

ارائه یک روش جدید فشرده‌سازی تصاویر چهره با استفاده از نمایش تنک سیگنال و الگوریتم یادگیری دیکشنری RLS-DLA

امیرمسعود طاهری^۱، همایون مهدوی نسب^{۲*}

۱: دانشکده مهندسی برق، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران a.m.taheri1984@gmail.com

۲: مرکز تحقیقات پردازش دیجیتال و بینایی ماشین، واحد نجف‌آباد، دانشگاه آزاد اسلامی، نجف‌آباد، ایران

mahdavinasab62@gmail.com

تاریخ دریافت: ۱۳۹۵/۸/۱۴ تاریخ پذیرش: ۱۳۹۶/۴/۲۱

چکیده

رشد فناوری و افزایش تصاعدی اطلاعات نیاز به حجم ذخیره‌سازی بیشتر داده‌های اطلاعاتی گوناگون را افزایش داده است. در این راستا فشرده‌سازی تصویر به‌عنوان ابزاری کارآمد جهت کاهش افزونگی و صرفه‌جویی در حجم ذخیره‌سازی و کاهش پهنای باند انتقالی داده‌های تصویری به کار می‌آید. هنگامی که فشرده‌سازی یک دسته یا خانواده از تصاویر، مانند پایگاه داده تصاویر چهره یک سازمان یا مؤسسه یا پایگاه داده MRI یک بیمارستان بزرگ یا پایگاه داده اثر انگشت مدنظر باشد افزونگی اطلاعات افزایش یافته و فشرده‌سازی اهمیت و الزام بیشتری پیدا می‌کند. در این میان تصاویر چهره با توجه به کاربرد وسیعی که به‌عنوان رایج‌ترین تصاویر پایگاه داده سازمان‌های و نهادهای مختلف مانند اداره‌های پلیس، نهادهای نظامی، دانشگاه‌ها و شرکت‌های بزرگ دارند مورد توجه بیشتری قرار دارند. به همین خاطر ارائه الگوریتمی که بتواند این دسته از تصاویر را با کیفیت بیشتر و نرخ بالاتری فشرده کند اهمیت بسیاری دارد. در این مقاله با استفاده از حوزه جدیدی از پردازش سیگنال به نام نمایش تنک و روش یادگیری دیکشنری RLS-DLA الگوریتم جدیدی برای فشرده‌سازی تصویر ارائه شده است که می‌تواند برای فشرده‌سازی پایگاه داده تصاویر به کار رود. در این الگوریتم تصاویر با به‌کارگیری چند دیکشنری به نحو وفقی بر اساس کیفیت بازسازی مورد نیاز آن‌ها فشرده می‌شوند. نتایج به‌دست‌آمده از الگوریتم پیشنهادی نشان‌دهنده عملکرد مؤثر و برتری معنی‌دار آن نسبت به روش‌های پیشرفته و مطرحی همچون JPEG2000 است به‌طوری‌که به افزایش کیفیتی در حدود 0.5 dB تا 1.2 dB در نرخ بیت یکسان دست می‌یابد.

واژه‌های کلیدی: نمایش تنک، کدگذاری تنک، فشرده‌سازی تصویر، یادگیری دیکشنری، RLS-DLA، JPEG2000.

۱- مقدمه

فشرده‌سازی تصویر همواره یکی از دغدغه‌های اصلی پژوهشگران حوزه پردازش سیگنال بوده و همچنان از زمینه‌های فعال تحقیقاتی است. در این راستا روش‌های زیادی ارائه شده و توسعه یافته‌اند که از جمله آن‌ها می‌توان به استانداردهای JPEG و JPEG2000 اشاره کرد که دومی هم‌اکنون یکی از بهترین روش‌های فشرده‌سازی تصویر مبتنی بر تبدیل موجک و با کارایی بالا است. هنگامی که فشرده‌سازی یک دسته یا خانواده از تصاویر مدنظر باشد افزونگی اطلاعات افزایش یافته و باعث می‌شود تا فشرده‌سازی اهمیت بیشتری پیدا کند. فشرده‌سازی تصاویر چهره به‌عنوان تصاویری پرکاربرد که به‌طور گسترده در پایگاه‌های داده بزرگ همچون شبکه‌های اجتماعی، وب سرورها و سازمان‌های مختلف مانند سازمان ثبت‌احوال مدارس، دانشگاه‌ها و شرکت‌های خصوصی مورد استفاده قرار می‌گیرند، بسیار سودمند و جذاب است. با توجه به اینکه روش‌ها و استانداردهای موجود کارایی لازم برای فشرده‌سازی خانواده‌های خاص تصاویر را ندارند،

نیاز به روش‌های جدید جهت بهره‌گیری از افزونگی درونی رده‌های تصاویر افزایش می‌یابد. خوشبختانه در سال‌های اخیر روش نوینی جهت مدل‌سازی اطلاعات به نام نمایش تُنک^۱ ابداع شده است که با توجه به ویژگی‌های آن این امر میسر شده است. حوزه نمایش تنک حوزه نسبتاً جدیدی در پردازش سیگنال است که به سرعت در حال رشد بوده و تحولات بسیاری را در آن ایجاد کرده است، اما به نظر می‌رسد که همچنان در ابتدای راه خود باشد [۱]، [۲]. هم‌اکنون پژوهش‌های زیادی برای کاربرد حوزه تنک در حوزه‌های مختلف در جریان است. ارتباطات بیسیم، پردازش سیگنال‌های رادار و سونار، پردازش سیگنال‌های گفتار و موسیقی، تصویربرداری پزشکی، کاربردهای بینایی ماشین، یادگیری ماشین، داده‌کاوی، پردازش آرایه‌ای و ... چند نمونه از زمینه‌هایی هستند که در آن‌ها از مدل‌های حوزه تنک استفاده شده است [۳]. همچنین در زمینه پردازش تصویر مدل‌سازی حوزه تنک در کاربردهای عملی و مهندسی دارای کارایی بالا و موفق بوده و در سال‌های اخیر محبوبیت زیادی به دست آورده است. از جمله کاربردهای این حوزه در پردازش سیگنال‌های تصویر می‌توان به کاربردهای آن در تشخیص چهره^۲ [۴] و ساخت تصاویر با مقیاس و کیفیت بالا^۳ [۵]، فشرده‌سازی تصویر [۶]–[۱۰]، طبقه‌بندی تصاویر^۴ [۱۱]، ردیابی اشیاء در تصاویر [۱۲]، بازیابی تصویر^۵ [۱۳] و رفع محوشدگی یا ماتی تصویر^۶ [۱۴] اشاره کرد. برای مطالعه بیشتر در زمینه نقش حوزه تنک در پردازش تصویر می‌توان به مراجع [۲] و [۳] مراجعه کرد. حوزه نمایش تنک ابزارهایی برای ساخت دیکشنری‌های اختصاصی برای یک دسته یا کلاس از سیگنال‌ها معرفی می‌کند که با استفاده از آن می‌توان آن‌ها را به صورت تنک‌تر نمایش داد و همین امر حوزه نمایش تنک را در کاربردهای فشرده‌سازی یک کلاس از تصاویر جذاب‌تر می‌کند. در این مقاله روش جدیدی مبتنی بر نمایش تنک برای فشرده‌سازی تصاویر چهره ارائه شده است که بر اساس انتخاب وقتی یک دیکشنری مؤثرتر از میان چند دیکشنری آموزش دیده و کدگذاری تنک متناظر با آن عمل کرده و چنانچه نشان داده شده است می‌تواند ضمن حفظ یا ارتقاء کیفیت فشرده‌سازی به نرخ فشرده‌سازی بالاتری دست یابد.

در بخش‌های بعد این مقاله در ابتدا به معرفی نمایش سیگنال‌ها در حوزه تنک خواهیم پرداخت و پس از آن در مورد یادگیری دیکشنری مطالب مقدماتی لازم بیان خواهد شد. کلیات روش فشرده‌سازی تصویر مبتنی بر نمایش تنک در بخش چهارم بیان شده و در بخش پنجم روش پیشنهادی برای فشرده‌سازی تصاویر چهره با استفاده از ساخت دیکشنری اختصاصی ارائه می‌شود. در بخش ششم نتایج به دست آمده از شبیه‌سازی ارائه و بررسی شده‌اند. در بخش آخر نیز نتیجه‌گیری تحقیق آمده است.

۲- نمایش سیگنال در حوزه تُنک

اکثر سیگنال‌های طبیعی در یک یا چند حوزه (مانند زمان، فرکانس، موجک و ...) دارای نمایش تنک هستند، به این معنا که می‌توان اطلاعات موجود در آن‌ها را با استفاده از تعداد اندکی ضریب در یک حوزه نمایش خاص بیان کرد. مثلاً اگرچه تصاویر در حوزه مکانی بسیار انبوه یا غنی و با اطلاعات زیاد به نظر می‌رسند، اما در حوزه فرکانس دارای اطلاعات فشرده و به اصطلاح تنک هستند، به طوری که اغلب ضرایب فرکانسی آن‌ها صفر یا نزدیک به صفر است. همین موضوع در مورد بسیاری از انواع دیگر سیگنال‌ها هم صادق است، به طوری که می‌توان سیگنالی بسیار حجیم به طول N را تنها با K ضریب که $K \ll N$ است نمایش داد.

از خاصیت تنک بودن می‌توان در نمونه‌برداری و اندازه‌گیری سیگنال‌ها هم استفاده کرد. حسگری فشرده^۷ حوزه‌ای از پردازش سیگنال است که به این موضوع می‌پردازد [۱۵]، [۱۶]. مطابق قضیه نمونه‌برداری نایکوئیست-شانون می‌بایست تعدادی نمونه با نرخ مشخص از یک سیگنال باند محدود را در اختیار داشت تا بتوان آن را بازسازی کرد. اما اگر سیگنال موردنظر در حوزه‌ای خاص تنک باشد، می‌توان در اختیار داشتن تعداد بسیار کمتری اندازه‌گیری خطی آن را بازسازی کرد. در واقع با استفاده از حسگری فشرده به جای آنکه اول با نرخ بالا نمونه‌برداری کرده و بعد نمونه‌های موجود را فشرده کنند، از همان ابتدا سیگنال به نحوی "فشرده" اندازه‌گیری

¹ Sparse Representation

² Face Recognition

³ Super Resolution Images

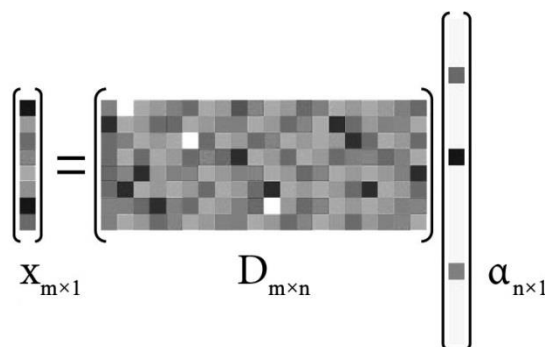
⁴ Image Classification

⁵ Image Restoration

⁶ Image Deblurring

⁷ Compressed Sensing

می‌شود. در شکل شماره (۱) یک دیاگرام نوعی از نمونه‌برداری به‌وسیله حسگری فشرده نشان داده شده است. در این دیاگرام سیگنال x با ابعاد $m \times 1$ به‌وسیله ماتریس D با ابعاد $m \times n$ ($m \ll n$) نمونه‌برداری می‌شود و سیگنال α را با اندازه‌ای به‌مراتب کوچک‌تر نتیجه می‌دهد. سیگنال اندازه‌گیری شده α را می‌توان به گیرنده ارسال کرد و اگر α به میزان کافی تنک بوده و ماتریس D با ابعاد $m \times n$ شرایط خاصی را ارضا کند، می‌توان x را از روی α بازسازی کرد. به الگوریتم‌های که برای یافتن α با مشاهده x و D به کار می‌روند، الگوریتم‌های بازبازی تنک^۱ یا کدگذاری تنک^۲ گفته می‌شود که دارای انواع مختلفی مانند الگوریتم‌های حریصانه^۳ و الگوریتم‌های محدب^۴ هستند [۳]، [۱۷]، [۱۸].



شکل (۱): نمایش تنک سیگنال X به‌وسیله ماتریس D

ماتریس D در واقع در بردارنده تمام پایه‌هایی است که برای نمایش داده‌ها استفاده می‌شود. این ماتریس دیکشنری^۵ نامیده می‌شود و تمام فضای برداری مربوط به داده‌ها را پوشش می‌دهد. هر کدام از ستون‌های ماتریس دیکشنری را یک اتم^۶ می‌نامند. در صورتی که تعداد اتم‌های دیکشنری برابر بعد فضای برداری باشد آنگاه آن دیکشنری را یک دیکشنری کامل^۷ می‌نامند و در این صورت هر کدام از داده‌ها نمایشی یکتا با استفاده از اتم‌های دیکشنری خواهد داشت [۱۹]. اگر تعداد اتم‌های یک دیکشنری کامل را بیشتر کنیم، به دیکشنری فراکامل^۸ گفته می‌شود که نمایش سیگنال با آن دارای جواب یکتا نخواهد بود. برای درک بهتر این موضوع دستگاه معادله خطی معادله (۱) را در نظر بگیرید. به دلیل اینکه ماتریس دیکشنری فراکامل است دستگاه معادله خطی در پیدا کردن α نامعین است و دارای بی‌شمار جواب خواهد بود. از آنجایی که به دنبال تنک‌ترین پاسخ برای این دستگاه هستیم، می‌توانیم مسئله را به فرم یک مسئله بهینه‌سازی به شکل معادله (۲) بیان کنیم (مسئله P_0 [۳]).

$$x = D\alpha, \quad (1)$$

$$P_0: \arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad x = D\alpha. \quad (2)$$

در این مسئله نرم l_0 به معنی تعداد عناصر غیر صفر بردار α است و به‌طور کلی مسئله بالا به معنی یافتن بردار α است که دارای کمترین مؤلفه غیر صفر باشد و یا به عبارت دیگر تنک‌ترین α ممکن. اگر در معادله (۲) میزان تنک بودن را به صورت یک قید یا محدودیت به مسئله اضافه کنیم آنگاه مسئله به فرم زیر درمی‌آید:

$$\arg \min_{\alpha} \|x - D\alpha\|_0 \quad s.t. \quad \|\alpha\|_0 \leq k. \quad (3)$$

¹ Sparse Recovery

² Sparse Coding

³ Greedy Algorithms

⁴ Convex Algorithms

⁵ Dictionary

⁶ Atom

⁷ Complete Dictionary

⁸ Over-Complete Dictionary

درواقع در اینجا هدف از حل مسئله کم کردن خطای حاصل از سیگنال بازسازی شده با نمایش تنک و سیگنال اولیه است. به خاطر اینکه حل مسئله P_0 در حالت کلی غیرممکن است ناگزیر باید مسئله را به فرم معادله (۳) حل کنیم و چون خطا و نویز بازسازی اجتنابناپذیر است در این حالت بردار α به صورت تخمینی بیان شده و معادله اولیه (۱) به فرم معادله زیر درمی آید.

$$x = D\alpha + s. \quad (4)$$

در این معادله s نویز نمایش نامیده می شود و $s \in \mathbb{R}^m$ است و دارای قید $\|s\|_2 \leq \varepsilon$ است. با حضور نویز پاسخ های تنک مسئله (۴) را می توان به صورت تقریبی با حل یکی از مسائل بهینه سازی زیر به دست آورد.

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|x - D\alpha\|_2^2 \leq \varepsilon, \quad (5)$$

$$\hat{\alpha} = \arg \min_{\alpha} \|x - D\alpha\|_2^2 \quad s.t. \quad \|\alpha\|_0 \leq \varepsilon. \quad (6)$$

با استفاده از قضیه ضرایب لاگرانژ می توان مسئله را با استفاده از ثابت λ به فرم زیر بازنویسی کرد [۳]، [۲۰]:

$$\hat{\alpha} = L(\alpha, \lambda) = \arg \min_{\alpha} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_0. \quad (7)$$

هرچند روش حل مسائل بهینه سازی تعریف شده در بالا (با استفاده از نرم l_0) می تواند از روش های بنیادین یافتن پاسخ تنک با استفاده از ماتریس D به دست آید ولی همچنان پاسخ مسئله غیرقطعی و به طور کلی مستلزم حل یک مسئله NP-HARD است [۲۱]. در [۲۲]، [۲۳] نشان داده شده است که پاسخ مسئله نرم l_0 را می توان با استفاده از نرم l_1 نیز به دست آورد که به مسئله بهینه سازی آن P_1 گفته می شود. در بسیاری از مسائل کاربردی پاسخ مسائل P_0 و P_1 معادل هستند و مسئله P_1 می تواند تقریب قابل قبولی برای یافتن نمایش تنک سیگنال باشد. اغلب ساختارهای مشهور نمایش تنک بر مبنای نرم l_1 بر اساس مسائل بهینه سازی زیر حل می شوند:

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|x - D\alpha\|_2^2 \leq \varepsilon \quad \text{یا} \quad \hat{\alpha} = \arg \min_{\alpha} \|x - D\alpha\|_2^2 \quad s.t. \quad \|\alpha\|_1 \leq \tau. \quad (8)$$

یکی دیگر از نرم هایی که توسط آن پاسخ تنک را به دست می آورند نرم l_2 است. پاسخ های به دست آمده توسط این نرم دارای دقت زیادی نبوده و ممکن است که به اندازه کافی تنک نباشند.

از معروف ترین الگوریتم های کدگذاری تنک که از روش های حریصانه بهره می برند می توان به الگوریتم های OMP^1 ، $ORMP^2$ و $WOMP^3$ اشاره کرد [۳]، [۲۴]، [۲۵]. همچنین الگوریتم های LASSO و گرادیان نزولی^۴ از تکنیک ملاپ سازی محدب^۵ برای حل مسائل فوق استفاده می کنند [۲۰].

۳- یادگیری دیکشنری

جهت نمایش تنک سیگنال ها به یک دیکشنری نیاز است. ساخت دیکشنری به وسیله الگوریتم های یادگیری دیکشنری صورت می گیرد. الگوریتم های یادگیری دیکشنری در دهه اخیر مورد توجه پژوهشگران این حوزه بوده است [۲۶]. از الگوریتم های ساخت دیکشنری می توان به الگوریتم MOD اشاره کرد که دارای کارایی بالایی است [۲۷]. از دیگر روش های ساخت دیکشنری الگوریتم K-SVD است که دارای پیچیدگی کمتر و کارایی مشابه الگوریتم MOD است [۲۸]، [۲۹].

¹ Orthogonal Matching Pursuit

² Order Recursive Matching Pursuit

³ Weak Orthogonal Matching Pursuit

⁴ Gradient Descent

⁵ Convex Relaxation

همان‌طور که در قسمت قبل گفته شد دیکشنری D ماتریسی به ابعاد $m \times n$ است که دارای n اتم به طول m هستند. در نمایش تنک سیگنال x می‌تواند به صورت ترکیب خطی یا تقریبی از اتم‌های دیکشنری باشد. تقریب یا بازسازی سیگنال x را با استفاده از دیکشنری D را می‌توان به صورت رابطه (۹) نشان داد که در آن برداری است شامل ضرایبی که بیشتر آن‌ها صفر هستند و r خطای نمایش است.

$$x = \hat{x} + r \quad (9)$$

$$\hat{x} = D\alpha$$

هدف از یادگیری دیکشنری در واقع یافتن ماتریسی است که کمترین ضرایب غیر صفر را در بردار α ایجاد نماید. در یادگیری دیکشنری یکی از مهم‌ترین مسائل دسترسی به یک مجموعه آموزشی است شامل بردارهای آموزشی که همگی با طول n هستند. اگر ماتریس داده‌های آموزش را به صورت $X = \{x_l\}_{l=1}^L, x_l \in \mathbb{R}^m$ نشان دهیم (شامل L بردار سیگنال به طول m)، هدف از پیدا کردن ماتریس D است این است که توسط آن بتوان سیگنال‌های آموزشی را به تنک‌ترین حالت ممکن نمایش داد و در واقع در مسئله بهینه‌سازی به شکل معادله (۱۰) صدق کند [۳].

$$\arg \min_{D,A} \|A\|_0 + \gamma \|X - DA\|_F^2 \quad (10)$$

در اینجا X ماتریس داده‌های آموزشی و $A = \{\alpha_l\}_{l=1}^L, \alpha_l \in \mathbb{R}^m$ ماتریس ضرایب (وزن‌های) تنک متناظر با آن است که در آن α_l به‌عنوان ستون‌های ماتریس ضرایب و $\|A\|_0 = \sum_{l=1}^L \|\alpha_l\|_0$ برابر تعداد ضرایب غیر صفر است. همچنین $\|X - DA\|_F^2 = \sum_{l=1}^L \|r_l\|_2^2$ برابر مجموع خطاهای به وجود آمده و γ ضریب لاگرانژ است. لازم به ذکر است معادله (۱۰) با توجه به جفت (D,A) یک مسئله بهینه‌سازی سخت هست و به‌طور کلی در اکثر روش‌های یادگیری دیکشنری برای حل این مسئله با استفاده از یک فرآیند بازگشتی دومرحله‌ای به حل آن می‌پردازند. این فرآیند با یک دیکشنری اولیه به صورت زیر آغاز می‌شود.

(a) با فرض ثابت بودن دیکشنری D ، ماتریس A پیدا می‌شود (مرحله کدگذاری تنک):

$$A^{(k+1)} = \arg \min_{D,A} \|A\|_0 + \gamma \|X - D^{(k)}A\|_F^2 \quad (11)$$

(b) با فرض ثابت بودن دیکشنری A ، ماتریس D پیدا می‌شود (مرحله آپدیت دیکشنری).

$$D^{(k+1)} = \arg \min_{D,A} \|A^{(k+1)}\|_0 + \gamma \|X - DA^{(k+1)}\|_F^2 \quad (12)$$

مرحله اول الگوریتم یک مسئله بهینه‌سازی کدگذاری تنک است که توسط روش‌های بیان‌شده در بخش قبل حل می‌شود و تقریباً در اکثر روش‌های یادگیری دیکشنری ثابت است. مرحله دوم مرحله به‌روزرسانی دیکشنری است. تفاوت بین اکثر روش‌های یادگیری در این مرحله است و در آن خطای نمایش مرحله اول کاهش می‌یابد.

علاوه بر روش‌های مورد اشاره در آموزش دیکشنری، یک روش یادگیری دیکشنری نسبتاً جدید دیگر در سال 2010 با نام 1 RLS-DLA معرفی شده است که روشی مؤثر و کارآمد در ساخت دیکشنری در شرایط برخط^۲، با قابلیت تطبیق با نمونه‌های آموزشی جدید و فراموشی مؤثر دیکشنری اولیه است [۳۰].

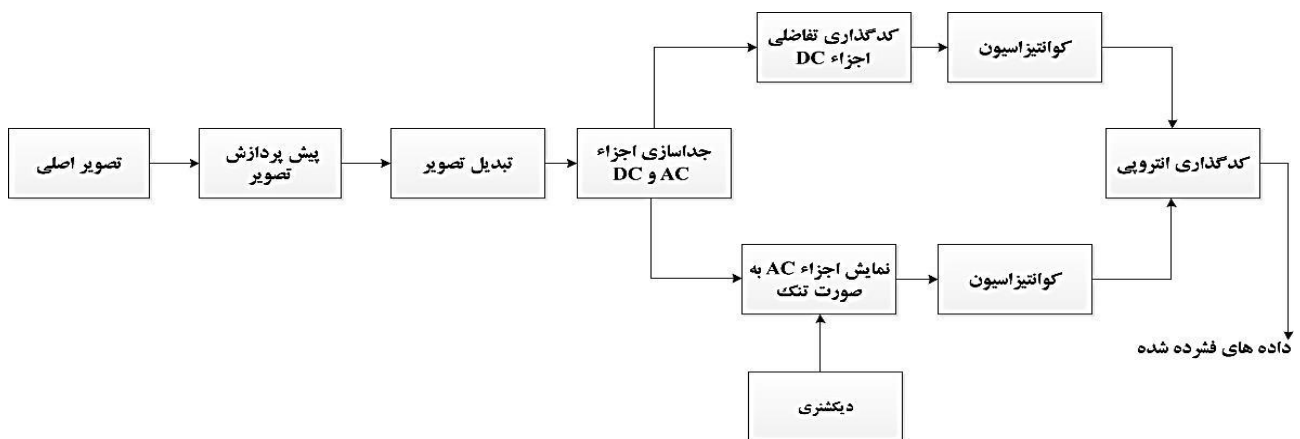
۴- روش مورد استفاده در فشرده‌سازی تصاویر با استفاده از دیکشنری

¹ Recursive Least Squares Dictionary Learning Algorithm

² Online

بلوک دیاگرام فرآیند کلی فشرده‌سازی تصویر با استفاده از نمایش تنک که در الگوریتم پیشنهادی از آن استفاده کرده‌ایم در شکل (۲) آمده است. این فرآیند به ترتیب شامل مراحل پیش‌پردازش^۱، تبدیل تصویر، جداسازی اجزا DC و AC، کدگذاری تنک، کدگذاری تفاضلی، کوانتیزاسیون و کدگذاری انتروپی است.

در مرحله پیش‌پردازش اطلاعات هر لایه تصویر به‌طور مستقل از دیگر لایه‌ها (یا اطلاعات روشنایی تصویر) به زیرتصویرهایی^۲ با اندازه 8×8 یا 16×16 یا 32×32 تقسیم شده و یک تبدیل روی هر زیرتصویر اعمال می‌شود. از تبدیل‌های مختلفی می‌توان در این مرحله استفاده کرد، مانند تبدیل کسینوسی (DCT) به‌کاررفته در JPEG و یا تبدیلات موجک مشهور همچون تبدیل موجک CDF 9/7 (4.4) (bi-orthogonal) یا CDF 5/3 (bi-orthogonal 2.2) که در JPEG2000 به‌کار رفته است. ضریب اول در تبدیل هر زیرتصویر که دارای بیشترین انرژی است ضریب DC و بقیه ضرایب AC نامیده می‌شوند. در این مرحله ضرایب DC و AC جداسازی و ذخیره می‌شوند. ضرایب DC توسط یک روش کدگذاری تفاضلی مانند DPCM کدگذاری و ذخیره شده و اجزاء AC به‌صورت بردارهایی ذخیره و با استفاده از یک دیکشنری و کدگذاری تنک نمایش داده می‌شوند. برای پیدا کردن ضرایب تنک بردارهای AC از الگوریتم OMP و یا ORMP می‌توان استفاده کرد که در اینجا به‌خاطر بالا بردن کیفیت سیگنال بازسازی از الگوریتم ORMP استفاده شده است. در کدگذاری تنک از خطای نمایش به‌عنوان معیار توقف الگوریتم استفاده می‌شود. بدین‌صورت که با استفاده از یک PSNR هدف داده‌شده، پارامتر ϵ یا معیار توقف تعیین می‌شود. پس از اعمال کدگذاری تنک به بردارهای ضرایب AC بردارها به‌صورت بردارهایی با کمترین تعداد عناصر غیر صفر درمی‌آیند. روشن است که هر چه بردارها تنک‌تر باشند عمل فشرده‌سازی تصویر به‌گونه‌ای بهتر انجام خواهد شد. پس از این که ضرایب AC و DC به‌طور جداگانه کد شدند، کدهای به‌دست‌آمده برای هر کدام به‌صورت جداگانه کوانتیزه می‌شوند. در ساده‌ترین حالت کوانتیزاسیون مقادیر ضرایب به نزدیک‌ترین مقدار صحیح گرد می‌شوند. همچنین می‌توان از کوانتایزر *midtread* بیان شده در [۳۱]، [۳۲] استفاده کرد. سپس ضرایب کوانتیزه به‌صورت یک رشته درآمده و توسط یک کدگذاری انتروپی برای فشرده‌سازی کد می‌گردند که برای این منظور در ابتدا عناصر غیر صفر و همچنین اطلاعات موقعیت آن‌ها در رشته به‌دست‌آمده مشخص و در رشته دیگر گذاشته می‌شوند. سپس هر دو رشته به‌دست‌آمده توسط یک کدگذاری انتروپی بهینه مانند هافمن کد شده و به‌صورت رشته‌های بیت درمی‌آیند که به برای کدبردار ارسال شده یا در یک فایل ذخیره می‌شوند. بدیهی است که در کدبردار دقیقاً عکس فرآیند مذکور انجام می‌گیرد تا تصویر اصلی بازیابی شود.



شکل (۲): بلوک دیاگرام الگوریتم فشرده‌سازی با استفاده از نمایش تنک

¹ Preprocessing

² Subimages

۵- روش پیشنهادی

در شکل (۳) سیستم پیشنهادی برای فشرده‌سازی تصاویر چهره با استفاده از چند دیکشنری نشان داده شده است. این سیستم شامل سه بخش کدگذار، آموزش و کدبردار است. در بخش کدگذار، تصویر چهره داده شده با استفاده از یک دیکشنری انتخابی فشرده می‌شود. در بخش آموزش، دیکشنری متناسب با تصاویر چهره‌ای که باید فرآیند فشرده‌سازی بر روی آن‌ها صورت گیرد ساخته می‌شود. الگوریتمی که برای آموزش دیکشنری به کار برده‌ایم الگوریتم RLS-DLA است که یکی از مزیت‌های مهم آن کاهش وابستگی به دیکشنری اولیه با استفاده از یک ضریب فراموشی λ است. در این الگوریتم از یک گام زمانی i برای معرفی ماتریس داده‌های آموزشی استفاده می‌شود که به صورت $X_i = [x_1, x_2, \dots, x_i]$ با ابعاد $m \times i$ تعریف می‌شود. هر گام زمانی i در واقع بیان‌کننده یک مرحله است. در هر حلقه تکرار یکی از بردارهای آموزشی به ماتریس افزوده می‌شود از این‌رو هر زمان که بخواهیم یک بردار آموزشی جدید به ماتریس داده‌ها اضافه کنیم کافی است که الگوریتم را با یک مرحله یا گام اضافی اجرا کنیم. به این دلیل این روش را یک روش یادگیری دیکشنری برخط می‌نامند [۳۳]. در این الگوریتم ماتریس ضرایب به صورت $A_i = [\alpha_1, \alpha_2, \dots, \alpha_i]$ با ابعاد $n \times i$ و ماتریس $C_i = (A_i A_i^T)^{-1}$ تعریف می‌شوند. در هر مرحله یا تکرار با استفاده از دیکشنری مرحله قبل یعنی D_{i-1} و یک الگوریتم کدگذاری تنک مانند OMP یا ORMP وزن‌ها یا ضرایب متناسب با بردارهای آموزشی جدید X_i پیدا می‌شوند. مرحله به‌روزرسانی دیکشنری با استفاده از یک لم ماتریس معکوس بر روی ماتریس C_i و به کار گرفتن ضریب فراموشی λ به صورت زیر است:

$$u = \left(\lambda_i^{-1} C_{i-1} \right) \alpha_i, \quad \lambda_i \leq 1, \quad (13)$$

$$\beta = \frac{1}{1 + \alpha_i^T u}, \quad (14)$$

$$r_i = x_i - D_{i-1} \alpha_i, \quad (15)$$

$$C_i = \left(\lambda_i^{-1} C_{i-1} \right) - \beta u u^T, \quad (16)$$

$$D_i = D_{i-1} + \beta r_i u^T. \quad (17)$$

در اینجا با به‌کارگیری یک ضریب فراموشی تطبیقی $\lambda_i \leq 1$ در مرحله نام میزان وابستگی به دیکشنری اولیه کمتر شده و همگرایی الگوریتم بهبود می‌یابد [۳۰]. مقدار λ در صورت ثابت در نظر گرفته شدن مقداری نزدیک به 1 مثلاً حدود $1 - \frac{1}{L} \approx \lambda$ در نظر گرفته شده و برای λ_i تطبیقی مقدار آن به صورت بازگشتی تغییر می‌کند.

الگوریتم RLS-DLA در کاربردهایی که در آن‌ها از داده‌های آموزشی بزرگ استفاده می‌شود، روش بسیار مناسبی است. برای مقداردهی دیکشنری اولیه از n بردار اولیه آموزشی استفاده شده و توسط آن‌ها ماتریس اولیه C_i نیز به دست می‌آید. در این الگوریتم از معیار توقف ϵ برای تقریب خطا استفاده می‌شود که بر اساس آن می‌توان یک PSNR هدف^۳ را در آغاز الگوریتم یادگیری دیکشنری محاسبه کرد. در این الگوریتم به کل بازگشت‌هایی که به تعداد بردارهای آموزش داریم یک آزمایش^۴ گفته می‌شود و همچنین یک تعداد تکرار برای هر آزمایش مشخص می‌شود تا الگوریتم به PSNR مطلوب یا هدف خود همگرا شود [۳۰]. برای ساخت دیکشنری باید از یک زیرمجموعه تصاویر نمونه از تصاویر چهره‌ای که باید عمل فشرده‌سازی بر روی آن‌ها صورت پذیرد استفاده شود. به هنگام ساخت دیکشنری موردنظر باید عملیات پیش‌پردازش و تبدیل تصویر متناظر با طرح فشرده‌سازی موردنظر انتخاب شود. به‌عنوان مثال در صورتی که در فشرده‌سازی تصویر چهره موردنظر در بخش پیش‌پردازش از زیرتصویرهایی با اندازه 8×8 استفاده کرده باشیم و در بخش تبدیل تصویر، تبدیل موجک CDF 9/7 را بر روی زیر تصویرها به کار رود در این صورت باید برای ساخت دیکشنری تصاویر آموزشی مورد نظر هم به

¹ Forgetting Factor

² Time Step

³ Target PSNR

⁴ Trial

زیرتصویرهایی با ابعاد 8×8 تقسیم‌بندی شوند و همان تبدیل موجک CDF 9/7 بر روی آن‌ها اعمال شود. همچنین در هنگام آموزش باید ضرایب DC هر زیرتصویر نیز جدا شود و سپس دیکشنری موردنظر آموزش داده شود تا در فرآیند کدگذاری تنک ضرایب AC به کار گرفته شود. پس از ساخت دیکشنری از آن برای فشرده‌سازی تصویر مورد نظر استفاده می‌شود.

داده‌های فشرده‌شده پس از ارسال در کانال توسط کدبردار به‌منظور بازیابی تصویر اولیه دریافت می‌شوند. در بخش کدبرداری انتروپی عکس عمل کدگذاری انتروپی بر روی داده‌ها صورت می‌گیرد و یک رشته از ضرایب کوانتیزه شده AC و DC به دست می‌آیند. سپس عمل کوانتیزاسیون معکوس بر روی آن‌ها انجام می‌شود تا ضرایب AC و DC کد شده به دست آیند. به‌منظور بازیابی عمل معکوس کدگذاری تفاضلی بر روی ضرایب DC صورت می‌گیرد. برای بازیابی اجزاء AC، با استفاده از دیکشنری متناظر موجود در کدبردار، کدبرداری انجام می‌شود. پس از به دست آوردن اجزاء AC و DC آن‌ها باهم ترکیب می‌شوند تا زیرتصویر تبدیل‌شده تصویر به دست آید. درنهایت با اعمال یک تبدیل موجک معکوس مقادیر شدت روشنایی زیرتصویرها به دست می‌آیند و با کنار هم قرار دادن آن‌ها تصویر بازیابی شده ایجاد می‌شود.

در طرح پیشنهادی برخلاف روش‌های موجود، به‌جای استفاده از یک دیکشنری برای فشرده‌سازی تصویر چهره موردنظر در تمامی کیفیت‌ها، به‌طور فوقی از چند دیکشنری که هر دیکشنری مطابق با کیفیت فشرده‌سازی تصویر چهره موردنظر طراحی شده است استفاده می‌شود. نحوه عملکرد آن بر اساس انتخاب یک دیکشنری منطبق با مقدار کیفیت تصویری که قرار است فشرده شود، (PSNR مطلوب تصویر موردنظر) است. درواقع برای فشرده‌سازی هر تصویر با کیفیت موردنظر با استفاده از یک الگوریتم تصمیم‌گیری یک دیکشنری از میان چند دیکشنری ساخته‌شده برای فشرده‌سازی تصویر موردنظر انتخاب می‌شود. علت استفاده از الگوریتم تصمیم‌گیری و انتخاب یک دیکشنری از بین چند دیکشنری آن است که در هنگام فشرده‌سازی تصاویر چهره با کیفیت‌های مختلف با استفاده از دیکشنری ساخته‌شده با الگوریتم یادگیری دیکشنری RLS-DLA به‌طور تجربی دریافته‌ایم که PSNR تصاویر فشرده‌شده، متناسب با Target PSNR دیکشنری ساخته‌شده در برخی بازه‌های کیفیت موردنظر در گیرنده (کدبردار) عملکرد نرخ-اعوجاج بالاتری (بالاتر از JPEG 2000) دارند و همین موضوع را مبنای کار الگوریتم تصمیم‌گیری قرار دادیم. رابطه کیفیت فشرده‌سازی موردنظر تصویر با Target PSNR دیکشنری ساخته‌شده به‌صورت زیر تعیین گردید:

$$Dic.TargetPSNR - 5 < Desired\ Compression\ PSNR \leq Dic.TargetPSNR \quad (18)$$

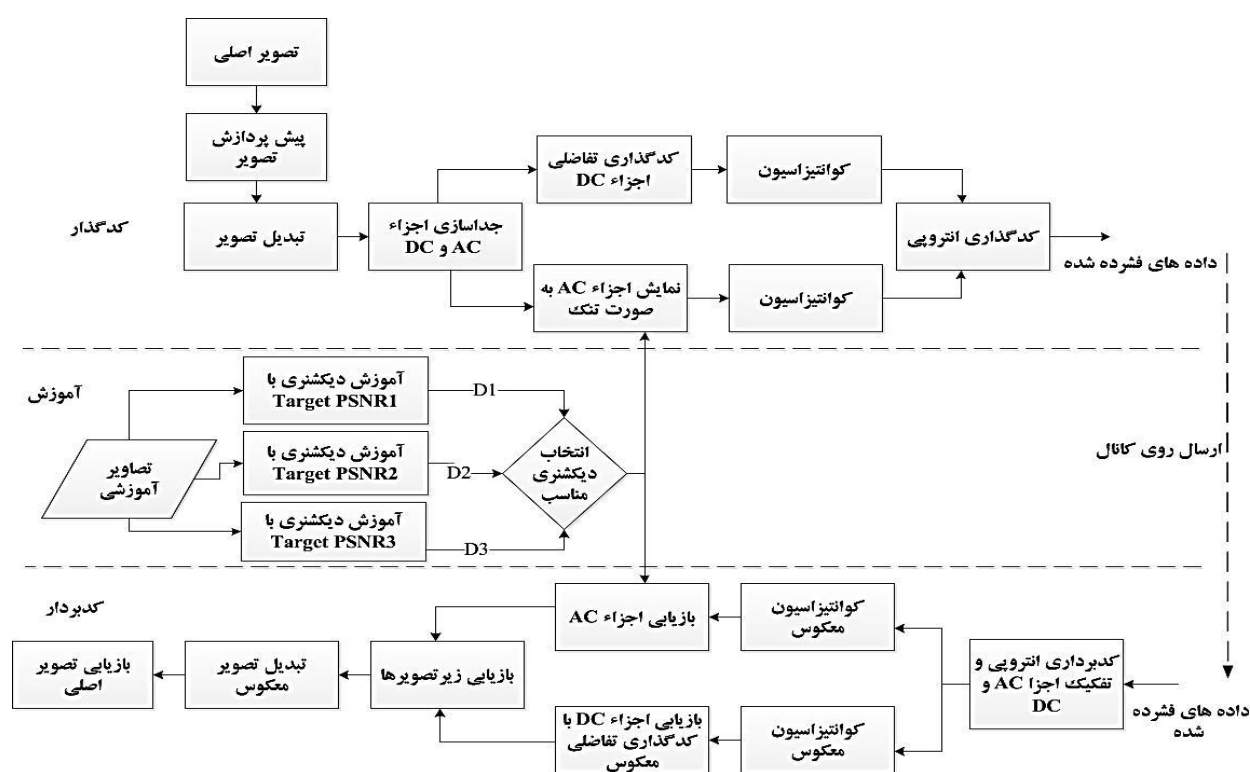
در روش پیشنهادی هنگام فشرده‌سازی تصویر چهره موردنظر با استفاده از نامساوی (۱۸) یک دیکشنری آموخته که در این رابطه صدق کند انتخاب می‌شود. مقادیر نوعی برای کیفیت تصویر فشرده‌شده یا مقادیر نوعی PSNR های تصویر در فشرده‌سازی با تالاف تصاویر چهره برای نرخ‌های فشرده‌سازی 0.10 تا 2.00 بیت بر پیکسل، بین 35 تا 55 دسی‌بل است. از این‌رو به نظر می‌رسد ساخت تنها سه دیکشنری برای سیستم پیشنهادی کفایت می‌کند. برای آموزش دیکشنری‌های موردنظر با PSNR های متفاوت باید از معیارهای توقف مختلف (E های متفاوت) استفاده شود و همچنین دیکشنری‌ها باید از تصاویر چهره نمونه از میان یک پایگاه داده یا تصاویر با موضوع یکسان ساخته شوند.

لازم به ذکر است که دیکشنری‌های موردنظر تنها یک‌بار ساخته شده و در ساختار الگوریتم فشرده‌سازی ذخیره می‌شوند و در کدگذار و هم در کدبردار به کار گرفته می‌شوند. لذا روش پیشنهادی در مرحله اجرای فشرده‌سازی صرفاً محدود به تعیین ضرایب تنک با استفاده از الگوریتم مؤثر و سریع ORMP است و نیازی به آموزش مجدد دیکشنری ندارد. به این ترتیب این روش بار محاسباتی یا زمان عملکرد بیشتری را در مقایسه با دیگر روش‌های متداول و پیشرفته فشرده‌سازی به سیستم تحمیل نمی‌کند (در کدبردار یا کدگذار) و تنها ذخیره دو ماتریس دیکشنری اضافی را لازم دارد.

۶- نتایج شبیه‌سازی الگوریتم ارائه‌شده فشرده‌سازی تصاویر چهره

به‌منظور پیاده‌سازی و شبیه‌سازی روش ارائه‌شده از نرم‌افزار MATLAB 2015a و یک کامپیوتر شخصی با توان پردازشی 2 GHz و مقدار حافظه رم 8 GB و سیستم‌عامل Windows 7 استفاده شده است. به‌منظور ساخت دیکشنری و نمایش تنک ضرایب از روش

کدگذاری تنک ORMP استفاده شده است. این کدگذار از سرعت و دقت بیشتری نسبت به روش‌های مشابه همچون OMP برخوردار است. برای ساخت دیکشنری‌های موردنظر از الگوریتم RLS-DLA استفاده شده است.^۱ برای یادگیری دیکشنری ۸ تصویر آموزشی، مطابق شکل (۴)، متشکل از تصاویر ۴ مرد و ۴ زن که به صورت تصادفی از میان تصاویر پایگاه داده FEI^۲ انتخاب شده‌اند، بکار رفته‌اند. هر کدام از این تصاویر دارای ابعاد 640×480 پیکسل هستند. لازم به ذکر است که استفاده از تصاویر بیشتر در آموزش دیکشنری به دلیل تنوع بیشتر ضرایب می‌تواند منجر به ساخت دیکشنری بهتر و فشرده‌سازی دقیق‌تر تصاویر شود، ولی این امر باعث افزایش زمان آموزش دیکشنری نیز می‌شود. در اینجا سه دیکشنری با استفاده از تصاویر آموزشی با Target PSNR های 55 و 50 و 45 دسی‌بل ساخته شدند و برای ساخت هر دیکشنری از 1000 حلقه تکرار به منظور دستیابی به کیفیت موردنظر (Target PSNR) استفاده شده است. زمانی که برای آموزش هر دیکشنری طی شده به ترتیب برابر 399.1 و 315.6 و 135.7 دقیقه است. دیکشنری‌های ساخته‌شده شامل 440 اتم با ابعاد 8×8 پیکسل می‌باشند (با ستون‌های به طول 64 درایه). هر دیکشنری از هر تصویر آموزشی 3000 تکه تصویر با ابعاد 8×8 پیکسل به‌طور تصادفی انتخاب می‌کند؛ یعنی جمعاً 24000 زیر تصویر برای آموزش دیکشنری‌ها به کار رفته است که قبل از ارائه به الگوریتم آموزش RLS-DLA به حوزه موجک CDF 9/7 تبدیل شده‌اند.



شکل (۳): بلوک دیاگرام سیستم پیشنهادی برای فشرده‌سازی تصاویر چهره

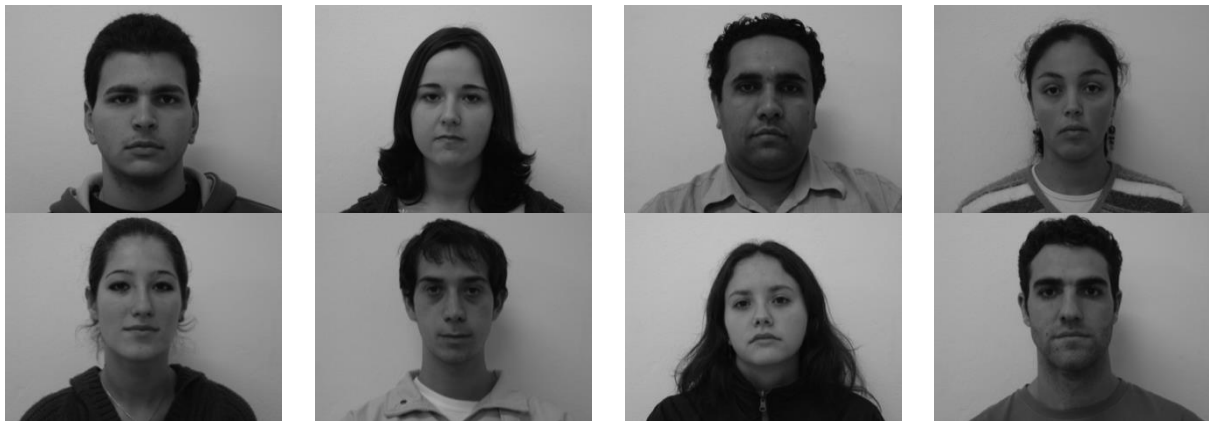
^۱ برای به‌کارگیری ابزارهای مختلف فشرده‌سازی و همچنین الگوریتم یادگیری دیکشنری از ابزارهای موجود در آدرس‌های زیر استفاده شده است:

<http://www.ux.uis.no/~karlsk/dle/index.html#ssec61>

<http://www.ux.uis.no/~karlsk/ICTools/ictools.html>

^۲ تصاویر چهره پایگاه داده FEI از طریق آدرس زیر موجود و قابل دسترسی می‌باشند:

<http://fei.edu.br/~cet/facedatabase.html>



شکل (۴): مجموعه تصاویر آموزش برای یادگیری دیکشنری با ابعاد 640×480 پیکسل

پس از ساخته شدن دیکشنری‌های مورد نظر، از دو تصویر چهره متفاوت از تصاویر آموزشی برای آزمایش الگوریتم فشرده‌سازی استفاده شد و فرآیند فشرده‌سازی و بازیابی تصویر بر روی آن‌ها با نرخ بیت‌های متفاوت صورت گرفته است و پس از آن کیفیت تصاویر بازیابی شده اندازه‌گیری شده‌اند. در شکل (۵) تصاویر آزمایشی مذکور ارائه شده است. ابعاد این تصاویر نیز 640×480 بوده و دارای عمق بیت 256 (8 بیت) هستند. کدگذاری انتروپی که در شبیه‌سازی الگوریتم ارائه شده بکار رفته، کدگذاری هافمن است که پیاده‌سازی آن نسبت به دیگر کدگذاری‌های انتروپی ساده‌تر است همچنین زمان پردازش آن نسب به روش‌های کدگذاری انتروپی دیگر مانند کدگذاری حسابی بسیار کمتر است. پس از فشرده‌سازی تصاویر مورد آزمایش با الگوریتم ارائه شده دو تصویر مورد نظر با دیگر روش‌های فشرده‌سازی JPEG2000 و JPEG و همچنین روش فشرده‌سازی مبتنی بر دیکشنری MOD با همان نرخ‌های فشرده‌سازی الگوریتم ارائه شده نیز فشرده و بازیابی شدند که نتایج حاصل از آن‌ها در شکل‌های (۶) و (۷) و جدول‌های (۱) و (۲) آمده است.

در شکل (۸) کیفیت تصاویر آزمایشی بازیابی شده در ناحیه چشم با الگوریتم‌های فشرده‌سازی در نرخ بیت یکسان 0.25 bpp نشان داده شده است. همان‌طور که مشاهده می‌شود کیفیت تصاویر بازیابی شده با استفاده از الگوریتم پیشنهادی از لحاظ بصری بهتر از بقیه الگوریتم‌ها است.

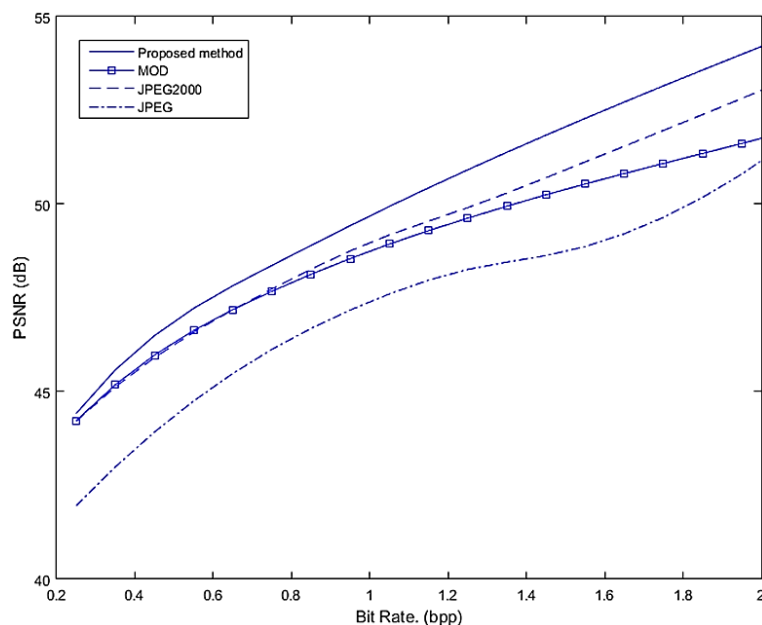


شکل (۵): دو تصویر مورد آزمایش در فشرده‌سازی با الگوریتم‌های مختلف، راست: تصویر شماره ۱، چپ: تصویر شماره ۲.

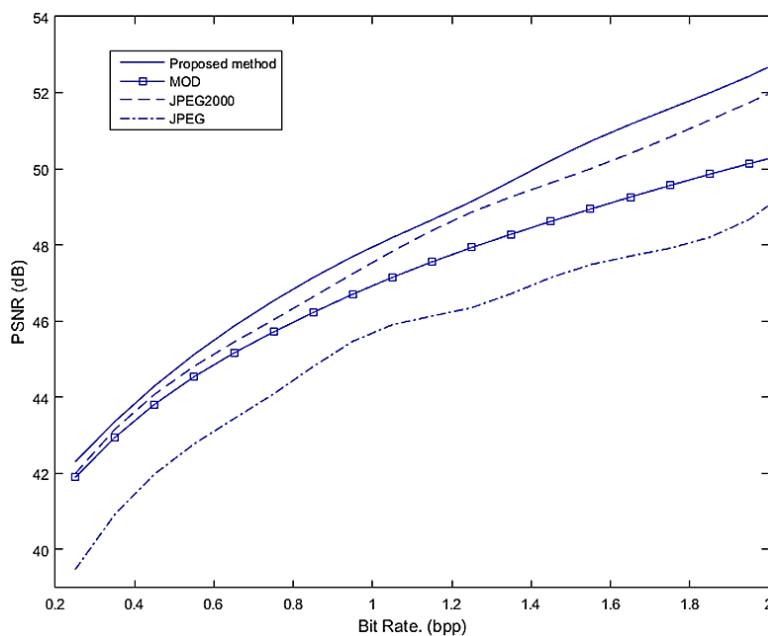
تصاویر مورد آزمایش در شبیه‌سازی‌ها با نرخ بیت 0.25 تا 2.00 بیت در پیکسل یعنی با نرخ فشرده‌سازی حدود 4 تا 32 برابر توسط روش‌های مطرح شده فشرده شدند. نتایج به دست آمده حاکی از آن است که روش فشرده‌سازی ارائه شده کیفیت بالاتری در حدود 0.5 تا 1.2 دسی‌بل نسبت به روش پیشرفته و متداول JPEG2000 ارائه می‌کند که برتری قابل توجهی است. می‌دانیم که با توجه به لگاریتمی بودن واحد مورد استفاده جهت کیفیت بازسازی، بهبود تنها نیم دسی‌بل معادل کاهش بیش از ۱۰٪ در میانگین مربع خطا و بهبود به میزان یک دسی‌بل معادل کاهش بیش از ۲۰٪ در آن است. نمونه این برتری را می‌توان در شکل (۸) به صورت کیفی دریافت. یکی دیگر از برتری‌های روش پیشنهادی نسبت به روش‌های مشابه فشرده‌سازی تصویر با استفاده از تجزیه تنک، نیاز به استفاده از مجموعه آموزشی کوچک تصاویر آموزشی است به طوری که نتایج به دست آمده تنها با ۸ تصویر آموزشی حاصل شده است. با به کارگیری روش‌های فشرده‌سازی انتروپی بهینه پیچیده‌تر مانند کدگذاری حسابی و استفاده از تصاویر آموزشی بیشتر می‌توان کیفیت فشرده‌سازی را بازهم بهبود داده و در فضای ذخیره‌سازی پایگاه داده تصاویر به مقدار زیادی صرفه‌جویی نمود.

ارائه یک روش جدید فشرده‌سازی تصاویر چهره با استفاده از نمایش تنک سیگنال و الگوریتم

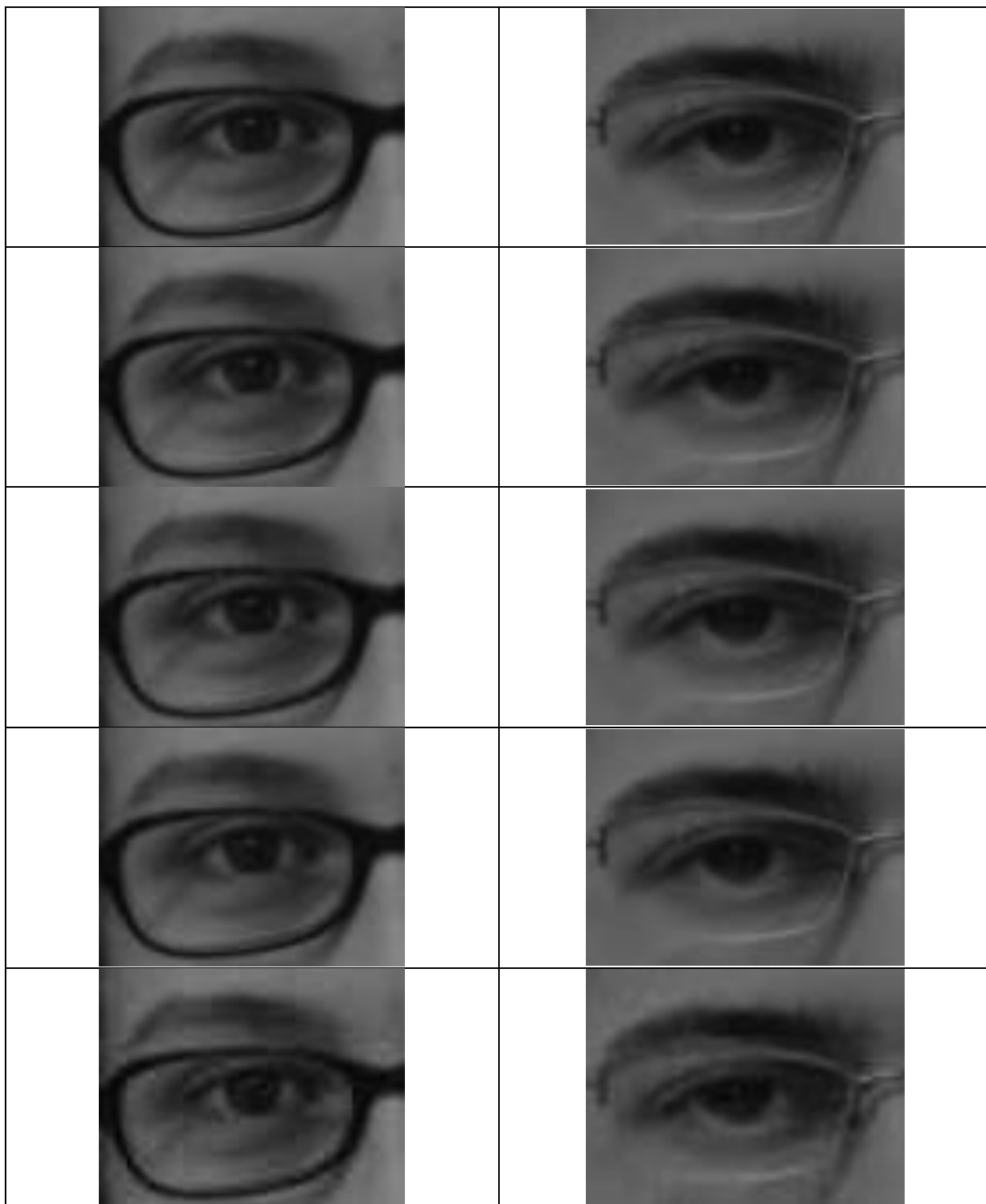
مجدداً متذکر می‌شویم که با توجه به عدم نیاز به آموزش دیکشنری در مرحله فشرده‌سازی، زمان اجرای این روش حتی می‌تواند کوتاه‌تر از دیگر روش‌های متداول باشد، اگرچه با توجه به بهینه نبودن برنامه ارائه‌شده در این مقاله (با اجرا در محیط MATLAB) و تفاوت شرایط پیاده‌سازی در مقایسه با روش‌هایی مانند JPEG2000، مقایسه آن‌ها در شرایط موجود عملاً ارزشمند و قابل قبول نخواهد بود.



شکل (۶) مقایسه فشرده‌سازی تصویر آزمایشی شماره ۱ توسط روش‌های مختلف



شکل (۷) : مقایسه فشرده‌سازی تصویر آزمایشی شماره ۲ توسط روش‌های مختلف



شکل (۸) : کیفیت دو تصویر آزمایشی فشرده شده در ناحیه چشمها (راست: تصویر آزمایشی شماره ۱، چپ: تصویر آزمایشی شماره ۲) با نرخ 0.25 bpp با الگوریتم‌های مختلف. تصاویر به ترتیب از بالا به پایین: تصویر اصلی، تصویر فشرده شده با روش ارائه شده، تصویر فشرده شده با MOD، تصویر فشرده شده با JPEG2000، تصویر فشرده شده با JPEG.

جدول (۱): نتایج کیفیت فشرده‌سازی (PSNR) تصویر آزمایشی ۱ با استفاده از الگوریتم‌های مختلف برحسب dB با نرخ بیت‌های مختلف

روش فشرده‌سازی	نرخ بیت فشرده‌سازی (bpp)							
	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
روش ارائه‌شده	44.71	46.87	48.36	49.68	50.91	52.06	53.15	54.20
MOD	44.22	46.31	47.67	48.74	49.62	50.39	51.08	51.75
JPEG2000	44.18	46.25	47.73	48.97	49.90	50.91	51.96	53.03
JPEG	41.95	44.34	46.12	47.39	48.25	48.73	49.64	51.14

جدول (۲): نتایج کیفیت فشرده‌سازی (PSNR) تصویر آزمایشی ۲ با استفاده از الگوریتم‌های مختلف برحسب dB با نرخ بیت‌های مختلف

روش فشرده‌سازی	نرخ بیت فشرده‌سازی (bpp)							
	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
روش ارائه‌شده	42.51	44.72	46.53	47.95	49.14	50.48	51.58	52.67
MOD	41.89	44.19	45.71	46.93	47.93	48.78	49.56	50.27
JPEG2000	41.99	44.46	46.03	47.54	48.86	49.81	50.83	51.97
JPEG	39.47	42.40	44.08	45.72	46.35	47.33	47.91	49.00

۷- نتیجه‌گیری

در این تحقیق با بررسی روش‌ها و الگوریتم‌های موجود برای فشرده‌سازی تصویر با استفاده از حوزه تنک روش جدیدی برای فشرده‌سازی تصاویر چهره پیشنهاد گردید که باعث بهبود کیفیت تصاویر فشرده‌شده نسبت به دیگر الگوریتم‌های فشرده‌سازی در نرخ بیت یکسان می‌شود. در این الگوریتم با بررسی و تجزیه و تحلیل نتایج متوجه شدیم که کیفیت تصاویر فشرده‌شده با کیفیت دیکشنری (مقدار خطای دیکشنری) ساخته‌شده مرتبط است و همین امر را اساس ساخت الگوریتم پیشنهادی قرار دادیم. بدین صورت که با استفاده از الگوریتم آموزش RLS سه دیکشنری با کیفیت‌های متفاوت ساخته شدند و هنگام فشرده‌سازی دیکشنری منطبق با کیفیت تصویر مورد نظر در فشرده‌سازی انتخاب گردید. نتایج شبیه‌سازی حاصل از الگوریتم پیشنهادی نشان‌دهنده بهبود کیفیت تصاویر فشرده‌شده و برتری آن نسبت به دیگر روش‌های فشرده‌سازی است، به طوری که نسبت به الگوریتم پیشرفته JPEG2000 که عملاً بالاترین کیفیت را در بین الگوریتم‌های فشرده‌سازی ارائه کرده و به عنوان یک استاندارد در فشرده‌سازی تصویر شناخته‌شده است، در حدود 0.5 تا 1.2 dB برتر است. در روش ارائه‌شده استفاده از کدگذاری انتروپی پیشرفته‌تری همچون کدگذاری حسابی یا هافمن تطبیقی می‌تواند مورد بررسی قرار گرفته و کیفیت و نرخ فشرده‌سازی را ارتقا دهد. همچنین می‌توان تأثیر دیکشنری‌های مختلف در فشرده‌سازی تصاویر با ابعاد متغیر را مورد بررسی قرار داد. از الگوریتم مطرح شده می‌توان در فشرده‌سازی تصاویر پایگاه‌های داده بزرگ با چند میلیون تصویر مانند داده‌های سازمان‌ها، دانشگاه‌ها و بیمارستان‌ها بهره جست و به میزان قابل توجهی در فضای ذخیره‌سازی صرفه‌جویی نمود که این امر نیز به نوبه خود باعث کاهش پهنای باند اشغالی جهت تبادل داده در شبکه‌های کامپیوتری و مخابرات تصویری می‌شود.

مراجع

- [1] M. Elad, "Sparse and redundant representation modeling-What next?," *IEEE Signal Process. Lett.*, vol. 19, pp. 922–928, 2012.
- [2] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, pp. 972–982, 2010.
- [3] M. Elad, *Sparse and redundant representations: From theory to applications in signal and image processing*. 2010.
- [4] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, pp. 210–227, 2009.
- [5] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image Super-Resolution via Sparse Representation," *IEEE Trans. Image Process.*, vol. 19, pp. 2861–2873, 2010.
- [6] K. Skretting and K. Engan, "Image compression using learned dictionaries by RLS-DLA and compared with K-SVD," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 1517–1520, 2011.
- [7] O. Bryt and M. Elad, "Compression of facial images using the K-SVD algorithm," *J. Vis. Commun. Image Represent.*, vol. 19, no. 4, pp. 270–282, 2008.
- [8] J.-Y. Zhu, Z.-Y. Wang, R. Zhong, and S.-M. Qu, "Dictionary based surveillance image compression," *J. Vis. Commun. Image Represent.*, vol. 31, pp. 225–230, 2015.
- [9] I. Horev, O. Bryt, and R. Rubinstein, "Adaptive image compression using sparse dictionaries," in *Systems, Signals and Image Processing (IWSSIP), 2012 19th International Conference on*, 2012, pp. 592–595.
- [10] G. Shao, Y. Wu, X. Liu, and T. Guo, "Fingerprint Compression Based on Sparse Representation," *Image Process. IEEE Trans.*, vol. 23, no. 2, pp. 489–501, 2014.
- [11] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 1794–1801.
- [12] X. Mei and H. Ling, "Robust visual tracking and vehicle classification via sparse representation," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 33, no. 11, pp. 2259–2272, 2011.
- [13] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *Image Process. IEEE Trans.*, vol. 17, no. 1, pp. 53–69, 2008.
- [14] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *Image Process. IEEE Trans.*, vol. 20, no. 7, pp. 1838–1857, 2011.
- [15] Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," *Signal Processing*, vol. 86, no. 3, pp. 549–571, 2006.
- [16] D. L. Donoho, "Compressed sensing," *Inf. Theory, IEEE Trans.*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [17] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *Signal Process. IEEE Trans.*, vol. 53, no. 7, pp. 2477–2488, 2005.
- [18] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, pp. 948–958, 2010.
- [19] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, pp. 1045–1057, 2010.
- [20] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *Access, IEEE*, vol. 3, pp. 490–530, 2015.
- [21] E. Amaldi and V. Kann, "On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems," *Theor. Comput. Sci.*, vol. 209, no. 1, pp. 237–260, 1998.
- [22] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2231–2242, 2004.
- [23] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, 2009.
- [24] M. Gharavi-Alkhansari and T. S. Huang, "A fast orthogonal matching pursuit algorithm," in *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, 1998, vol. 3, pp. 1389–1392.
- [25] E. Liu and V. N. Temlyakov, "The orthogonal super greedy algorithm and applications in compressed sensing," *Inf. Theory, IEEE Trans.*, vol. 58, no. 4, pp. 2040–2047, 2012.
- [26] I. Tošić and P. Frossard, "Dictionary learning," *Signal Process. Mag. IEEE*, vol. 28, no. 2, pp. 27–38, 2011.
- [27] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," *1999 IEEE Int. Conf. Acoust. Speech, Signal Process. Proceedings. ICASSP99 (Cat. No.99CH36258)*, vol. 5, 1999.
- [28] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, pp. 4311–4322, 2006.
- [29] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse

- model,” *IEEE Trans. Signal Process.*, vol. 61, pp. 661–677, 2013.
- [30] K. Skretting and K. Engan, “Recursive least squares dictionary learning algorithm,” *Signal Process. IEEE Trans.*, vol. 58, no. 4, pp. 2121–2130, 2010.
- [31] D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice: Image Compression Fundamentals, Standards and Practice*, vol. 642. Springer Science & Business Media, 2012.
- [32] Y. Q. Shi and H. Sun, “Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards,” 2008.
- [33] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online dictionary learning for sparse coding,” in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 689–696.

A novel face images compression scheme using sparse signal representation and RLS_DLA dictionary learning algorithm

Amir Masoud Taheri¹, Homayoun Mahdavi-Nasab^{2*}

1,2: Islamic Azad University, Najaf Abad Branch, Isfahan, Iran

Abstract

Due to the rapid growth of information technology and exponential increasing of information the need for more and more storage capacity and efficiency has increased. Image compression is an important tool to reduce the redundancy of images data in order to be able to store or transmit them in an efficient manner. When images are limited to a specific and limited family of images like MRI databases of a hospital or facial image database of a university or an organization or fingerprint image databases, this limitation increases the total spatial redundancy. Thus, efficient storage of such images is beneficial, and their compression becomes an appealing application, and this urges algorithms specially tailored for the task of content base image compression to surpass general purpose compression algorithms. The facial images, due to their wide application as the most common images in the organizations and companies are more considerable for image compression. In this paper a new image compression scheme using sparse coding and RLS-DLA redundant dictionary learning is proposed that can be used for compressing of face image databases. In the proposed method, several dictionaries are exploited adaptively based on the required image quality to enhance the overall rate-distortion. The simulation results show that this scheme outperforms the state-of-art algorithms like JPEG2000 by about 0.5 to 1.2 dB for reconstructed images PSNR.

Keywords: Sparse representation, sparse coding, image compression, dictionary learning, jpeg2000