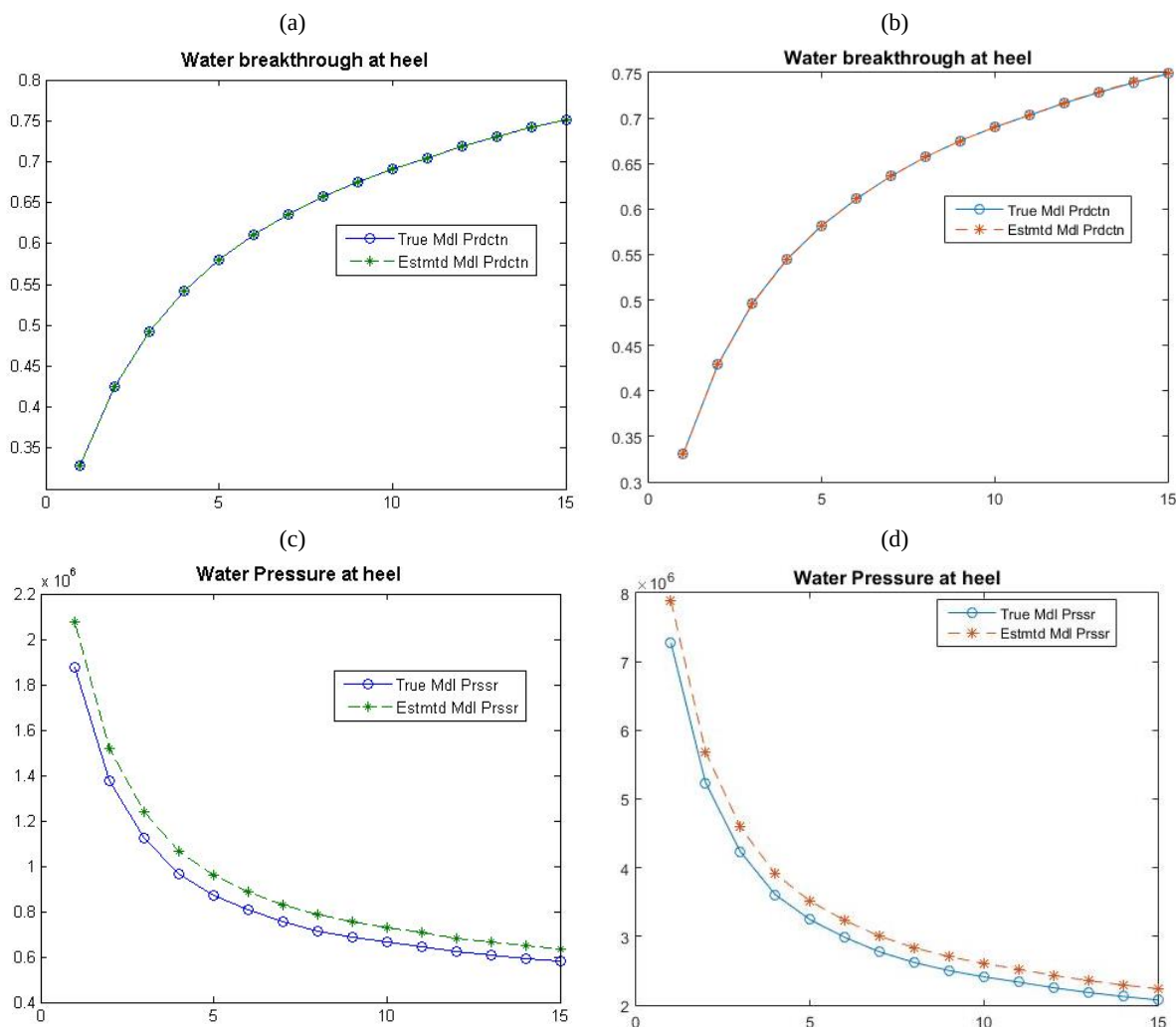


شکل ۵: (a) مدل تراوایی فشرده تنک برای شبکه شکستگی با استفاده از لغت نامه آموزش یافته با روش ILS-DLA و (b) همان، برای سیستم دلتایی؛ (c) تفاوت بین مدل اصلی و مدل فشرده برای سیستم شکستگی؛ (d) همان، برای سیستم دلتایی

برای هر مورد، مجموعه بزرگی از تحقیق‌های تصادفی ایجاد شد (شکل-های ۲ و ۳). احتمال رخ دادن هر کدام از تحقیق‌های تصادفی تولید شده از طریق روش زمین آمار چند نقطه‌ای با هم برابر است. شبیه‌سازی زمین آمار چند نقطه‌ای نسبت به هیچ داده سخت یا نرمی مشروط نشد تا خاصیت تصادفی بیشتری در مدل‌های تحقیق القا شود. به منظور ایجاد تحقیق‌های ذکر شده از کدهای متلب مرتبط (Honarkhah, 2011) استفاده شد.

۳-۴- تحقیق‌های زمین آمار چند نقطه‌ای

برای آزمایش‌های انجام شده، دو تصویر آموزشی از مکعب سه‌بعدی تجزیه طیفی داده‌های لرزه‌ای استخراج گردید؛ که ارائه دهنده شبکه شکستگی (شکل b۱) و سیستم دلتایی (شکل d۱) است. شبکه شکستگی، ایستا ولی سیستم دلتایی، ناپایستا فرض شده است. این دو تصویر به عنوان ورودی الگوریتم زمین آمار چند نقطه‌ای DisPat بکار گرفته شدند و



شکل ۶: (a) پروفیل تولید در چاه مشاهده‌ای برای مدل شبکه شکستگی و مقایسه آن با پروفیل تولید مدل اصلی؛ (b) همان، برای سیستم دلتایی؛ (c) پروفیل فشار در چاه مشاهده‌ای برای مدل شبکه شکستگی فشرده و مقایسه آن با پروفیل فشار مدل اصلی؛ (d) همان، برای سیستم دلتایی

تصویر آموزشی استفاده شد (شکل‌های ۲ و ۳). لغت نامه‌ها با استفاده از روش ILS-DLA روی این تحقیق‌ها تعلیم داده شدند (شکل‌های a۴ و b۴). هر تصویر را می‌توان بر اساس ترکیب خطی وزن یافته‌ای از تعداد خاصی از این اتم‌ها با استفاده از روش LARS تقریب زد. بر این اساس، مدل‌های فشرده تقریب تنک برای هر دو مدل شکستگی و دلتا بدست آمدند. از نرم‌افزار متلب به منظور اعمال الگوریتم‌های تقریب تنک

۴-۴- اعمال تقریب تنک روی تحقیق‌های زمین آمار چند نقطه‌ای

در گردش کار اعمال شده روی مخزن مورد مطالعه، مکعب تجزیه طیفی به عنوان منبع استخراج تصویر آموزشی شبکه ایستا و سیستم دلتایی ناپایستا بکار برده شد (شکل‌های b۱ و d۱). الگوریتم DisPat به منظور ایجاد مجموعه بزرگی از تحقیق‌های زمین آمار چند نقطه‌ای برای هر

مدل بازسازی شده شکستگی در شکل ۵a ارائه شده است و شکل ۵b نشانگر تفاوت بین تصویر اصلی و تصویر بازسازی شده است. همانطور که می توان در تصاویر تفاضل مشاهده نمود، مقادیر تفاضل در سراسر تصویر به صورت تصادفی است و ساختار خاصی را نشان نمی دهد و این نشان دهنده کیفیت خوب تصویر بازسازی شده است. همچنین مقادیر بیشینه و کمینه تفاضل، زیاد نیستند و تفاضل عمومی حول صفر باقی می ماند (به رنگ زمینه توجه نمایید). همین تفسیرها را می توان برای اشکال b5 و d5 متناظر با سیستم نایستای دلتا ارائه داد.

هر تحقیق ایجاد شده با روش های زمین آمار چند نقطه ای به همراه دو تصویر مدل فشرده تنک نهایی (شکل های a5 و b5) به عنوان مدل های ورودی به شبیه سازی مخزن بکار برده شدند و تاریخچه تولید و فشار هر مدل با تاریخچه تولید و فشار تصاویر مدل اصلی مقایسه گردید (شکل ۶). فرآیند شبیه سازی مخزن با نرم افزار متلب به انجام رسید (Lie, 2015).

بر اساس نتایج ارائه شده در این مقاله، با توجه به مقایسه پروفیل های فشار و تولید مدل های تقریب زده شده نسبت به مدل های اصلی (شکل ۶)، عملکرد الگوریتم تقریب تنک برای مدل مربوط به شکستگی ایستا نسبت به مدل مربوط به دلتا بهتر بوده است. با رجوع به شکل های ۲ و ۳ مشخص می شود که تحقیق های ایجاد شده برای مدل دلتا در مقایسه با تحقیق های ایجاد شده نسبت به مدل شکستگی، نسبت به همدیگر شبیه تر و از تنوع کمتری برخوردار هستند. از این رو، لغت نامه تعلیم داده شده بر اساس مدل های شکستگی با مدل های متنوع تری تعلیم داده شده و در نتیجه توانایی بیشتری در بازسازی مدل ها دارد.

لغت نامه هایی که بر اساس انواع خاصی از مدل ها با ویژگی های خاص تعلیم داده شده باشند در بازسازی مدل هایی که فاقد آن ویژگی یا از لحاظ آن ویژگی اندکی متفاوت باشند، عملکرد ضعیف تری خواهند داشت. هر چه مجموعه آموزشی (که یک لغت نامه بر اساس آن تعلیم داده می شود) حاوی مدل های متنوع تر با شاخص های متفاوت باشد، به نحوی که مدل های موجود در مجموعه آموزشی ارائه دهنده مناسبی از فضای نمونه ای قابل رخداد باشند، لغت نامه تعلیم داده شده در بازسازی مدل ها طبیعتاً عملکرد بهتری خواهد داشت.

۵-۴- مدل فشرده تنک

گردش کار مدل سازی و شناسایی مخزن فشرده تنک روی چند مجموعه از تحقیق های ایستا و نایستا تولید شده توسط روش های زمین آمار چند نقطه ای اعمال شد. رویه فشرده سازی تنک با استفاده از روش های

تعلیم لغت نامه موجود روی این مجموعه های آموزشی انجام شد. اطلاعات آماری بدست آمده از مدل های شبیه سازی برای هر تصویر مدل فشرده شده در شکل ۷ نشان داده شده است. برای هر تصویر مدل، دو معیار برای مقایسه لحاظ شده است: تولید و فشار. محور افقی نشان دهنده درصد بهتر بودن تصویر مدل فشرده تنک از لحاظ تولید و در مقایسه با تصاویر مدل در فضای نمونه ای است. محور عمودی مانند محور افقی ولی از لحاظ فشار است. تهییج کننده است که بدانیم از میان ۴۸ آزمایش انجام شده، ۵۴٫۱۷٪ از کل مدل های فشرده تنک در بخش ۹۵-۹۶٪ (۵٪ بالایی) بهتر بودن فضای نمونه ای قرار می گیرند، ۸۹٫۵۸٪ در بخش ۹۰-۹۰٪ (۱۰٪ بالایی)، و ۹۵٫۸۳٪ در بخش ۸۵-۸۵٪ (۱۵٪ بالایی).

حتی تهییج کننده تر است که توجه کنیم هر نقطه در شکل ۷ با استفاده از روش های مختلف فشرده سازی تصویر تنک ایجاد شده است. در تنظیمات هر آزمایش سه روش دخیل بوده است: روش های زمین آمار چند نقطه ای، الگوریتم های تعلیم لغت نامه، و روش های کدگذاری تنک. برای هیچ کدام از این آزمایش ها، بهینه سازی پارامتری انجام نگرفته است. بنابراین بهبود نتایج در صورت انجام بهینه سازی پارامتری امری کاملاً قابل انتظار است. بسیار قابل انتظار است که در صورت اعمال یک روش تقریب تنک روی مجموعه تحقیق های ایجاد شده از روش های زمین آمار چند نقطه ای با استفاده از چندین تصویر آموزشی مختلف، نتایج هنوز هم تهییج کننده تر باشند. نکته مهم دیگری که باید بدان اشاره کرد این است که تحقیق های زمین آمار چند نقطه ای ایجاد شده کاملاً تصادفی هستند، بدین معنی که هیچ محدودیت داده نرم یا سخت روی تصاویر اعمال نشده اند و بر این اساس، انتظار بر این است که در صورت اعمال محدودیت روی روش های زمین آمار چند نقطه ای و تولید تحقیق ها تحت این شرایط، هنوز نتایج بهتر و تهییج کننده تر گردند.

با در نظر گرفتن تطابق تاریخچه تولید و فشار به عنوان معیار مقایسه، نتایج آزمایش های اعمال الگوریتم تعلیم لغت نامه ILS-DLA با روش کدگذاری تنک LARS روی تحقیق های ایجاد شده توسط روش زمین آمار چند نقطه ای DisPat، برای تصویر آموزشی شکستگی ایستا نشان دهنده بهتر بودن تصویر مدل فشرده تنک نسبت به ۹۷٫۰۸٪ از مجموعه تحقیق های ایجاد شده است. این بدین معنی است که از میان تعداد ۳۰۰۰ تحقیق زمین آمار چند نقطه ای ایجاد شده، تنها ۸۸ تحقیق منجر به تطابق تاریخچه بهتر می شوند. همین مقایسه برای تصویر سیستم دلتایی نایستا نشان دهنده بهتر بودن تصویر فشرده نهایی نسبت به ۹۴٫۴۱٪ از کل تحقیق های زمین آمار چند نقطه ای ایجاد شده است، یعنی تعداد ۱۶۷ مدل از میان تحقیق های زمین آمار چند نقطه ای ایجاد شده از لحاظ تطابق تاریخچه از مدل فشرده تنک نهایی بهتر هستند.

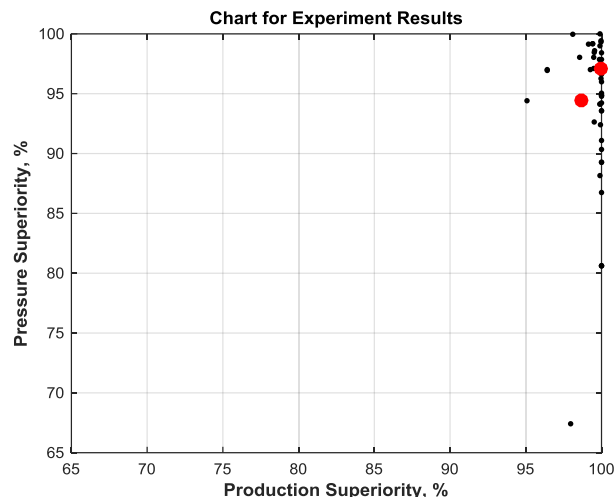
نقطه‌ای مربوط به شبکه شکستگی ایستا، تصویر مدل فشرده تنک نهایی نسبت به ۹۷,۰۸٪ نمونه‌ها از لحاظ تطابق تاریخچه برتری دارد؛ در مورد مجموعه تحقیق‌های ایجاد شده برای سیستم دلتایی نایستا، این عدد ۹۴,۴۱٪ است.

۶- سپاس‌گزاری

نویسندگان این مقاله از موسسه ژئوفیزیک دانشگاه تهران و معاونت‌های زمین‌شناسی و مهندسی نفت شرکت ملی مناطق نفت‌خیز جنوب (NISOC) برای فراهم آوردن داده‌ها، منابع، و راهنمایی در انجام این مطالعه تشکر می‌نمایند.

۷- منابع

- Aharon, M., Elad M., and Buckstein, A.M., 2006a, On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them, *Journal of Linear Algebra and Applications*, 416 (2006), 48–67.
- Aharon, M., Elad, M., and Bruckstein, A. M., 2006b, The K-SVD: an algorithm for designing of over complete dictionaries for sparse representation, *IEEE Trans. Signal Process.*, vol. 54, no. 11, 4311–4322.
- Bryt, O. and Elad, M., 2008, Compression of facial images using the K-SVD algorithm. *Journal of Visual communication and Image Representation*, 19(4), 270–283.
- Chen, S. S., Donoho, D. L., and Saunders, M. A., 1999, Atomic decomposition by basis pursuit, *SIAM J.Sci. Comp.*
- Deutsch, C. V., Journel, A. G., 1998, *GSLIB: Geostatistical Software Library and Users Guide*, second edition, Oxford, New York.
- Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R., 2004, Least angle regression (with discussion), *Annals of Statistics* 32(2), 407–499.
- Elad, M., 2010, *Sparse and Redundant Representations, from Theory to Applications in Signal and Image Processing*, Springer, New York, USA.
- Elad, M. and Aharon, M., 2006, Image denoising via learned dictionaries and sparse representation, *International Conference on Computer Vision and pattern Recognition*, New-York, June 17–22.
- Engan, K., Skretting, K. and Husøy, J. H., 2007, Family of iterative LS-based dictionary learning algorithms, ILS-DLA, for sparse signal representation, *Digital Signal Process.*, vol. 17, no. 1, 32–49.
- Haldorsen, H. H., Damsleth, E., 1990, Stochastic modelling, *Journal of Petroleum Technology*, 42 (4), 404–412.
- Hastie, T., Tibshirani, R., Friedman, J., 2009, *The elements of statistical learning*, 2nd Edition, *Springer Series in Statistics Springer New York Inc., New York, NY, USA*
- Honarkhah, M., 2011, *Stochastic simulation of patterns using distance-based pattern modeling*, Ph.D. thesis, Stanford University.
- Horev, I., Bryt, O., and Rubinstein, R., 2012, Adaptive image compression using sparse dictionaries. In *International Conference on Systems, Signals and Image Processing*, 592–595, IEEE.
- Lie, K. A., 2015, *An Introduction to Reservoir Simulation Using MATLAB: User guide for the Matlab Reservoir Simulation Toolbox (MRST)*, SINTEF ICT.
- Mairal, J., Bach, F., Ponce, J., Sapiro, G., 2009, Online dictionary



شکل ۷: نتایج مجموعه آزمایش‌های دیگر که از نتایج این مطالعه حمایت می‌کند. نقاط قرمز مربوط به نتایج آزمایش‌های ارائه شده در این مقاله است.

۵- نتیجه‌گیری

گردش کار مدل سازی و شناسایی مخزن فشرده و تنک معرفی، ارائه، و روی مدل‌های نیمه-صنعتی یکی از مخازن نفتی ایران اعمال شد. این گردش کار به منظور بدست آوردن ترجمان نزدیکی از مدل اصلی، که نامعلوم است، از میان مجموعه مدل‌هایی که هر کدام تبدیلی از مدل اصلی هستند، برقرار شده است. بلوک‌های اصلی تشکیل دهنده این الگوریتم، مدل سازی زمین آمار چند نقطه‌ای و الگوریتم‌های تقریب تنک هستند. الگوریتم زمین آمار چند نقطه‌ای مجموعه مدل‌های زمین آمار چند نقطه‌ای را ایجاد می‌کنند که هر کدام تبدیلی از مدل اصلی است؛ این مجموعه مدل‌ها به عنوان مجموعه آموزشی در الگوریتم تقریب تنک برای تولید تصویر مدل فشرده تنک نهایی بکار برده می‌شود. استخراج خواص تخلخل و اشباع آب مخزن تحت مطالعه از روی داده‌های لرزه‌ای و وارون سازی شده به صورت قطعی انجام شد.

از الگوریتم DisPat به منظور تولید مجموعه‌ای از تحقیق‌های تصادفی بر اساس تصویر آموزشی ایستا استفاده شد. مجموعه مشابهی از تحقیق‌ها برای سیستم تصویر آموزشی سیستم دلتایی نایستا ایجاد شد. یک مدل تزریق دو بعدی بر اساس هر دو نوع تصویر آموزشی طراحی و آزمایش شد و نتایج شبیه سازی به عنوان اساس مقایسه بکار برده شدند. همین آزمایش‌ها برای هر کدام از تحقیق‌های ایجاد شده در دو مجموعه تصاویر ایستا و نایستا انجام شد و فضای نمونه‌ای مرتبط با نتایج تطابق تاریخچه تشکیل شد. الگوریتم تعلیم لغت نامه ILS-DLA در روبه فشرده سازی تصویر تنک استفاده شد و مدل‌های فشرده نهایی بدست آمده در هر دو مورد مجموعه‌های ایستا و نایستا در فرآیند شبیه سازی مخزن بکار برده شدند و نتایج تطابق تاریخچه آنها مورد بررسی قرار گرفتند. بر اساس فضای نمونه‌ای حاصل از ایجاد تحقیقات زمین آمار چند

- dictionaries by RLS-DLA and compared with K-SVD, in Proceedings ICASSP 2011, Pargue, Czech Republic, June 2011, vol. 1, 1517-1520.
- Skretting, K., Engan, K., 2011b, Learned dictionaries for sparse image representation: properties and results, Proceedings of the SPIE, Volume 8138, id. 81381N.
- Skretting, K., Engan, K., 2010, Recursive least squares dictionary learning algorithm, IEEE Transactions on Signal Processing, vol. 58, 2121-2130
- Skretting, K., Husøy, J.H., 2003, Partial search vector selection for sparse signal representation, NORSIG-03, Bergen, Norway.
- Skretting, K., Husøy, J.H., Aase, S.O., 1999, Improved Huffman coding using recursive splitting, NORSIG-99 Conference on Signal Processing and Signal Compression.
- Starck, J.L., Murtagh, F., Fadili, J., 2010, Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity, Cambridge: Cambridge University Press.
- Strebelle, S., 2000, Sequential simulation drawing structures from training images, Ph.D. thesis, Stanford University.
- Tibshirani, R., 1996, Regression Shrinkage and Selection via the LASSO, J. Royal Statistical Soc. B, vol. 58, no. 1, 267-288.
- learning for sparse coding, in ICML '09: Proceedings of the 26th Annual International Conference on Machine Learning, New York, NY, USA, 689-696, ACM.
- Mairal, J., Elad, M., and Sapiro, G., 2008, Sparse representation for color image restoration, IEEE Trans. on Image Processing, 17(1), 53-69.
- Mariethoz, G., Caers, J., 2015, Multiple-Point Geostatistics: Stochastic Modeling with Training Images. John Wiley & Sons, Ltd, Chichester, UK.
- Rubinstein, R., Bruckstein, A.M., and Elad, M., 2010a, Dictionaries for sparse representation modeling, Proc. IEEE, vol. 98, no. 6, 1045-1057.
- Rubinstein, R., Peleg, T., and Elad, M., 2013, Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model, IEEE Trans. Signal Process., vol. 61, no. 3, 661-677.
- Rubinstein, R., Zibulevsky, M., and Elad, M., 2010b, Double sparsity: Learning sparse dictionaries for sparse signal approximation, IEEE Trans. Signal Process., vol. 58, no. 3, 1553-1564
- Rubinstein, R., Zibulevsky, M., and Elad, M., 2008, Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit, Technical report.
- Skretting, K., Engan, K., 2011a, Image compression using learned



Approximation of the reservoir permeability model using ILS-DLA learning algorithm and LARS sparse coding method

Mohammad Hosseini¹ and Mohammad Ali Riahi^{2*}

1- Ph.D. Candidate, Institute of Geophysics, University of Tehran, Tehran, Iran

2- Professor, Institute of Geophysics, University of Tehran, Tehran, Iran

Received: 7 March 2018; Accepted: 9 December 2018

Corresponding author: mariahi@ut.ac.ir

Keywords

Sparse approximation

Dictionary learning

Sparse coding

Multiple-point geostatistics

Reservoir characterization

Seismic inversion

Reservoir simulation

Summary

In this paper, the task is to return from a set of multiplicities from a model to obtain an approximation of that model using sparse approximation. The term 'approximation' indicate the sufficiency of an interpretation that is close enough to the true mode, i.e. reality. In geosciences, the multiplicities are provided by multiple-point statistical (MPS) methods. Realistic modeling of the earth interior demands for more sophisticated geostatistical methods based on true available images, i.e. the training images. Among the available MPS methods, the DisPat algorithm is a distance-based MPS method, which generates appealing realizations for stationary and non-stationary training

images by classifying the patterns based on distance functions using kernel methods. Advances in non-stationary image modeling is an advantage of the DisPat method. Realizations generated by the MPS methods form the training set for the sparse approximation. The sparse approximation is comprising of two steps, sparse coding and dictionary update, which are alternately used to optimize the trained dictionary. Model selection algorithms like LARS are used for sparse coding. LARS optimizes the regression model sequentially by choosing a proper number of variables and adding the best variable to the active set in each iteration. The ILS-DLA dictionary learning algorithm addresses the internal structure of the dictionary by considering the overlapping or non-overlapping blocks and the inversion task according to the internal structure of the trained dictionary. The ILS-DLA is fast in the sense that it inverts smaller blocks constructing the trained dictionary rather than inverting the entire dictionary. The trained dictionary is sequentially updated by alternating between sparse coding and dictionary training steps. According to the experiments, the compressed sparsity-based image model is superior to 90% of the generated realizations by 90% probability.

Introduction

In reservoir characterization and modeling, providing a large set of population samples is practically impossible unless introducing stochasticity in the models using the MPS methods. DisPat algorithm is one the MPS methods that is robust for integrating data based on visual system of the human being and representing new algorithms for modeling images with non-stationary properties. The task is to achieve a single model image as the representation of the true model from a large set of MPS realizations, which are considered as the manipulations of the true model. The mathematical tools to perform such a task is known as the sparse approximation.

Methodology and Approaches

DisPat is a distance-based pattern-based multiple-point geostatistical method. In DisPat, the training image is scanned by the designed template and the patterns are extracted from the training image. The patterns are then classified based on distance functions using kernel methods. A distance function measures the distance between each pair of patterns in the metric space. Any two close points in the metric space refer to two similar patterns from the pattern database. The kernel k -means clustering algorithm is used to classify the patterns in the pattern database.

The sparse approximation is a sequential alternation between two steps: sparse coding and dictionary update. The LARS algorithm is used for the sparse coding that is a model selection algorithm used to select a parsimonious set of covariates among a larger set of them. The ILS-DLA algorithm is used for dictionary updating that solves the sparse approximation problem according to the internal structure of the dictionary. The dictionary matrix is blocked into a definite number of sub-matrices, which are either blocky or overlapping. The inversion is performed for each block rather than the whole dictionary, and therefore, it is a fast algorithm.

Results and Conclusions

The three-dimensional (3D) cube seismic data were deterministically inverted and acoustic impedance (AI), porosity, and saturation cubes were obtained. The spectral decomposition cube was generated and studied. Two reality-based permeability models were selected from the spectral decomposition cube based on which, many MPS realizations were generated using DisPat-MPS methodologies. The ILS-DLA learning algorithm was applied on the set of realizations along with the LARS algorithm for sparse coding. The resultant picture model was a compressed model that is superior to 90% of the model images in the training sample by 89.58% probability.
