

طراحی یک سیستم فشرده سازی/بازسازی تصاویر متنی با درجه‌ی تفکیک مکانی بالا مبتنی بر فرا تفکیک پذیری

سعید مرادی^۱، فارغ التحصیل کارشناسی ارشد الکترونیک، هادی گرایلو^۲، استادیار

۱- دانشکده مهندسی برق - دانشگاه صنعتی شاهرود - شاهرود - ایران - saeid006@gmail.com

۲- دانشکده مهندسی برق - دانشگاه صنعتی شاهرود - شاهرود - ایران - grailu@shahroodut.ac.ir

چکیده: در این مقاله یک سیستم فشرده سازی/بازسازی تصاویر متنی با درجه‌ی تفکیک مکانی بالا مبتنی بر فرا تفکیک پذیری پیشنهاد شده است. در روش پیشنهادی، برای رسیدن به میزان فشرده سازی بیشتر از ایده کاهش ابعاد در تصاویر متنی استفاده شده است. کاهش ابعاد در کنار عمل فشرده سازی ممکن است باعث تنزل در کیفیت تصویر شود. بنابراین باید روشی انتخاب شود که واحد بازسازی بتواند در کنار افزایش ابعاد تصویر، اثرات مخرب تأثیر گذار بر تصویر را نیز اصلاح کند. در مرحله بازسازی از روش فرا تفکیک پذیری استفاده شده است. در این روش، تصویر و ضوح پایین ورودی به سه لایه تقسیم و سپس هر لایه براساس اهمیت اطلاعاتی آن با یک روش خاص بزرگ‌نمایی شده است. در نهایت لایه‌های بزرگ‌نمایی شده با هم ترکیب و تصویر و ضوح بالای نهایی تشکیل شده است. یک ویژگی مهم روش پیشنهادی، قابلیت ترکیب آن با روش‌های فشرده سازی مختلف است. در این مقاله، ترکیب روش پیشنهادی با هر یک از روش‌های فشرده سازی JPEG، JPEG2000 و SPIHT بررسی و ملاحظه می شود، جواب قابل قبولی از نظر معیارهای باز شناسی متن (OCR) و متوسط امتیاز نظر سنجی (MOS) بدست آمده است؛ گرچه از نظر معیار قله سیگنال به نویز (PSNR) روش‌های دیگر بهتر از روش پیشنهادی عمل کرده‌اند.

واژه‌های کلیدی: فشرده سازی تصاویر متنی، فشرده سازی JPEG، فشرده سازی JPEG2000، فشرده سازی SPIHT، فراتفکیک پذیری، شناسایی نوری کاراکترها.

Designing a Codec System for High Resolution Textual Images Based on Super Resolution

Saeid Moradi, Graduated¹, Hadi Grailu, Assistant Professor²

1- Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran, Email: saeid006@gmail.com

2- Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran, Email: grailu@shahroodut.ac.ir

Abstract: In this paper, a CODEC system based on super resolution, is proposed for compression of high resolution textual images. It employs image resizing to decrease image dimensions and consequently, to improve the compression ratio; but at the other hand, it may reduce the image quality. Therefore, the decompression unit employs super resolution to simultaneously increase the reconstructed image dimensions and quality. In the employed interpolation-based super-resolution method, using an efficient textual image matting algorithm, the input low-resolution textual image is decomposed into three layers after which, each layer is enlarged using a particular method. Finally, the enlarged layers are combined to build the high resolution reconstructed textual image. An interesting feature of the proposed method is the ability to use existing compression methods such as JPEG, JPEG2000 and SPIHT. We have employed the aforementioned compression methods in the proposed CODEC system and evaluated the compression results with respect to OCR rate, Mean Opinion Score (MOS), and PSNR measures. Considering the OCR and MOS measures, the proposed method outperformed the others but not so with respect to PSNR.

Keywords: Textual image compression, JPEG compression, JPEG2000 compression, SPIHT compression, super resolution, optical character recognition.

تاریخ ارسال مقاله: ۱۴۰۰/۰۲/۱۵

تاریخ اصلاح مقاله: ۱۴۰۰/۱۱/۲۶

تاریخ پذیرش مقاله: ۱۴۰۰/۱۲/۱۳

نام نویسنده مسئول: هادی گرایلو

نشانی نویسنده مسئول: ایران - شاهرود - دانشگاه صنعتی شاهرود - دانشکده مهندسی برق - گروه الکترونیک.

۱- مقدمه

حجم عظیم تصاویر اسناد متنی که به طور روزانه در مراکزی مانند کتابخانه‌های دیجیتال و مراکز بایگانی اسناد در حال تولید است، سبب می‌شود که نگهداری و مدیریت آن‌ها چالشی مهم محسوب شود. این حجم عظیم از داده‌ها در کنار محدود و پرهزینه بودن حافظه‌های دیجیتالی امروزی، موجب توجه جدی به روش‌های تخصصی فشرده سازی تصاویر اسناد متنی شده است.

روش‌های فشرده سازی به دو دسته‌ی بااتلاف و بدون اتلاف تقسیم می‌شوند. در کاربردهایی که اشاره شد، عمدتاً از روش‌های فشرده سازی بااتلاف استفاده می‌شود زیرا هم میزان فشرده سازی بسیار بالاتری نسبت به روش‌های بدون اتلاف فراهم می‌کنند و هم این که کیفیت تصویر فشرده شده تا حدی، قابل کنترل است. در این مقاله از سه روش فشرده سازی JPEG، JPEG2000 و SPIHT در ترکیب با روش پیشنهادی استفاده شده است. هر کدام از این روش‌ها در نرخ بیت‌های پایین می‌توانند موجب اثرات نامطلوبی مانند اثرات بلوکی یا نشتی روی تصاویر متنی شوند. از طرف دیگر در دو روش JPEG و JPEG2000 حداکثر میزان فشرده سازی قابل حصول محدود است. با توجه به اینکه تصاویر مورد بررسی، تصاویر با درجه‌ی تفکیک مکانی بالا بوده است و این تصاویر حجم فایل زیادی برای ذخیره سازی اشغال می‌کنند، هدف این است تا جایی که امکان دارد حجم فایل تصویر فشرده شده جهت ذخیره سازی کمتر شود.

با توجه به مشکلات ذکر شده برای روش‌های فشرده سازی استفاده شده، به منظور رسیدن به نرخ بیت‌های کمتر، ایده‌ی اصلی مورد استفاده در روش پیشنهادی کاهش ابعاد تصاویر متنی ورودی است؛ ابتدا تصویر ورودی را کاهش ابعاد داده و سپس آن را فشرده می‌کنیم. باید توجه داشت که روش‌های کاهش ابعاد نیز ممکن است موجب اثرات نامطلوبی در تصویر شوند. بنابراین در مرحله بازسازی نیاز به روشی است که بتواند هم اثرات نامطلوب ناشی از فشرده سازی و کاهش ابعاد را بهبود دهد و هم ابعاد تصویر را به ابعاد اولیه بازگرداند. برای این منظور در مرحله بازسازی از روش فراتفکیک پذیری استفاده شده است.

۲- مرور روش‌های موجود

یک روش فشرده سازی تصاویر مبتنی بر روش فشرده سازی فراکتالی با استفاده از شبکه‌های عصبی مصنوعی در [۱] پیشنهاد شده است. استفاده از شبکه عصبی به منظور سرعت بخشیدن به روش فشرده سازی فراکتالی در مرحله فشرده سازی بوده است. یک روش فشرده سازی بدون اتلاف تصاویر متنی مبتنی بر ترکیب روش کدگذاری طول اجرا* و روش کدگذاری هافمن در [۲] پیشنهاد شده است.

روش‌های امروزی موجود برای فشرده سازی تصاویر طبیعی را نمی‌توان برای فشرده سازی تصاویر اسناد با نرخ فشرده سازی زیاد مورد توجه قرار

داد، زیرا چنین روش‌هایی از ویژگی‌های مخصوص به متن برای افزایش نرخ فشرده سازی استفاده نمی‌کنند.

در تصاویر متنی عمده‌ی اطلاعات در لبه‌ها قرار داشته و نواحی داخلی نویسه‌ها و علائم متنی چندان حاوی اطلاعات نمی‌باشند. اما در بیشتر تصاویر طبیعی، قسمت مهمی از اطلاعات در بدنه‌ی اشیاء موجود در تصویر قرار دارد. پس می‌توان گفت که، طیف این نوع از تصاویر، طیفی میان‌گذر -یا شامل مولفه‌های فرکانسی میانی مهم- است اما طیف تصاویر طبیعی، پایین‌تر است [۳]. یک ویژگی مهم دیگر تصاویر متنی این است که در این تصاویر، برخلاف بیشتر تصاویر طبیعی، درجه‌ی تفکیک مکانی بیشتر از درجه‌ی تفکیک مقادیر شدت روشنایی در حفظ کیفیت یا اطلاعات اولیه اهمیت دارد [۴]. تفاوت دیگر بین تصاویر متنی و تصاویر طبیعی این است که تصاویر متنی، برخلاف بیشتر تصاویر طبیعی، شامل جزئیات دیداری مهم مانند شکل و انحنا، رنگ، بافت کاغذ و بافت کاغذ هستند. این مسأله در مورد تصاویر متنی تاریخی، تصاویر متنی دست‌نویس، و نیز تصاویر شامل جداول و روابط ریاضی اهمیت ویژه‌ای پیدا می‌کند.

ویژگی‌های تصاویر متنی آن قدر مهم و متمایز از دیگر انواع تصاویر است که بخش ۶ از استاندارد JPEG2000 ویژه تصاویر متنی بوده و پسوند فایل JPM را به خود اختصاص داده است [۵].

۲-۱- فشرده سازی تصاویر متنی

روش‌های فشرده سازی تصاویر متنی را می‌توان به دو دسته‌ی روش‌های مبتنی بر مدل و روش‌های غیرمبتنی بر مدل تقسیم‌بندی کرد. روش‌های مبتنی بر مدل، از مدل محتوای ترکیبی (MRC) استفاده نموده و تصویر متنی را به لایه‌های مختلفی تجزیه می‌کنند. روش‌های غیرمبتنی بر مدل، عمدتاً به دنبال ارتقاء کارایی فشرده سازی یکی از روش‌ها یا استانداردهای نسبتاً عمومی موجود به منظور استفاده در کاربرد فشرده سازی تصاویر متنی می‌باشند.

روش‌های فشرده سازی مبتنی بر تبدیل موجک جزو روش‌های امروزی و موثر محسوب می‌شوند [۶]. امروزه مشخص شده است که ترکیب تبدیل موجک و کدگذاری بخش‌بندی مجموعه (SPC) [۸،۷] در حالت کلی، در بیشتر کاربردها کارایی بیشتری نسبت به دیگر انواع تبدیل‌ها (مانند DCT) و کدگذاری‌ها (مانند هافمن و حسابی وقفی) دارد [۶-۸]. کدگذاری SPC عموماً از ویژگی وجود شباهت در زیرباندهای هم‌راستا در تجزیه‌ی هرمی موجک بهره می‌برند. عملکرد این روش‌ها عمدتاً مبتنی بر روشی قانون‌مند برای جستجوی مهم‌ترین ضرایب موجک و تصمیم‌گیری در مورد کدگذاری مناسب آنها است طوری که واحد کدگشا، با در دست داشتن نتایج این جستجو و تصمیم‌گیری، قادر به بازسازی ماتریس ضرایب موجک و سپس، اعمال عکس تبدیل موجک باشد [۹]. از جمله مهم‌ترین روش‌های کدگذاری SPC، روش کدگذاری SPIHT است [۱۰]. از این کدگذار تاکنون در

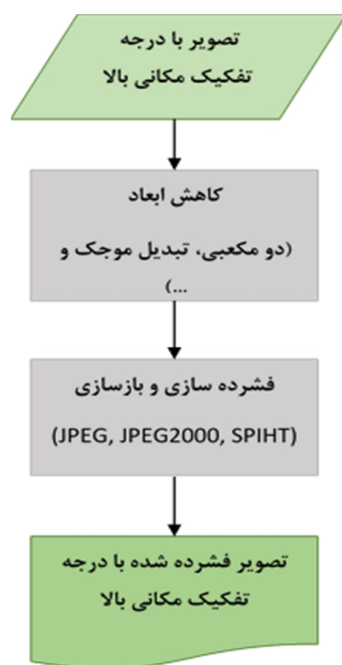
فشرده‌سازی تصاویر مختلفی مانند تصاویر هوایی، تصاویر پزشکی، تصاویر عمومی و نیز برای سیگنال‌هایی مانند سیگنال‌های قلبی و مغزی استفاده شده است [۸]. یک روش فشرده‌سازی تصاویر متنی مبتنی بر تبدیل موجک و کدگذاری SPIHT به‌منظور حفظ یا بهبود کیفیت تصویر، خوانایی و تشخیص نوری کاراکترها در [۱۱] پیشنهاد شده‌است.

در [۱۲] یک روش قطعه‌بندی بلوکی تصاویر متنی به منظور فشرده‌سازی این نوع از تصاویر معرفی شده است. در این روش از استاندارد JPEG برای کدگذاری استفاده شده است. روش پیشنهاد شده در [۱۳] به دنبال کاهش یا حذف اثرات نامطلوب ایجاد شده از اعمال استاندارد JPEG که ویژه‌ی فشرده‌سازی تصاویر طبیعی است، روی تصاویر متنی می‌باشد. استانداردهایی که تاکنون در کاربرد فشرده‌سازی تصاویر متنی استفاده شده‌اند، محدود به استانداردهای فشرده‌سازی تصاویر نمی‌شوند. برای نمونه، در [۱۴] از کدک ویدیویی H.264/AVC برای فشرده‌سازی تصاویر متنی استفاده شده است. یک روش فشرده‌سازی تصاویر متنی در نرخ بیت‌های پایین براساس کدگذاری ناحیه مورد علاقه در [۱۵] ارائه شده‌است. در این روش از کدگذار SPIHT در چارچوب فشرده‌سازی ناحیه مورد علاقه استفاده شده‌است. به‌منظور بهبود اثرات ناشی حاصل از روش SPIHT از روش‌های بهبود تصاویر استفاده شده‌است.

در تصاویر متنی، کیفیت لبه‌های بازیابی شده در تصاویر بازسازی شده (پس از فشرده‌سازی) مهم است. روش‌های عمومی فشرده‌سازی تصاویر و نیز برخی روش‌های تخصصی فشرده‌سازی تصاویر متنی به این نکته توجهی نشان نداده‌اند. برای مثال در فشرده‌سازی به کمک استاندارد JPEG، اثر بلوکی و حذف فرکانس‌های میانی به بالا مشاهده می‌شود. لبه‌های تصاویر متنی عمدتاً متناظر با فرکانس‌های میانی می‌باشند. در خروجی روش JPEG که مبتنی بر تبدیل DCT است، به دلیل بلوک‌بندی تصویر، اثر بلوکی شدن مشاهده می‌شود. همچنین این روش به دلیل کوانتیزاسیون انجام شده در آن، عملکردی معادل با یک فیلتر پایین‌گذر دارد [۳].

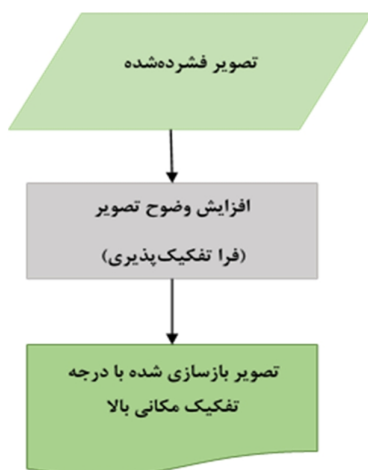
در استاندارد JPEG2000 از تبدیل موجک و کدگذار EBCOT که نوعی کدگذار بخش‌بندی مجموعه است [۱۶]، استفاده شده است. موجک‌ها از ویژگی حضور محلی و نیز ویژگی متراکم‌سازی انرژی برخوردار هستند؛ بنابراین در JPEG2000، میزان فشرده‌سازی بیشتر شده و کیفیت لبه‌ها نیز بهتر حفظ می‌شود؛ اما متأسفانه به دلیل وقوع اثر ناشی در نرخ بیت‌های پایین، کیفیت لبه‌ها و نیز کیفیت دیداری تصویب، آفت پیدا می‌کنند [۳].

در روش‌های فشرده‌سازی JPEG2000 و SPIHT مهم‌ترین عامل محدودیت ساز در رسیدن به مقادیر بزرگ میزان فشرده‌سازی، اثر نشیمنی است. لذا اگر بتوان این عامل افت کیفیت را با برخی روش‌ها



شکل ۱. مرحله فشرده سازی روش پیشنهادی

دهد. مطابق با شکل ۲، تصویر فشرده شده بعنوان ورودی مرحله بازسازی در نظر گرفته می شود. در تنها بلوک مهم مورد استفاده مرحله بازسازی، به منظور افزایش وضوح و نیز افزایش ابعاد تصویر فشرده شده از روش فرا تفکیک پذیری مبتنی بر درون یابی استفاده می شود. در مرحله بازسازی از روش فرا تفکیک پذیری مبتنی بر درون یابی استفاده شده است. در روش استفاده شده ابتدا تصویر ورودی به سه زیر لایه شامل؛ (۱) لایه شامل اطلاعات لایه تصویر متن (۲) لایه پیش زمینه و (۳) لایه پس زمینه تجزیه می شود. سپس با توجه به اهمیت هر لایه، آن لایه به روشی خاص بزرگ نمایی می شود. سپس لایه ها با هم ترکیب شده و تصویر وضوح بالای نهایی بدست می آید.



شکل ۲. مرحله برگشت

۲-۲-۲ روش های مبتنی بر درون یابی

این روش ها، روش های ساده و از لحاظ محاسباتی، روش های مؤثری هستند. در این روش ها مدل تصویر برداری ساده در نظر گرفته می شود [۲۱، ۲۰]. در واقع در این روش ها به نوعی ترکیبی از الگوریتم های درون یابی و بهبود تصاویر را داریم. عبارت دیگر یا تصویر را با روش های معمول درون یابی مانند دو مکعبی بزرگ می کنند و سپس تصویر بزرگ شده را با تأکید بر لایه های متن بهبود می دهند و یا برعکس، ابتدا تصویر را بهبود می دهند و سپس بزرگ نمایی را انجام می دهند.

اخیراً از روش های درون یابی مبتنی بر ویژگی های آماری لایه به منظور فرا تفکیک پذیری مبتنی بر درون یابی استفاده شده است. اما این روش ها، روش هایی زمان بر هستند [۲۲]. در [۲۳] روش افزایش ابعاد ساده اما مؤثری برای بهبود خودکار وضوح تصویر/ویدئو، در حالی که اطلاعات ساختاری اساسی را حفظ می کند، پیشنهاد شده است. مزیت اصلی روش پیشنهاد شده چارچوب بازخورد کنترلی است که به خوبی اطلاعات تصویر وضوح بالا را بدون در نظر گرفتن هیچ آموزشی از داده ورودی باز یابی می کند. این امر باعث می شود که روش پیشنهادی هیچ گونه وابستگی به کیفیت و تعداد نمونه های انتخاب شده که مسئله رایج در روش های مبتنی بر آموزش است، نداشته باشد. مزیت دیگر روش پیشنهاد شده در [۲۳] سرعت بالای آن است.

در [۲۴] یک الگوریتم جدید زمان واقعی لایه بندی مبتنی بر مدل خطی-محلی به منظور تجزیه تصویر متنی ورودی به لایه های مختلف آن طراحی شده است. در این روش برای افزایش ابعاد هر لایه یک روش بزرگ نمایی خاص در نظر گرفته شده است. در نهایت تصویر متنی با وضوح بالای نهایی حاصل از فرا تفکیک پذیری با ترکیب لایه های وضوح بالای ایجاد شده، بدست می آید.

۳- روش پیشنهادی

روش پیشنهادی شامل دو بخش، مرحله اول شامل فشرده سازی تصویر متنی با درجه تفکیک مکانی بالا (مرحله فشرده سازی) و مرحله دوم شامل بازسازی تصویر متنی فشرده شده و سپس رسیدن به تصویر با درجه تفکیک مکانی بالای اولیه است (مرحله بازسازی). در مرحله فشرده سازی مطابق با شکل ۱، ابتدا تصویر ورودی که با درجه تفکیک مکانی بالا روبش شده است بوسیله یکی از روش های کاهش ابعاد، ابعادش کاهش داده می شود. سپس، تصویر حاصل از مرحله کاهش ابعاد با یکی از روش های فشرده سازی، فشرده می شود. از جمله مزیت های روش پیشنهادی قابلیت ترکیب این روش با انواع مختلف روش های فشرده سازی موجود است.

در رسیدن به نرخ بیت های پایین، اعمال فشرده سازی و تکنیک پیشنهادی کاهش ابعاد، اثرات مخرب و نامطلوبی روی تصویر متنی ورودی ایجاد می کنند. بنابراین در مسیر بازسازی نیاز به روشی است که بتواند بطور همزمان هم اثرات مخرب و نامطلوب ایجاد شده را بهبود دهد و هم ابعاد تصویر فشرده شده را به ابعاد تصویر ورودی اولیه افزایش

۳-۱- لایه بندی تصویر

لایه بندی تصویر بیان می کند که تصویر I می تواند به صورت مجموع وزن داری از تصویر پیش زمینه F و تصویر پس زمینه B با استفاده از لایه بند آلفا α به عنوان وزن مطابق رابطه (۱) مدل شود [۲۵].

$$I = \alpha F + (1 - \alpha)B \quad (1)$$

در رابطه (۱)، α ماتریسی است که هر پیکسل آن می تواند مقداری در بازه $[0, 1]$ داشته باشد. اگر $\alpha_{i,j} = 0$ در این صورت پیکسل واقع در مختصات (i, j) پس زمینه و اگر $\alpha_{i,j} = 1$ پیکسل واقع در مختصات (i, j) را پیش زمینه می نامند. در غیر این صورت پیکسل ترکیبی از پیش زمینه و پس زمینه است. لایه بند تصویر با هدف بدست آوردن لایه آلفا α ، پیش زمینه F و پس زمینه B روی تصویر ورودی I اعمال می شود.

شکل ۳ مثالی از لایه بندی تصویر متنی را نشان می دهد. شکل ۳(الف) تصویر ورودی را نشان می دهد. شکل ۳(ب) لایه آلفا را بعد از لایه بندی تصویر نشان می دهد. شکل های ۳(ج) و ۳(د) به ترتیب لایه های پیش زمینه و پس زمینه را نشان می دهند. لایه آلفا شامل اطلاعات مهم مربوط به نواحی متنی و لایه های پیش زمینه و پس زمینه نیز به ترتیب شامل اطلاعات مربوط به پیش زمینه و پس زمینه هستند. همان طور که از شکل ۳ مشخص است، لایه آلفا شامل نواحی هموار متن می باشد. لایه های پیش زمینه و پس زمینه تنها شامل اطلاعات رنگ متن و پس زمینه هستند. این نکته برای فرا تفکیک پذیری تصاویر متنی بسیار مفید است و به جلوگیری از مصنوعات و ماتی با پردازش جداگانه روی لایه متن و اجزای رنگ کمک می کند.

همان طور که گفته شد اولین مرحله از روش ارائه شده، تجزیه تصویر ورودی I' به سه لایه متفاوت شامل لایه وضوح پایین رنگ پیش زمینه F' ، لایه وضوح پایین رنگ پس زمینه B' و لایه وضوح



شکل ۳ مثالی از لایه بندی تصویر متنی. (الف) تصویر ورودی. (ب) لایه آلفا. (ج) لایه پیش زمینه. (د) لایه پس زمینه. [۲۵]

Linear Ridge Regression ‡

۳-۱-۱- لایه شامل اطلاعات لایه متن

ابتدا تصویر ورودی را به یک تصویر دودویی تبدیل می کنیم. تصویر دودویی A به دست آمده به منظور تصویر پوشش متن در نظر گرفته می شود. [۲۵]. به منظور استخراج اطلاعات لایه تصویر متن اصلی، از فیلتر حفظ لایه یعنی فیلتر هدایت [۲۷] استفاده می شود. کلید اصلی فیلتر هدایت مدل خطی محلی بین تصویر راهنما I و تصویر خروجی q است. فرض می شود q تبدیل خطی از I در پنجره w_k به مرکز k باشد:

$$q_i = a_k I_i + b_k, \quad \forall i \in w_k \quad (2)$$

در رابطه (۲) a_k و b_k ضرایب خطی ثابت در w_k هستند. پنجره w_k به صورت مربعی و به شعاع r در نظر گرفته شده است. مدل خطی محلی اطمینان می دهد که خروجی فیلتر یعنی q دارای لایه می باشد تنها اگر I لایه داشته باشد، زیرا $\nabla q = a \nabla I$. برای تعیین ضرایب خطی a_k و b_k به ورودی p فیلتر نیاز داریم. خروجی q بصورت تفریق المان های ناخواسته n مانند نویز از ورودی p مدل می شود.

$$q_i = p_i - n_i \quad (3)$$

راه حلی که در نظر گرفته شده است، روشی است که در رابطه (۳) اختلاف بین p و q را کمینه کند در حالیکه مدل خطی (۲) را شامل شود. به طور خاص، تابع هزینه ارائه شده در رابطه (۴) در پنجره w_k کمینه شده است.

$$E(a_k, b_k) = \sum_{i \in w_k} ((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2) \quad (4)$$

ε پارامتر تنظیم جهت a_k های بزرگ است. رابطه (۵) رگرسیون خطی ستیغی[‡] است که در [۲۸، ۲۹] مدل شده است و راه حل آن بصورت زیر است:

$$a_k = \frac{1}{|w| \sum_{i \in w_k} I_i p_i - \mu_k \overline{p_k}} \quad (5)$$

$$b_k = \overline{p_k} - a_k \mu_k \quad (6)$$

در روابط (۵) و (۶)، μ_k و σ_k^2 به ترتیب میانگین و واریانس I در w_k هستند. $|w|$ تعداد پیکسل ها در w_k و $\overline{p_k} = 1/|w| \sum_{i \in w_k} p_i$ میانگین

P در w_k است. با بدست آوردن ضرایب خطی a_k و b_k می توان خروجی q_i فیلتر را با استفاده از رابطه (۲) محاسبه کنیم. شکل ۴ فرایند

با کمینه کردن خطای باز سازی و عبارت تنظیم، باز سازی پیش زمینه و پس زمینه می تواند با کمینه کردن تابع انرژی رابطه (۱۲) حل شود [۲۵]:

$$\sum_{i \in W_k} (((F_k - B_k) \alpha_i + B_k - I_i)^2 + \gamma (F_k - B_k)^2) / \gamma (F_k - B_k)^2 \quad (12)$$

رابطه (۱۲) بیان گر عبارت تنظیم است و γ ثابتی است که معمولاً برابر ۰.۰۱ تنظیم می شود. تابع انرژی نمایش داده شده در رابطه (۱۲) می تواند در شکل بسته حل شود و حل آن به صورت رابطه (۱۳) و (۱۴) است [۳۰]:

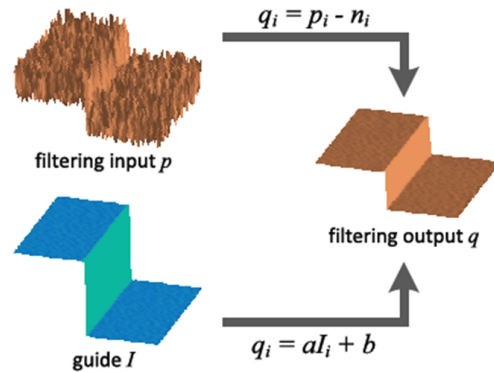
$$F_k - B_k = \frac{1}{\sigma_k^2 + \gamma} \sum_{i \in W_k} \alpha_i I_i - \mu_k \bar{I}_k \quad (13)$$

$$B_k = \bar{I}_k - (F_k - B_k) \mu_k \quad (14)$$

در رابطه (۱۳) W_k پنجره 3×3 است زیرا استفاده از پنجره بزرگ تر ممکن است فرض خطی محلی را نقض کند. μ_k و σ_k^2 در روابط (۱۳) و (۱۴) به ترتیب میانگین و واریانس α در W_k می باشند، تعداد پیکسل ها در پنجره محلی است، i شماره پیکسل ها در همسایگی W_k و $\bar{I}_k$ میانگین شدت روشنایی پیکسل های I در همسایگی W_k می باشد.

۳-۲- بزرگ نمایی جداگانه لایه ها و ترکیب آن ها

این قسمت از روش فرا تفکیک پذیری شامل افزایش ابعاد لایه های به دست آمده از قسمت قبل است. همان طور که در شکل ۳ (ج) و ۳ (د) دیده می شود، لایه های پیش زمینه و پس زمینه به طور معمول هموار هستند و اطلاعات مفید کمتری دارند در حالی که لایه های حاوی اطلاعات لایه متن نشان داده شده در شکل ۳ (ب) شامل عمده اطلاعات مربوط به شکل متن است. بنابراین برای افزایش ابعاد هر یک از سه لایه پیش زمینه، پس زمینه و لایه حاوی اطلاعات متن آلفا، روش های مختلفی در نظر گرفته می شود. با توجه به اهمیت لایه شامل اطلاعات لایه متن، ابعاد این لایه با تأکید بر تقویت لبه ها بزرگ شده است. از طرف دیگر به دلیل اینکه لایه های پیش زمینه و پس زمینه از اهمیت اطلاعاتی کمی در رابطه با ویژگی های متن برخوردار هستند، از روش درون یابی دومکعبی برای بزرگ نمایی آنها استفاده می شود. چون درون یابی دومکعبی لبه های تصویر متنی را مات می کند بنابراین لایه های حاوی اطلاعات لایه متن را قبل از درون یابی ابتدا بهبود می دهیم. برای بهبود لبه های تصویر متنی از فیلتر تیگر [۳۱] استفاده می کنیم. فیلتر تیگر، فیلتری با تأکید بر تقویت لبه های تصویر متنی و از بین بردن نویز است. این فیلتر به منظور تقویت لبه های متن مورد استفاده قرار گرفته شده است. فیلتر تیگر برای یک پیکسل تصویر I به مختصات (i, j) در رابطه (۱۵) معرفی شده است:



شکل ۴ فیلتر هدایت

فیلتر هدایت را نشان می دهد. تصویر دودویی به دست آمده به عنوان تصویر اصلی و تصویر ورودی به عنوان تصویر راهنما به فیلتر داده می شود. فیلتر هدایت تا زمانی که براساس مدل خطی محلی است می تواند تصویر متن دودویی A را به عنوان تقریبی از لبه های تصویر متن اصلی در نظر بگیرد و اطمینان دهد که در نواحی محلی تغییرات تصویر ورودی و تصویر پوشش با هم برابر است یعنی $\nabla A = \nabla I'$. اگر G را معرف فیلتر هدایت در نظر بگیریم، لایه حاوی اطلاعات لایه متن a' را می توان از رابطه (۷) بدست آورد:

$$a' = G A I'; r, \epsilon \quad (7)$$

در رابطه (۷) دو پارامتر r و ϵ به ترتیب شعاع همسایگی و پارامتر تنظیم ضرایب فیلتر هدایت هستند [۲۷]. شکل ۳ (ب) لایه حاوی اطلاعات لایه متن a' را نشان می دهد.

۳-۱-۲- لایه های پیش زمینه و پس زمینه

مدل خطی محلی بیان می کند که در پنجره های محلی کوچک، کانال α می تواند به صورت ترکیب خطی از تصویر I نمایش داده شود.

$$\alpha_k = Z I_k^c + H \quad k \in \omega_i \quad (8)$$

در رابطه (۸) ω_i پنجره کوچکی حول پیکسل i و c بیان گر کانال رنگ است. H و G در پنجره محلی، ثابت می باشند. برای سادگی، در ادامه شاخص کانال c حذف می شود. براساس رابطه (۱) و (۸)، می توان Z و H را به صورت زیر بدست آورد. با جای گذاری رابطه (۸) در رابطه (۱) داریم:

$$I = Z I F + H F + B - Z I B - H B \quad (9)$$

با در نظر گرفتن $I = 0$ برای مقادیر پس زمینه و $I = 1$ برای مقادیر پیش زمینه و جایگذاری این مقادیر در رابطه بالا داریم:

$$I = 0 \rightarrow H(F - B) = -B \rightarrow H = \frac{B}{B - F} \quad (10)$$

$$I = 1 \rightarrow 1 = Z(F - B) + H(F - B) + B$$

$$1 = Z(F - B) + \frac{-B}{F - B}(F - B) + B \rightarrow Z = \frac{1}{F - B} \quad (11)$$

در این مقاله از سه روش فشرده سازی JPEG، JPEG2000 و SPIHT در ترکیب با روش پیشنهادی استفاده شده است. به منظور فشرده سازی تصاویر از نرم افزارهای Advanced JPEG Compressor، JPEG2000 Compressor و روش SPIHT [۱۰] استفاده شده است. به منظور ارزیابی روش پیشنهادی، معیارهای باز شنا سی متن (OCR)، متوسط امتیاز نظرسنجی (MOS) و پیک سیگنال به نویز (PSNR) مورد استفاده قرار گرفته اند.

معیار OCR

برای OCR، از نرم افزار ABBY Fine Reader استفاده شده است. برای بدست آوردن صحت OCR، تعداد کاراکترهایی که نرم افزار معرفی شده به درستی تشخیص داده است به تعداد کل کاراکترها تقسیم می شود.

معیار MOS

برای معیار MOS از ۲۰ نفر خواسته شد که به تصاویری که به آنها نشان داده می شود از نظر خوانایی و کیفیت از ۱ تا ۵ نمره دهند. به این صورت که ۵ بهترین نمره ای است که می توان به یک تصویر با کیفیت عالی داد و ۱ ضعیف ترین نمره ای است که می توان به یک تصویر داد. در نهایت برای هر تصویر کل ۲۰ نمره ای که به آن اختصاص داده شده است را میانگین گیری کرده و نمره نهایی برای تصویر مورد نظر گزارش می شود.

معیار PSNR

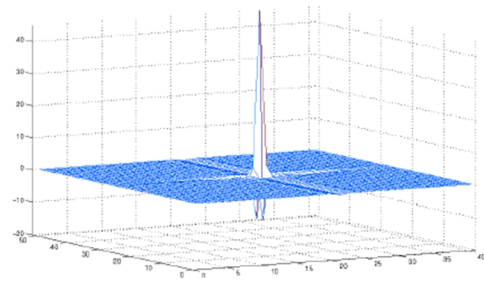
این معیار میزان نزدیک بودن تصویر باز سازی شده به تصویر اصلی را از لحاظ عددی نشان می دهد. برای محاسبه PSNR از رابطه (۱۵) استفاده می شود:

$$psnr = 10 \log_{10} \left(\frac{255^2 mn}{\| \hat{X} - I \|^2} \right) \quad (15)$$

در رابطه (۱۵)، I و $\hat{X}$ به ترتیب تصویر اصلی و تصویر باز سازی شده و mn تعداد پیکسل های تصویر اصلی هستند.

$$y^T(i,j) = 3y^2(i,j) - \frac{1}{2}y(i+1,j+1)y(i-1,j-1) - \frac{1}{2}y(i+1,j-1)y(i-1,j+1) \quad (15)$$

ساختار فیلتر تیگر در شکل ۳-۳ نشان داده شده است.



شکل ۵ فیلتر تیگر [۳۱]

فیلتر تیگر یک فیلتر غیر خطی است که ما را قادر می سازد تا لبه های تصویر متن را تقویت و نویز را حذف کنیم. پس از بهبود لایه حاوی اطلاعات لبه متن، به منظور بزرگ نمایی این لایه از درون بای دو مکعبی استفاده می کنیم. افزایش نمونه های لایه لبه متن به صورت رابطه (۱۶) انجام می شود:

$$a^h = \text{Bicubic}(a' + \beta \text{Teager}(a')) \quad (16)$$

در رابطه (۱۶) β ضریب وزن دهی برای کنترل تیزی لبه ها است که معمولاً برابر ۱ قرار داده می شود [۲۵]. $\text{Teager}()$ فیلتر تیگر و $\text{Bicubic}()$ درون بای دو مکعبی را نشان می دهد. پس از آنکه لایه های وضوح بالای a^h ، F^h و B^h به دست آمدند، تصویر وضوح بالایی با ترکیب این لایه ها و به صورت رابطه (۱۷) بدست می آید [۲۵]:

$$I^h = a^h F^h + (1 - a^h) B^h \quad (17)$$

۴- نتایج

جهت تولید داده های آموزشی، پایگاه داده ای شامل ۱۰۱ تصویر از تصاویر متنی واقعی در درجه ی تفکیک مکانی ۶۰۰ dpi و ۳۰۰ dpi ایجاد شد. در شکل های ۶ و ۷ به ترتیب تصاویری از پایگاه داده که با درجه ی تفکیک ۳۰۰ dpi و ۶۰۰ dpi روبش شده اند، نشان داده شده اند. برای نشان دادن اختلاف بین این تصاویر، قسمتی از این تصاویر در شکل های ۸ و ۹ بزرگ نمایی شده اند.

به منظور اجرای الگوریتم فراتفکیک پذیری از کد نوشته شده در نرم افزار متلب استفاده شده است که در سیستمی با مشخصات پردازنده Intel Core i5-3210M @ 2.5GHz و 4G RAM اجرا شده است و ضریب کاهش ابعاد برای نمودارهای گزارش شده در این تحقیق ۵ می باشد. زمان اجرای روش پیشنهادی به طور میانگین ۵.۲۳ ثانیه بوده است. نتایج ارائه شده در این مقاله بهترین پاسخ به دست آمده از اجرای روش پیشنهادی روی تصاویر پایگاه داده می باشد.

170 CHAPTER 7 Clustering

- k -means cannot guarantee convergence to the global minimum of $J(m, U)$ (which, hopefully, corresponds to the best possible clustering). In other words, it returns clusterings corresponding to local minima of $J(m, U)$. Consequently, different initializations of the algorithm may lead to different final clusterings. Care must be taken in the initialization of u_j (see the practical hints that follow). If the initial values for say, u_1 , of u_j lie away from the region where the data vectors lie, they may never be updated. As a consequence, k -means algorithm will proceed as if there were only $m - m_1$ n y 's.
 - Accurate estimation of the number of clusters (representatives) is crucial for the algorithm, since a poor estimate will prevent it from unraveling the clustering structure of X . More specifically, if a larger number of representatives is used, it is likely that at least one "physical" cluster will be split into two or more. On the other hand, if a smaller number of representatives is used, two or more physical clusters are likely to be represented by a single representative, which in general will lie in a sparse region (with respect to the number of data points) between those clusters.
 - The algorithm is sensitive to the presence of outliers (that is, points that lie away from almost all data vectors in X) and "noisy" data vectors. Such points are the result of a noisy process unrelated to the clustering structure of X . Since both outliers and noisy points are necessarily assigned to a cluster, they influence the respective mean representative.
 - k -means is suitable for real-valued data and, in principle, should not be used with discrete-valued data.
- Practical Hints**
- Assuming that m is fixed, and to increase our chances of obtaining a reliable clustering, we may run k -means several times, each time using different initial values for the representatives, and select the best possible clustering (according to J). These simple methods for choosing initial values for u_j 's are (a) random initialization, (b) random selection of m data vectors from X as the initial estimates of u_j 's, and (c) utilization of the clustering output of a simpler (e.g., sequential) algorithm as input.
 - Two simple ways to estimate m are:
 - Use the methodology described for the ISAS algorithm.
 - For each value of m , in a suitably chosen range $[m_{\min}, m_{\max}]$, run the k -means algorithm n_{times} times (each time using different initial values) and determine the clustering (among the n_{times} produced) that minimizes the cost function J . Let J_m be the value of J for the latter clustering. Plot J_m versus m and search for a significant local change: (i) appears as a significant "knee". If such a knee occurs, its position indicates the desired number of clusters. Otherwise, it is an indication that there is no clustering structure (that contains compact clusters) in the data set.
 - Two simple ways to deal with outliers are (a) to determine the points that lie at "large" distances from most of the data vectors in X and discard them, or (b) to run k -means and identify the clusters with very few elements. An alternative is to use algorithms that are less sensitive to outliers (this is the case with the PAM algorithm, discussed later).

Example 7.5.1. Generate and plot a data set X_n that consists of $N = 400$ 2-dimensional points. These points form four equally sized groups. Each group contains vectors that stem from Gaussian distributions

شکل ۶- تصویر ۳۰۰ dpi از پایگاه داده ایجاد شده

در شکل‌های ۱۱ تا ۱۳ نتایج حاصل از ترکیب روش پیشنهادی به ترتیب با روش‌های فشرده‌سازی JPEG، JPEG2000 و SPIHT در نرخ بیت‌های ۰/۱، ۰/۱۵ و ۰/۱ نشان داده شده است. با دقت در این شکل‌ها که برای دیدن بهتر اثر روش‌های فشرده‌سازی و کارایی روش پیشنهادی بزرگ‌نمایی شده است، مشخص است که روش پیشنهادی در ترکیب با هر یک از روش‌های فشرده‌سازی بکار گرفته شده، توانسته است اثرات مخرب ناشی از این روش‌ها و همچنین اثرات مخرب ناشی از کاهش ابعاد را به خوبی بهبود دهد. در شکل ۱۱ روش پیشنهادی توانسته است بخوبی اثرات بلوکی ناشی از روش فشرده‌سازی JPEG را بهبود دهد. شکل ۱۲ اثرات ناشی از اعمال روش فشرده‌سازی JPEG2000 روی تصویر ورودی را در نرخ بیت ۰/۱۵ نشان می‌دهد. کارایی روش پیشنهادی در ترکیب با روش فشرده‌سازی JPEG2000 در این شکل دیده می‌شود. در شکل ۱۳ اثرات ناشی از اعمال روش فشرده‌سازی مبتنی بر تبدیل موجک و سپس کدگذاری SPIHT روی تصویر ورودی در نرخ بیت ۰/۱ قابل مشاهده است. روش پیشنهادی در ترکیب با روش فشرده‌سازی SPIHT توانسته است اثرات ناشی مخرب ناشی از این روش فشرده‌سازی را بهبود دهد.

sentences

الف

sentences

ب

sentences

ج

شکل ۱۱ بزرگ‌نمایی بخشی از تصویر در ترکیب روش پیشنهادی با JPEG (الف) تصویر اصلی (ب) تصویر فشرده‌شده با JPEG. در نرخ بیت ۰/۱ (ج) روش پیشنهادی در ترکیب با روش فشرده‌سازی JPEG در نرخ بیت ۰/۱

sentences

الف

sentences

ب

sentences

ج

شکل ۱۲ بزرگ‌نمایی بخشی از تصویر در ترکیب روش پیشنهادی با JPEG2000 (الف) تصویر اصلی (ب) تصویر فشرده‌شده با JPEG2000 در نرخ بیت ۰/۱۵ (ج) ترکیب روش پیشنهادی با روش فشرده‌سازی JPEG2000 در نرخ بیت ۰/۱۵

170 CHAPTER 7 Clustering

- k -means cannot guarantee convergence to the global minimum of $J(\theta, U)$ (which, hopefully, corresponds to the best possible clustering). In other words, it returns clusterings corresponding to local minima of $J(\theta, U)$. Consequently, different initializations of the algorithm may lead to different final clusterings. Care must be taken in the initialization of θ_j (see the practical hints that follow). If the initial values for, say, m_i , of θ_j lie away from the region where the data vectors lie, they may never be updated. As a consequence, k -means algorithm will proceed as if there were only $m_i - m_j \theta_j$'s.
- Accurate estimation of the number of clusters (representatives) is crucial for the algorithm, since a poor estimate will prevent it from unravelling the clustering structure of X . More specifically, if a larger number of representatives is used, it is likely that at least one "physical" cluster will be split into two or more. On the other hand, if a smaller number of representatives is used, two or more physical clusters are likely to be represented by a single representative, which in general will lie in a sparse region (with respect to the number of data points) between those clusters.
- The algorithm is sensitive to the presence of outliers (that is, points that lie away from almost all data vectors in X) and "noisy" data vectors. Such points are the result of a noisy process unrelated to the clustering structure of X . Since both outliers and noisy points are necessarily assigned to a cluster, they influence the respective mean representatives.
- k -means is suitable for real-valued data and, in principle, should not be used with discrete-valued data.

Practical Hints

- Assuming that m is fixed, and to increase our chances of obtaining a reliable clustering, we may run k -means several times, each time using different initial values for the representatives, and select the best possible clustering (according to J). Three simple methods for choosing initial values for θ_j 's are (a) random initialization, (b) random selection of m data vectors from X as the initial estimates of θ_j 's, and (c) utilization of the clustering output of a simpler (e.g., sequential) algorithm as input.
- Two simple ways to estimate m are:
 - Use the methodology described for the BSAS algorithm.
 - For each value of m , in a suitably chosen range $[m_{\min}, m_{\max}]$, run the k -means algorithm n_{runs} times (each time using different initial values) and determine the clustering (among the n_{runs} produced) that minimizes the cost function J . Let J_m be the value of J for the latter clustering. Plot J_m versus m and search for a significant local change (it appears as a significant "knee"). If such a knee occurs, its position indicates the desired number of clusters. Otherwise, it is an indication that there is no clustering structure (that contains compact clusters) in the data set.
- Two simple ways to deal with outliers are (a) to determine the points that lie at "large" distances from most of the data vectors in X and discard them, or (b) to run k -means and identify the clusters with very few elements. An alternative is to use algorithms that are less sensitive to outliers (this is the case with the PAM algorithm, discussed later).

Example 7.5.1. Generate and plot a data set, X_N , that consists of $N = 400$ 2-dimensional points. These points form four equally sized groups. Each group contains vectors that stem from Gaussian distributions

شکل ۷- تصویر ۶۰۰dpi از پایگاه داده ایجاد شده

This procedure is

شکل ۸ تصویر ۳۰۰dpi بزرگ‌نمایی شده از پایگاه داده

This procedure is

شکل ۹ تصویر ۶۰۰dpi بزرگ‌نمایی شده از پایگاه داده

در شکل ۱۰ به ترتیب اثر روش‌های فشرده‌سازی JPEG، SPIHT و زمانی که این روش‌ها بطور مستقیم روی تصاویر متنی با درجه تفکیک ۶۰۰dpi اعمال شوند، نشان داده شده است. همان‌طور که از این شکل مشخص است، در نرخ بیت‌های پایین روش فشرده‌سازی JPEG روی تصویر متنی اثرات بلوکی و روش‌های فشرده‌سازی JPEG2000 و SPIHT اثر ناشی ایجاد می‌کنند.

sentences

الف

sentences

ب

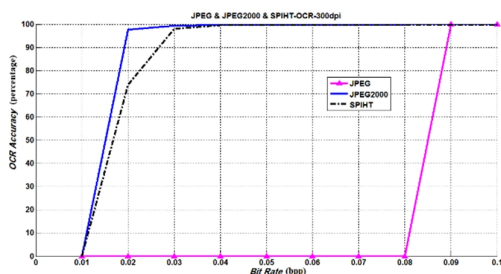
sentences

ج

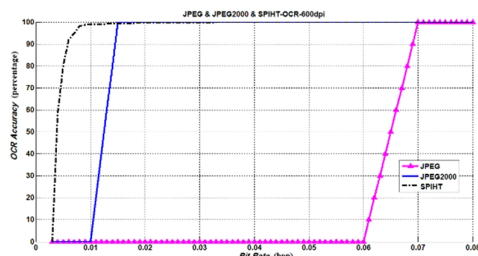
sentences

د

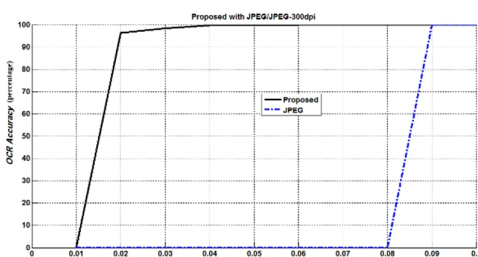
شکل ۱۰ اثرات ناشی از روش‌های فشرده‌سازی مختلف روی تصاویر متنی (الف) تصویر اصلی (ب) اثر بلوکی حاصل از JPEG با نرخ بیت ۰/۰۹ (ج) اثر ناشی حاصل از JPEG2000 با نرخ بیت ۰/۰۸ (د) اثر ناشی فشرده‌سازی به کمک SPIHT با نرخ بیت ۰/۰۸



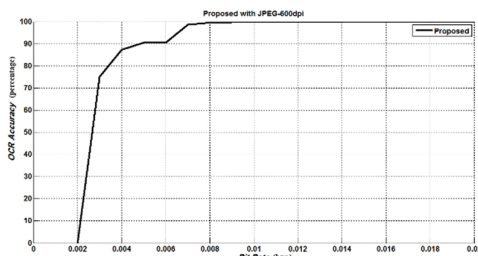
شکل ۱۴ مقایسه روش‌های فشرده‌سازی از نظر معیار OCR برای تصویر با درجه‌ی تفکیک مکانی ۳۰۰dpi



شکل ۱۵ مقایسه روش‌های فشرده‌سازی از نظر معیار OCR برای تصویر با درجه‌ی تفکیک مکانی ۶۰۰dpi



شکل ۱۶ نتایج OCR روش پیشنهادی در ترکیب با فشرده‌سازی JPEG، روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



شکل ۱۷ نتایج OCR روش پیشنهادی در ترکیب با فشرده‌سازی JPEG، روی تصویر با درجه تفکیک مکانی ۶۰۰dpi

فشرده سازی JPEG. توانسته‌ایم نرخ بیت را از ۰/۰۷ به ۰/۰۴ کاهش دهیم. در این شکل به دلیل نشان دادن دقت روش پیشنهادی تنها نمودار مربوط به روش پیشنهادی رسم شده است. در ادامه در شکل‌های ۱۸ و ۱۹ نتایج حاصل از روش فشرده‌سازی JPEG2000 به تنهایی و در ترکیب با روش پیشنهادی نشان داده شده است. همان‌طور که از این شکل‌ها مشخص است در ترکیب روش پیشنهادی با JPEG2000 می‌توان به نرخ بیت کمتری نسبت به اعمال روش JPEG2000 به تنهایی رسید.

sentences
الف
sentences
ب
sentences
ج

شکل ۱۳ بزرگ‌نمایی بخشی از تصویر در ترکیب روش پیشنهادی با SPIHT (الف) تصویر اصلی (ب) تصویر فشرده‌شده با SPIHT در نرخ بیت ۰/۱ (ج) ترکیب روش پیشنهادی با روش فشرده‌سازی SPIHT در نرخ بیت ۰/۱

۴-۱- ارزیابی روش پیشنهادی براساس معیار OCR

ابتدا دو دسته تصویر با درجه‌ی تفکیک مکانی متوسط و بالا (به ترتیب ۳۰۰ dpi و ۶۰۰ dpi) به طور مستقیم به روش‌های فشرده سازی JPEG، JPEG2000 و کدگذاری SPIHT داده شده و خروجی حاصل از هر روش براساس معیار OCR ارزیابی می‌شود. نتایج به دست آمده به ترتیب در شکل‌های ۱۴ و ۱۵ نشان داده شده است.

همان‌طور که از این شکل‌ها دیده می‌شود، روش فشرده‌سازی JPEG از نظر صحت OCR جواب مطلوبی دارد. اما به دلیل محدودیت ذاتی این استاندارد رسیدن به نرخ بیت کمتر از ۰/۰۹ bpp و ۰/۰۷ bpp به ترتیب برای تصاویر با درجه تفکیک مکانی ۳۰۰ dpi و ۶۰۰ dpi امکان پذیر نبوده است. روش فشرده‌سازی JPEG2000 نیز به ترتیب تا نرخ بیت ۰/۰۲ bpp و ۰/۰۱۵ bpp برای تصاویر مذکور جواب مطلوبی داشته است. بنا به همان دلیل قبلی برای این روش فشرده سازی نیز رسیدن به نرخ بیت کمتر از ۰/۰۲ bpp و ۰/۰۱۵ bpp امکان پذیر نبوده است. اما روش SPIHT برخلاف دو روش قبل محدودیتی روی نرخ بیت آن وجود ندارد و تا هر نرخ بیت دلخواهی، می‌توان تصویر ورودی را فشرده کرد. در این روش برای نرخ بیت‌های کمتر از ۰/۰۲ bpp و ۰/۰۰۶ bpp به ترتیب برای تصاویر ۳۰۰ dpi و ۶۰۰ dpi از نظر معیار OCR جواب مطلوبی به دست نمی‌آید.

در رسم نمودارها برای نرخ بیت‌هایی که رسیدن به آنها امکان پذیر نبوده است داده‌های مصنوعی صفر ایجاد شده است. به طور مثال برای نرخ بیت‌های کمتر از ۰/۰۹ bpp در روش فشرده‌سازی JPEG یا نرخ بیت‌های کمتر از ۰/۰۲ bpp در روش فشرده‌سازی JPEG2000 در شکل ۱۴ داده‌های مصنوعی صفر قرار داده شده است.

در شکل‌های ۱۶ و ۱۷ نتایج حاصل از OCR روش فشرده‌سازی JPEG به تنهایی و در ترکیب با روش پیشنهادی روی تصاویر متنی به ترتیب با درجه‌ی تفکیک مکانی ۳۰۰ dpi و ۶۰۰ dpi، نشان داده شده است. همان‌طور که از شکل ۱۶ مشخص است روش پیشنهادی توانسته فشرده‌سازی را از ۰/۰۹ تا نرخ بیت ۰/۰۲ پایین بیاورد. از مقایسه شکل ۱۷ با شکل ۱۵ می‌توان دید با ترکیب روش پیشنهادی با روش

جدول ۱: مقادیر میانگین OCR ترکیب روش پیشنهادی با روش های فشرده-

سازي روی تصاویر با درجه تفکیک مکانی ۳۰۰dpi									
نرخ بیت	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰
JPEG	۹۹.۰	۹۹.۰	۹۹.۰	۹۸.۰	۹۸.۰	۹۷.۰	۹۷.۰	۹۶.۰	۰
JPEG 2000	۹۹.۰	۹۹.۰	۹۹.۰	۹۹.۰	۹۹.۰	۹۹.۰	۹۸.۰	۹۶.۰	۶۲
SPIHT	۹۹.۰	۹۹.۰	۹۹.۰	۹۹.۰	۹۸.۰	۹۸.۰	۹۸.۰	۹۷.۰	۷۱

جدول ۲: مقادیر انحراف معیار OCR ترکیب روش پیشنهادی با روش های

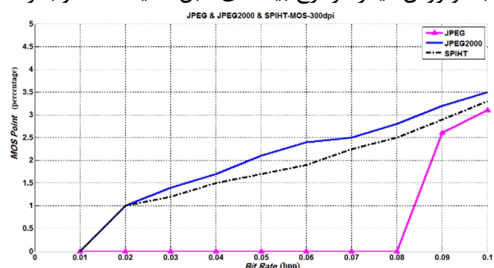
فشرده سازی روی تصاویر با درجه تفکیک مکانی ۳۰۰dpi									
نرخ بیت	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰	۰.۰
JPEG	۰.۴	۰.۵	۰.۷	۰.۷	۰.۸	۰.۹	۱.۰	۱.۱	۰
JPEG 2000	۰.۱	۰.۲	۰.۲	۰.۳	۰.۳	۰.۳	۰.۵	۰.۶	۰.۸
SPIHT	۰.۰	۰.۳	۰.۴	۰.۵	۰.۶	۰.۶	۰.۷	۰.۷	۰.۸

شده است. در این حالت هم استفاده از روش پیشنهادی موجب بهبود عملکرد فشرده سازی و بازشناسی شده است.

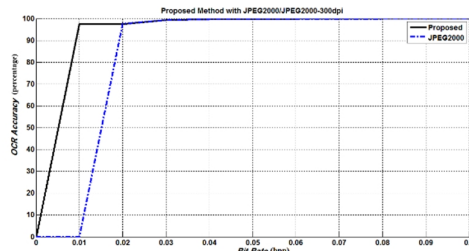
در جداول ۱ و ۲ به ترتیب مقادیر مربوط به میانگین و انحراف معیار نتایج حاصل از OCR روش پیشنهادی در ترکیب با سه روش فشرده سازی JPEG2000 و SPIHT روی تصاویر با درجه تفکیک مکانی ۳۰۰ dpi ارائه شده است. جداول ۳ و ۴ به ترتیب مقادیر مربوط به میانگین و انحراف معیار نتایج حاصل از OCR روش پیشنهادی در ترکیب با سه روش فشرده سازی JPEG2000 و SPIHT روی مجموعه تصاویر با درجه تفکیک مکانی ۶۰۰ dpi را ارائه می دهند.

۲-۴- ارزیابی روش پیشنهادی براساس معیار MOS

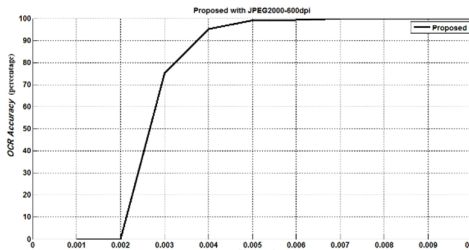
در شکل های ۲۲ و ۲۳ نتایج محاسبه معیار MOS حاصل برای حالت اعمال روش های JPEG2000 و SPIHT به تنهایی روی دو تصویر به ترتیب با درجه تفکیک مکانی ۳۰۰ dpi و ۶۰۰ dpi نشان داده شده است. نتایج نشان می دهند که افراد کیفیت روش JPEG2000 را در مقایسه با دو روش دیگر در نرخ بیت های قابل مقایسه، مطلوب تر



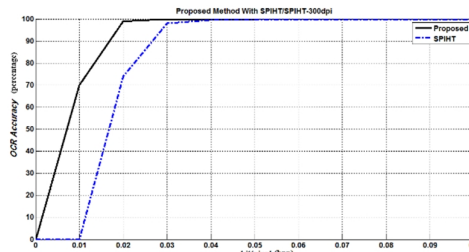
شکل ۲۲ مقایسه روش های فشرده سازی از نظر معیار MOS برای تصویر با درجه تفکیک مکانی ۳۰۰dpi



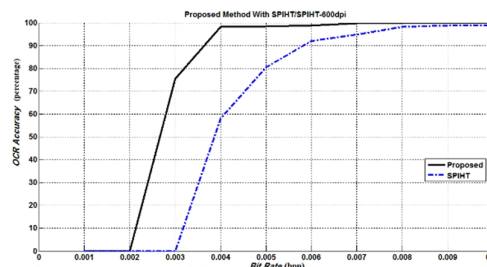
شکل ۱۸ نتایج OCR روش پیشنهادی در ترکیب با فشرده سازی JPEG2000 روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



شکل ۱۹ نتایج OCR روش پیشنهادی در ترکیب با فشرده سازی JPEG2000 روی تصویر با درجه تفکیک مکانی ۶۰۰dpi



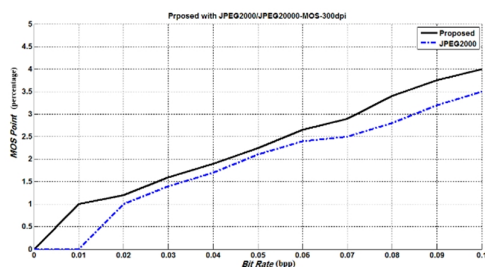
شکل ۲۰ نتایج OCR روش پیشنهادی در ترکیب با فشرده سازی SPIHT روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



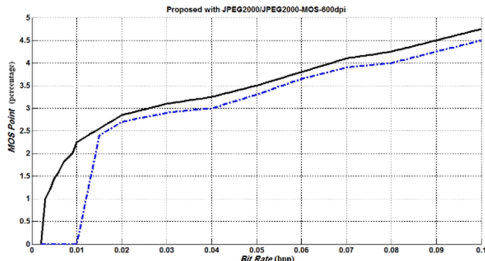
شکل ۲۱ نتایج OCR روش پیشنهادی در ترکیب با فشرده سازی SPIHT روی تصویر با درجه تفکیک مکانی ۶۰۰dpi

از مقایسه شکل های ۱۹ و ۱۵ می توان دید روش پیشنهادی در ترکیب با JPEG2000 در مقایسه با اعمال روش فشرده سازی JPEG2000 به تنهایی روی تصویری با درجه تفکیک مکانی ۶۰۰ dpi، توانسته نرخ بیت را از ۰/۰۱۵ به ۰/۰۰۴ برساند.

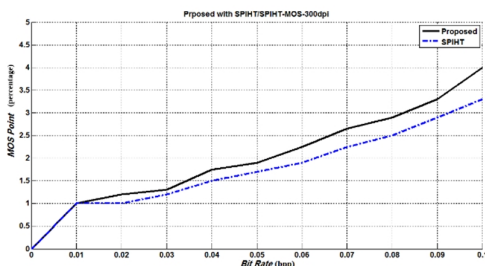
در شکل های ۲۰ و ۲۱ نتایج حاصل از ترکیب روش پیشنهادی با روش SPIHT در مقایسه با اعمال روش SPIHT به تنهایی نشان داده



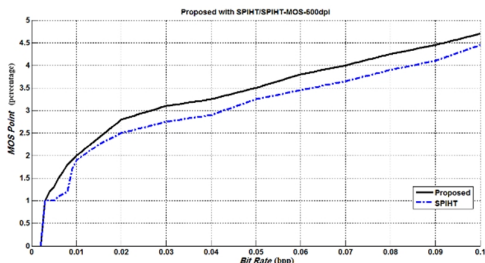
شکل ۲۶ نتایج MOS روش پیشنهادی در ترکیب با روش JPEG2000 روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



شکل ۲۷ نتایج MOS روش پیشنهادی در ترکیب با روش JPEG2000 روی تصویر با درجه تفکیک مکانی ۶۰۰dpi



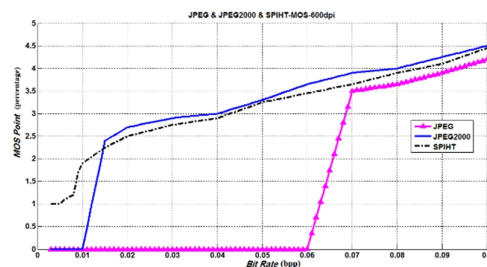
شکل ۲۸ نتایج MOS روش پیشنهادی در ترکیب با روش SPIHT روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



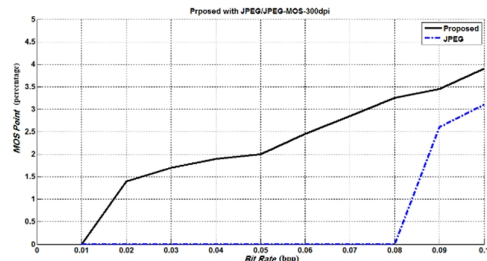
شکل ۲۹ نتایج MOS روش پیشنهادی در ترکیب با روش SPIHT روی تصویر با درجه تفکیک مکانی ۶۰۰dpi

نمودارهای مربوط به ارزیابی معیار MOS روش پیشنهادی در ترکیب با روش فشرده سازی SPIHT برای تصاویر با درجه تفکیک مکانی ۳۰۰ dpi و ۶۰۰ dpi به ترتیب در شکل های ۲۸ و ۲۹ نشان داده شده است. در این شکل ها خط ممتد نمودار مربوط به روش پیشنهادی در ترکیب با روش SPIHT و خط-نقطه مربوط به روش فشرده سازی SPIHT به تنهایی است.

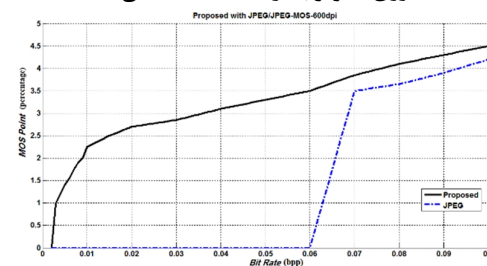
همان طور که از این شکل ها مشخص است روش پیشنهادی در ترکیب با روش SPIHT برای هر دو تصویر از نظر MOS جواب بهتری را داشته است.



شکل ۲۳ مقایسه روش های فشرده سازی از نظر معیار MOS برای تصویر با درجه تفکیک مکانی ۶۰۰dpi



شکل ۲۴ نتایج MOS روش پیشنهادی در ترکیب با روش فشرده سازی JPEG، روی تصویر با درجه تفکیک مکانی ۳۰۰dpi

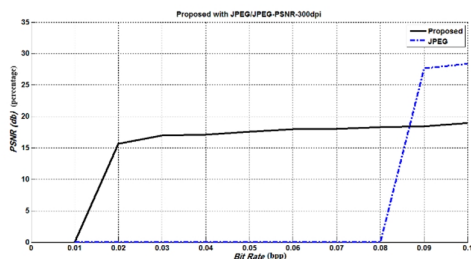


شکل ۲۵ نتایج MOS روش پیشنهادی در ترکیب با روش فشرده سازی JPEG، روی تصویر با درجه تفکیک مکانی ۶۰۰dpi

می دانسته اند. در شکل های ۲۴ و ۲۵ نتایج حاصل از ترکیب روش پیشنهادی با فشرده سازی JPEG در مقایسه با اعمال روش JPEG به تنهایی نشان داده شده است. در هر دو نمودار خط ممتد مربوط به روش پیشنهادی در ترکیب با روش فشرده سازی JPEG و خط-نقطه مربوط به روش فشرده سازی JPEG به تنهایی است.

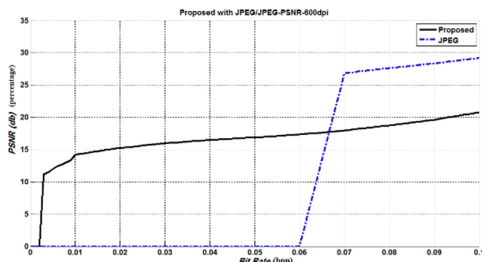
همان طور که از این شکل ها دیده می شود، ترکیب روش پیشنهادی با روش فشرده سازی JPEG از نظر خوانایی انسانی در جامعه ۲۰ نفره ذکر شده، نتایج مطلوب تری نسبت به روش فشرده سازی JPEG به تنهایی داشته است.

در شکل های ۲۶ و ۲۷ نتایج حاصل از ترکیب روش پیشنهادی با روش فشرده سازی JPEG2000 در مقایسه با اعمال روش فشرده سازی JPEG2000 به تنهایی نشان داده شده است. همان طور که از این شکل ها مشخص است روش پیشنهادی در ترکیب با روش فشرده سازی JPEG2000 توانسته است خوانایی تصویری فشرده شده ورودی را بهبود دهد. در شکل های ۲۶ و ۲۷، خط ممتد مربوط به روش پیشنهادی در ترکیب با روش فشرده سازی JPEG2000 و خط-نقطه مربوط به روش فشرده سازی JPEG2000 به تنهایی است.



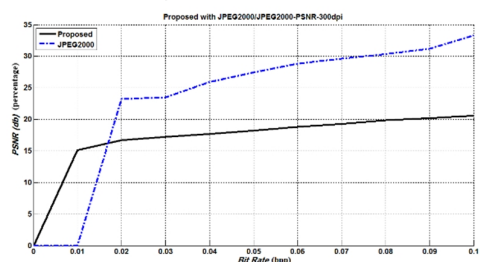
شکل ۳۲ نتایج PSNR روش پیشنهادی در ترکیب با روش JPEG، روی

تصویر با درجه تفکیک مکانی ۳۰۰dpi



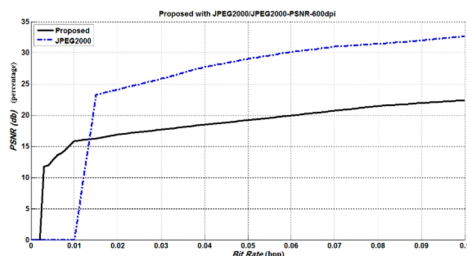
شکل ۳۳ نتایج PSNR روش پیشنهادی در ترکیب با روش JPEG، روی

تصویر با درجه تفکیک مکانی ۶۰۰dpi



شکل ۳۴ نتایج PSNR روش پیشنهادی در ترکیب با روش JPEG2000

روی تصویر با درجه تفکیک مکانی ۳۰۰dpi



شکل ۳۵ نتایج PSNR روش پیشنهادی در ترکیب با روش JPEG2000

روی تصویر با درجه تفکیک مکانی ۶۰۰dpi

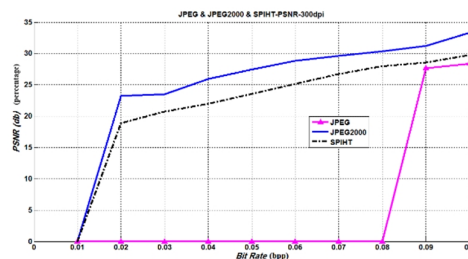
شکل های ۳۴ و ۳۵ نتایج حاصل از ترکیب روش پیشنهادی با روش فشرده سازی JPEG2000 را در مقایسه با اعمال روش فشرده سازی JPEG2000 به تنهایی بر مبنای معیار PSNR نشان می دهند. از این شکل ها نیز می توان دید که به دلیل گفته شده در بالا، روش پیشنهادی در ترکیب با روش فشرده سازی JPEG2000 از نظر معیار PSNR نتایج ضعیف تری نسبت به روش فشرده سازی JPEG2000 به تنهایی دارد. نتایج حاصل از ترکیب روش پیشنهادی با روش SPIHT بر مبنای معیار PSNR در شکل های ۳۴ و ۳۵ نشان داده شده است.

۳-۴- ارزیابی روش پیشنهادی براساس معیار PSNR

این معیار در واقع میزان نزدیک بودن تصویر بازسازی شده و تصویر اصلی را نشان می دهد. نتایج گزارش شده برای این معیار براساس رابطه (۱۵) بدست آمده است. در ابتدا نتایج محاسبه معیار PSNR حاصل از اعمال روش های فشرده سازی JPEG2000، SPIHT و به تنهایی گزارش شده اند. سپس نتایج حاصل از ترکیب روش پیشنهادی با هر یک از روش های فشرده سازی گفته شده گزارش شده است.

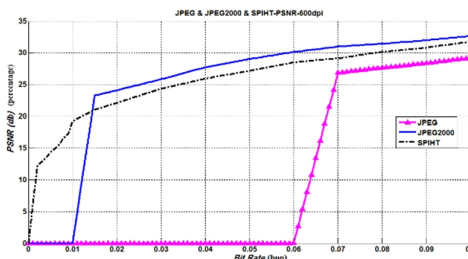
شکل های ۳۰ و ۳۱ نتایج حاصل از اعمال روش های فشرده سازی گفته شده به تنهایی را روی تصاویر براساس معیار PSNR نشان می دهند. از نتایج بدست آمده در شکل های ۳۰ و ۳۱ دیده می شود که روش JPEG2000 نسبت به دو روش دیگر از نظر معیار PSNR جواب بهتری داشته است. در این نمودارها باز هم محدودیت فشرده سازی برای روش های JPEG و JPEG2000 قابل مشاهده است.

در شکل های ۳۲ و ۳۳ نتایج حاصل از ترکیب روش پیشنهادی با روش JPEG در مقایسه با اعمال روش فشرده سازی JPEG به تنهایی بر مبنای معیار PSNR نشان داده شده است. نتایج بدست آمده برای روش پیشنهادی براساس معیار PSNR در مقایسه با اعمال روش فشرده سازی JPEG به تنهایی ضعیف تر بوده است. دلیل این نتایج این است که در روش پیشنهادی در مرحله بازسازی ضخامت حروف در تصویر متنی سبب می شود که در زمان محاسبه PSNR نقاطی که در تصویر بازسازی شده به ضخامت حروف افزوده شده اند، خطای بازسازی را نسبت به تصویر اصلی زیاد کنند.



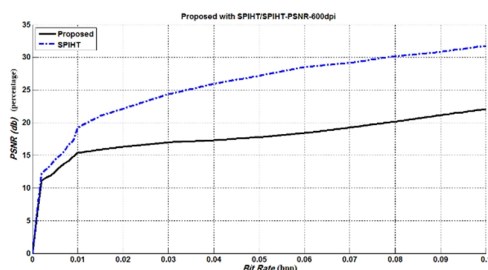
شکل ۳۰ مقایسه روش های فشرده سازی از نظر معیار PSNR برای

تصویر با درجه تفکیک مکانی ۳۰۰dpi

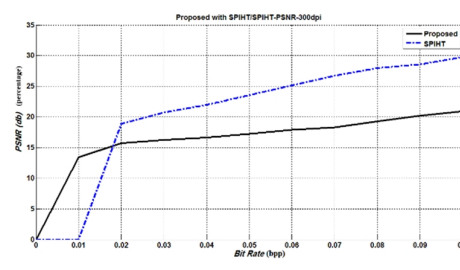


شکل ۳۱ مقایسه روش های فشرده سازی از نظر معیار PSNR برای

تصویر با درجه تفکیک مکانی ۶۰۰dpi



شکل ۳۷ نتایج PSNR روش پیشنهادی در ترکیب با روش SPIHT روی تصویر با درجه تفکیک مکانی ۶۰۰dpi



شکل ۳۶ نتایج PSNR روش پیشنهادی در ترکیب با روش SPIHT روی تصویر با درجه تفکیک مکانی ۳۰۰dpi

جدول ۳: مقادیر میانگین OCR ترکیب روش پیشنهادی با روش های فشرده سازی روی تصاویر با درجه تفکیک مکانی ۶۰۰dpi

نرخ بیت	۰.۰۰۴	۰.۰۰۵	۰.۰۰۶	۰.۰۰۷	۰.۰۰۸	۰.۰۰۹	۰.۰۱	۰.۰۱۵	۰.۰۲	۰.۰۳	۰.۰۴	۰.۰۵	۰.۰۶	۰.۰۷	۰.۰۸
JPEG	۸۶.۷	۸۷.۴۱	۸۷.۸۴	۹۷.۶۵	۹۷.۸۴	۹۸.۱۳	۹۸.۴۸	۹۸.۸۱	۹۹.۰۵	۹۹.۲۱	۹۹.۴۳	۹۹.۶۱	۹۹.۷	۹۹.۸۴	۹۹.۹۴
JPEG2000	۹۳.۷۶	۹۷.۶۷	۹۷.۹۸	۹۸.۵۶	۹۸.۷۸	۹۹.۰۴	۹۹.۲۵	۹۹.۵۵	۹۹.۶۴	۹۹.۷۱	۹۹.۷۸	۹۹.۸۴	۹۹.۹	۹۹.۹۸	۹۹.۹۹
SPIHT	۹۶.۹۵	۹۷.۷۶	۹۷.۹۹	۹۸.۶۱	۹۸.۸۸	۹۸.۹۶	۹۹.۱۴	۹۹.۳۴	۹۹.۴۸	۹۹.۵۸	۹۹.۶۵	۹۹.۸۴	۹۹.۸۸	۹۹.۹۷	۹۹.۹۹

جدول ۴: مقادیر انحراف معیار OCR ترکیب روش پیشنهادی با روش های فشرده سازی روی تصاویر با درجه تفکیک مکانی ۶۰۰dpi

نرخ بیت	۰.۰۰۴	۰.۰۰۵	۰.۰۰۶	۰.۰۰۷	۰.۰۰۸	۰.۰۰۹	۰.۰۱	۰.۰۱۵	۰.۰۲	۰.۰۳	۰.۰۴	۰.۰۵	۰.۰۶	۰.۰۷	۰.۰۸
JPEG	۱.۲۴۳	۱.۱۲۷	۰.۹۷۸	۰.۷۵۲	۰.۷۰۱	۰.۶۵۳	۰.۵۷۲	۰.۵۲۰	۰.۴۶۹	۰.۴۱۵	۰.۳۶۶	۰.۳۰۳	۰.۲۳۸	۰.۱۲۴	۰.۰۴
JPEG2000	۱.۱۱۱	۰.۷۶۱	۰.۷۰۳	۰.۶۵۱	۰.۶۱۳	۰.۵۵۸	۰.۵۰۳	۰.۴۴۱	۰.۳۵۲	۰.۲۸۱	۰.۲۱۳	۰.۱۱۳	۹۹.۹	۰.۰۱۱	۰.۰۰۲
SPIHT	۱.۰۰۳	۰.۸۷۱	۰.۸۰۱	۰.۷۲۱	۰.۶۶۲	۰.۶۰۱	۰.۵۴۱	۰.۵۰۰	۰.۴۵۲	۰.۳۴۱	۰.۲۶۳	۰.۱۳۴	۰.۱۰۱	۰.۰۱	۰.۰۰۵

نتایج بدست آمده با روش پیشنهادی براساس معیارهای OCR، PSNR و MOS در مقایسه با دیگر روش های فشرده سازی شامل JPEG، SPIHT و JPEG2000 مورد ارزیابی قرار گرفت. روش پیشنهادی براساس معیارهای OCR و MOS و افزایش میزان فشرده سازی تصاویر متنی نسبت به روش های فشرده سازی استفاده شده نتایج بهتری را داشته است اما از نظر معیار PSNR روش پیشنهادی نسبت به روش های استفاده شده ضعیف تر بوده است که دلیل این امر ضخیم شدن حروف در مرحله ی بازسازی است.

مراجع

- [1] D. W. Naaman, "Image Compression Technique Based on Fractal Image Compression Using Neural Network – A Review", *AJRCoS*, vol. 10, no. 4, pp. 47-57, Jul. 2021.
- [2] H. H. Huda, "Lossless Text Image Compression using Two Dimensional Run Length Encoding", *Journal Online Anformatika (JOIN)*, vol. 4, n0. 2, pp.75-78, 2019.
- [3] H. Grailu, M. Lotfizad and H. S. Yazdi, "Farsi and arabic document images lossy compression based on the mixed raster content model," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol. 12, no. 4, pp. 227-248, 2009.
- [4] N. C. Francisco, N. M. M. Rodrigues, E. A. B. Silva, M. B. Carvalho, S. M. M. Faria and V. M. M. Silva, "Scanned compound document encoding using multiscale recurrent patterns," *IEEE Transactions on Image Processing*, vol. 19, no. 10, pp. 2712-2724, 2010.
- [5] L. H. Sharpe and B. Manns, "JPEG 2000 options for document image compression," *Document Recognition and Retrieval IX*, Paul B. Kantor, Tapas Kanungo, Jiangying Zhou, Editors, Proceedings of SPIE, vol. 4670, pp. 167-173, 2002.

از شکل های ۳۶ و ۳۷ نیز دیده می شود که روش فشرده سازی SPIHT به تنهایی از نظر معیار PSNR نسبت به روش پیشنهادی در ترکیب با روش SPIHT جواب بهتری داشته است.

۵- نتیجه گیری و پیشنهادات

فشرده سازی می تواند موجب اثرات مخربی روی تصاویر متنی شود. بنابراین بسته به کاربرد تا جایی می توان تصاویر متنی را فشرده کرد که خوانایی انسانی/ماشینی این تصاویر کم نشود. علاوه بر مشکلات ذکر شده، برای روش های فشرده سازی مانند JPEG و JPEG2000، مشکل دیگر، محدود بودن میزان فشرده سازی یا نرخ بیت خروجی است. در این مقاله روشی جهت فشرده سازی تصاویر متنی با درجه تفکیک مکانی بالا پیشنهاد شد. ایده اصلی روش پیشنهادی در این مقاله این بوده است که با ارائه روشی مناسب بتوان میزان فشرده سازی را بیشتر و اثرات مخرب روی تصویر فشرده شده را نیز کم کرد. برای رسیدن به میزان فشرده سازی بیشتر از ایده کاهش ابعاد در تصاویر متنی استفاده شد. البته باید به این موضوع توجه شود که کاهش ابعاد روی کیفیت تصویر بی تأثیر نیست و ممکن است باعث تنزل در کیفیت تصویر شود. بنابراین باید روشی انتخاب شود که بتواند هم ابعاد تصویر را افزایش دهد و هم اثرات مخرب تأثیر گذار بر تصویر را اصلاح کند. به همین دلیل در این مقاله از ایده فرا تفکیک پذیری در ترکیب با روش های فشرده سازی استفاده شد.

- [18] J. Ma, Sh. Guo, L. Zhang, "Text Prior Guided Scene Text Image Super-resolution", arXiv:2106.15368, 2021.
- [19] J. Cai, et al., "Toward Real-World Single Image Super-Resolution: A New Benchmark and A New Model", arXiv:1904.00523, 2019.
- [20] M. C. Chiang and T. E. Boulte, "Efficient super-resolution via image warping", *Image and Vision Computing*, vol. 18, no. 10, pp. 761-771, 2000.
- [21] S. Farsiu, M. D. Robinson, M. Elad and P. Milanfar, "Fast and robust multiframe super resolution", *IEEE Transactions on Image processing*, vol. 13, pp. 1327-1344, 2004.
- [22] R. Fattal, "Image upsampling via imposed edges statistics", *ACM Transaction on Graphics*, vol. 26, no. 3, pp. 95:1-95:8, 2007.
- [23] Q. Shan, et al., "Fast image/video upsampling", *ACM Transaction on Graphics*, vol. 25, no. 5, pp. 153:1-153:7, 2008.
- [24] A. Levin, et al., "A closed-form solution to natural image matting", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228-242, 2008.
- [25] Y. zheng, X. Kang Shato Li and Y. He Jun Sun, "Real-time document image super-Resolution by fast matting" 11th IAPR International Workshop on Document Analysis System, 2014.
- [26] J. Wang and F. Cohen, "Optimized color sampling for robust matting", *International Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2007.
- [27] K. He, J. Sun, X. Tang, "Guided image filtering", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013.
- [28] Khangar S. V. and Malik L. G., "Handwritten Text Image Compression for Indic Script Document, 2012, International Journal of Computer Applications, Vol. 47, No. 5, pp. 11-16.
- [29] de Queiroz R., Buckley R. and Xu M., "Mixed Raster Content (MRC) Model for Compound Image Compression, Proc. IS&T/SPIE Symposium on Electronic Imaging Science & Technology Visual Communications and Image Processing, San Jose, CA, Vol. 3653, pp. 1106-1177, 1999.
- [30] J. Yang, J. Wright, T. Huang and Y. Ma, "Image superresolution via sparse representation of raw image patches", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [31] C. Mancas-Thillou and M. Majid, "Super-resolution text using the teager filter." *First International Workshop on Camera-Based Document Analysis and Recognition*. pp. 10-16, 2005.
- [6] S. Dhawan, "A review of image compression and comparison of its algorithms," *International Journal of Electronics and Communication Technology*, vol. 2, no. 1, pp. 22-26, 2011.
- [7] W. A. Pearlman and A. Said, "Set partitioning coding: part I of set partition coding and image wavelet coding systems", *Foundations and Trends in Signal Processing*, vol. 2, no. 2, pp. 95-180, 2008.
- [8] W. A. Pearlman and A. Said, "Image wavelet coding systems: part II of set partition coding and image wavelet coding systems", *Foundations and Trends in Signal Processing*, vol. 2, no. 3, pp. 181-246, 2008.
- [9] W. A. Pearlman and A. Said, *Digital Signal Compression: Principles and Practice*, Cambridge University Press, New York, 2011.
- [10] A. Said and W. A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243-250, 1996.
- [11] H. Grailu, "Textual Image Compression for Maintaining or Improving the Recognition Performance", *Circuits Syst Signal Process*, no. 36, pp. 658-674, 2017.
- [12] P. Mahesh, P. Rajesh and I. Suneetha, "Improved block based segmentation for JPEG compressed document images", *International Journal of Research in Engineering and Technology*, vol. 2, no. 11, pp. 669-673, 2013.
- [13] B. Oztan, A. Malik, Z. Fan and R. Eschbach, "Removal of artifacts from JPEG compressed document images", *Proceedings of SPIE-IS&T Electronic Imaging*, vol. 6493, pp. 649306-1: 649306-9, 2007.
- [14] A. Zaghetto and R. L. Queiroz, "Scanned document compression using a block-based hybrid video codec", *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2420-2428, 2010.
- [15] H. Grailu, "Textual image compression at low bit rates based on region-of-interest coding", *IJDAR*, no. 19, pp. 65-81, 2016.
- [16] D. Taubman, "High performance scalable image compression with EBCOT", *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1151-1170, 2000.
- [17] K. Lim, Y. Liu, T. H. Li, Sh. Liu, G. Li, , "Text Image Super-Resolution by Image Matting and Text Label Supervision", *Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2019.