

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

امیرمسعود طاهری<sup>(۲)</sup>، همایون مهدوی نسب<sup>(۱،۲)\*</sup>

(۱) مرکز تحقیقات پردازش دیجیتال و بینایی ماشین، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران

(۲) دانشکده مهندسی برق، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران

### چکیده

رشد فناوری و افزایش تصاعدی اطلاعات نیاز به حجم ذخیره سازی بیشتر داده های اطلاعاتی گوناگون را افزایش داده است. در این راستا فشرده سازی تصویر به عنوان ابزاری کارآمد جهت کاهش افزونگی و صرفه جویی در حجم ذخیره سازی و کاهش پهنای باند انتقالی داده های تصویری به کار می آید. هنگامی که فشرده سازی 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 فشرده سازی اهمیت و الزام بیشتری پیدا می کند. در این میان تصاویر چهره با توجه به کاربرد وسیعی که به عنوان رایج ترین تصاویر پایگاه داده سازمان های و نهادهای مختلف مانند اداره های پلیس، نهادهای نظامی، دانشگاه ها و شرکت های 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 فشرده سازی تصویر ارائه شده است که می تواند برای فشرده سازی پایگاه داده تصاویر به کار رود. در این الگوریتم تصاویر با به کارگیری چند دیکشنری به نحو وفقی بر اساس کیفیت بازسازی مورد نیاز آن ها فشرده می شوند. نتایج به دست آمده از الگوریتم پیشنهادی نشان دهنده عملکرد مؤثر و برتری معنی دار آن نسبت به روش های پیشرفته و مطرحی همچون JPEG2000 است به طوری که به افزایش کیفیتی در حدود 0.5 dB تا 1.2 dB در نرخ بیت یکسان دست می یابد.

واژه های کلیدی: نمایش تنک، کدگذاری تنک، فشرده سازی تصویر، یادگیری دیکشنری، RLS-DLA، JPEG2000.

### ۱- مقدمه

بر تبدیل موجک و با کارایی بالا است. هنگامی که فشرده سازی یک دسته یا خانواده از تصاویر مدنظر باشد افزونگی اطلاعات افزایش یافته و باعث می شود تا فشرده سازی اهمیت بیشتری پیدا کند. فشرده سازی تصاویر چهره به عنوان تصاویری پر کاربرد که به طور گسترده در پایگاه های داده بزرگ همچون شبکه های اجتماعی، وب سرورها و سازمان های مختلف مانند سازمان ثبت احوال مدارس، دانشگاه ها و شرکت های خصوصی

فشرده سازی تصویر همواره یکی از دغدغه های اصلی پژوهشگران حوزه پردازش سیگنال بوده و همچنان از زمینه های فعال تحقیقاتی است. در این راستا روش های زیادی ارائه شده و توسعه یافته اند که از جمله آن ها می توان به استانداردهای JPEG و JPEG2000 اشاره کرد که دومی هم اکنون یکی از بهترین روش های فشرده سازی تصویر مبتنی

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 وفقی 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 تنک

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

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 تنک<sup>۱</sup> 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<sup>۲</sup> [۴] و ساخت تصاویر با مقیاس و کیفیت بالا<sup>۳</sup> [۵]، فشرده‌سازی تصویر [۶]-[۱۰]، طبقه‌بندی تصاویر<sup>۴</sup> [۱۱]، ردیابی اشیا در تصاویر [۱۲]، بازیابی تصویر<sup>۵</sup> [۱۳] و رفع محوشدگی یا ماتی تصویر<sup>۶</sup> [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

<sup>1</sup> Sparse Representation

<sup>2</sup> Face Recognition

<sup>3</sup> Super Resolution Images

<sup>4</sup> Image Classification

<sup>5</sup> Image Restoration

<sup>6</sup> Image Deblurring

نمایش داده‌ها استفاده می‌شود. این ماتریس دیکشنری<sup>۶</sup> نامیده می‌شود و تمام فضای برداری مربوط به داده‌ها را پوشش می‌دهد. هرکدام از ستون‌های ماتریس دیکشنری را یک اتم<sup>۷</sup> می‌نامند. در صورتی که تعداد اتم‌های دیکشنری برابر بعد فضای برداری باشد آنگاه آن دیکشنری را یک دیکشنری کامل<sup>۸</sup> می‌نامند و در این صورت هرکدام از داده‌ها نمایشی یکتا با استفاده از اتم‌های دیکشنری خواهد داشت [۱۹]. اگر تعداد اتم‌های یک دیکشنری کامل را بیشتر کنیم، به دیکشنری فراکامل<sup>۹</sup> 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  $\alpha$  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]).

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

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

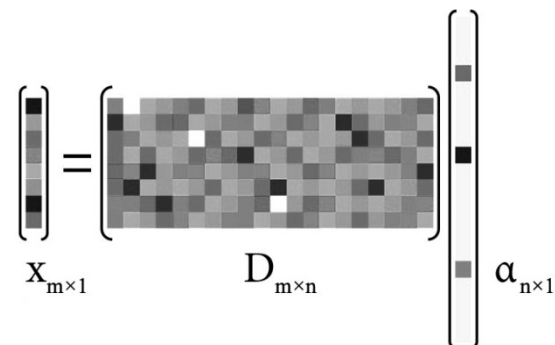
در این مسئله  $l_0$  به معنی تعداد عناصر غیر صفر بردار  $\alpha$  است و به‌طور کلی مسئله بالا به معنی یافتن بردار  $\alpha$  است که دارای کمترین مؤلفه غیر صفر باشد و یا به عبارت دیگر تنک‌ترین  $\alpha$  ممکن. اگر در معادله (۲) میزان تنک بودن را به صورت یک قید یا محدودیت به مسئله اضافه کنیم آنگاه مسئله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

$$\arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \alpha_0 \leq k. \quad (3)$$

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  $\alpha$

سیگنال‌ها هم استفاده کرد. حسگری فشرده<sup>۱</sup> 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 با در اختیار داشتن تعداد بسیار کمتری اندازه‌گیری خطی آن را بازسازی کرد. درواقع 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 می‌شود. در شکل شماره (۱) یک دیاگرام نوعی از نمونه‌برداری به وسیله حسگری فشرده نشان داده شده است. در این دیاگرام سیگنال  $x$  با ابعاد  $m \times 1$  به وسیله ماتریس  $D$  با ابعاد  $m \times n$  ( $m \ll n$ ) نمونه‌برداری می‌شود و سیگنال  $\alpha$  را با اندازه‌ای به مراتب کوچک‌تر نتیجه می‌دهد. سیگنال اندازه‌گیری شده  $\alpha$  را می‌توان به گیرنده ارسال کرد و اگر  $\alpha$  به میزان کافی تنک بوده و ماتریس  $D$  با ابعاد  $m \times n$  شرایط خاصی را ارضا کند، می‌توان  $x$  را از روی  $\alpha$  بازسازی کرد.

به الگوریتم‌های که برای یافتن  $\alpha$  با مشاهده  $x$  و  $D$  به کار می‌روند، الگوریتم‌های بازیابی تنک<sup>۲</sup> یا کدگذاری تنک<sup>۳</sup> گفته می‌شود که دارای انواع مختلفی مانند الگوریتم‌های حریصانه<sup>۴</sup> و الگوریتم‌های محدب<sup>۵</sup> هستند [۳]، [۱۷]، [۱۸].



شکل (۱): نمایش تنک سیگنال  $x$  به وسیله ماتریس  $D$

ماتریس  $D$  درواقع در بردارنده تمام پایه‌هایی است که برای

<sup>6</sup> Dictionary

<sup>7</sup> Atom

<sup>8</sup> Complete Dictionary

<sup>9</sup> Over-Complete Dictionary

<sup>1</sup> Compressed Sensing

<sup>2</sup> Sparse Recovery

<sup>3</sup> Sparse Coding

<sup>4</sup> Greedy Algorithms

<sup>5</sup> Convex Algorithms

حریصانه بهره می‌برند می‌توان به الگوریتم‌های  $OMP^1$ ،  $ORMP^2$  و  $WOMP^3$  اشاره کرد [۳]، [۲۴]، [۲۵]. همچنین الگوریتم‌های  $LASSO$  و گرادیان نزولی<sup>۴</sup> از تکنیک ملایم‌سازی محدب<sup>۵</sup> برای حل مسائل فوق استفاده می‌کنند [۲۰].

### ۳- یادگیری دیکشنری

جهت نمایش تنک سیگنال‌ها به یک دیکشنری نیاز است. ساخت دیکشنری به وسیله الگوریتم‌های یادگیری دیکشنری صورت می‌گیرد. الگوریتم‌های یادگیری دیکشنری در دهه اخیر مورد توجه پژوهشگران این حوزه بوده است [۲۶]. از الگوریتم‌های ساخت دیکشنری می‌توان به الگوریتم  $MOD$  اشاره کرد که دارای کارایی بالایی است [۲۷]. از دیگر روش‌های ساخت دیکشنری الگوریتم  $K-SVD$  است که دارای پیچیدگی کمتر و کارایی مشابه الگوریتم  $MOD$  است [۲۸]، [۲۹].

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  $D$  می‌توان به صورت رابطه (۹) نشان داد که در آن  $\hat{x}$  برداری است شامل ضرایبی که بیشتر آن‌ها صفر هستند و  $\Gamma$  خطای نمایش است.

$$\begin{matrix} x \\ \hat{x} \end{matrix} \quad (9)$$

هدف از یادگیری دیکشنری در واقع یافتن ماتریسی است که کمترین ضرایب غیر صفر را در بردار  $\alpha$  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  $X = \{x_l\}_{l=1}^L, x_l \in \mathbb{R}^m$  (شامل  $L$  بردار

به صورت تخمینی بیان شده و معادله اولیه (۱) به فرم معادله زیر درمی‌آید.

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

در این معادله  $s$  نویز نمایش نامیده می‌شود و  $s \in \mathbb{R}^m$  است و دارای قید  $\|s\|_2 \leq \varepsilon$  است. با حضور نویز پاسخ‌های تنک مسئله (۴) را می‌توان به صورت تقریبی با حل یکی از مسائل بهینه‌سازی زیر به دست آورد.

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

$$\hat{\alpha} = \arg \min_{\alpha} \|x - D\alpha\|_2 \quad s.t. \quad \alpha_0 \leq \varepsilon. \quad (6)$$

با استفاده از قضیه ضرایب لاگرانژ می‌توان مسئله را با استفاده از ثابت  $\lambda$  به فرم زیر بازنویسی کرد [۳]، [۲۰]:

$$\hat{\alpha} = L(\alpha, \lambda) = \arg \min_{\alpha} \|x - D\alpha\|_2 + \lambda \alpha_0. \quad (7)$$

هرچند روش حل مسائل بهینه‌سازی تعریف‌شده در بالا (با استفاده از نرم  $l_0$ ) می‌تواند از روش‌های بنیادین یافتن پاسخ تنک با استفاده از ماتریس  $D$  به دست آید ولی همچنان پاسخ مسئله غیرقطعی و به‌طور کلی مستلزم حل یک مسئله  $NP-HARD$  است [۲۱]. در [۲۲]، [۲۳] نشان داده شده است که پاسخ مسئله نرم  $l_0$  را می‌توان با استفاده از نرم  $l_1$  نیز به دست آورد که به مسئله بهینه‌سازی آن  $P_1$  گفته می‌شود. در بسیاری از مسائل کاربردی پاسخ مسائل  $P_0$  و  $P_1$  معادل هستند و مسئله  $P_1$  می‌تواند تقریب قابل قبولی برای یافتن نمایش تنک سیگنال باشد. اغلب ساختارهای مشهور نمایش تنک بر مبنای نرم  $l_1$  بر اساس مسائل بهینه‌سازی زیر حل می‌شوند:

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

$$\hat{\alpha} = \arg \min_{\alpha} \|x - D\alpha\|_2^2 \quad s.t. \quad \|\alpha\|_1 \leq \tau.$$

یکی دیگر از نرم‌هایی که توسط آن پاسخ تنک را به دست می‌آورند نرم  $l_2$  است. پاسخ‌های به‌دست‌آمده توسط این نرم دارای دقت زیادی نبوده و ممکن است که به‌اندازه کافی تنک نباشند.

از معروف‌ترین الگوریتم‌های کدگذاری تنک که از روش‌های

<sup>1</sup> Orthogonal Matching Pursuit

<sup>2</sup> Order Recursive Matching Pursuit

<sup>3</sup> Weak Orthogonal Matching Pursuit

<sup>4</sup> Gradient Descent

<sup>5</sup> Convex Relaxation

نام  $^1$  RLS-DLA معرفی شده است که روشی مؤثر و کارآمد در ساخت دیکشنری در شرایط برخط<sup>۲</sup>، 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

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<sup>۳</sup>، 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<sup>۴</sup> با اندازه  $8 \times 8$  یا  $16 \times 16$  یا  $32 \times 32$  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 (bi-orthogonal 4.4) یا CDF 5/3 (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 به خاطر بالا بردن کیفیت سیگنال بازسازی از الگوریتم ORMP استفاده شده

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 درواقع در مسئله بهینه‌سازی به شکل معادله (۱۰) صدق کند [۳].

$$a \quad (10)$$

در اینجا  $X$  ماتریس داده‌های آموزشی و  $A = \{\alpha_l\}_{l=1}^L, \alpha_l \in \mathbb{R}^n$  ماتریس ضرایب (وزن‌های) تنک متناظر با آن است که در آن  $\alpha_l$  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  $\gamma$  ضریب لاگرانژ است. لازم به ذکر است معادله (۱۰) با توجه به جفت  $(D, A)$  یک مسئله بهینه‌سازی سخت هست و به‌طورکلی در اکثر روش‌های یادگیری دیکشنری برای حل این مسئله با استفاده از یک فرآیند بازگشتی دو مرحله‌ای به حل آن می‌پردازند. این فرآیند با یک دیکشنری اولیه به‌صورت زیر آغاز می‌شود.

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

$$A$$

$$(11)$$

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

$$L$$

$$(12)$$

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

<sup>1</sup> Recursive Least Squares Dictionary Learning Algorithm

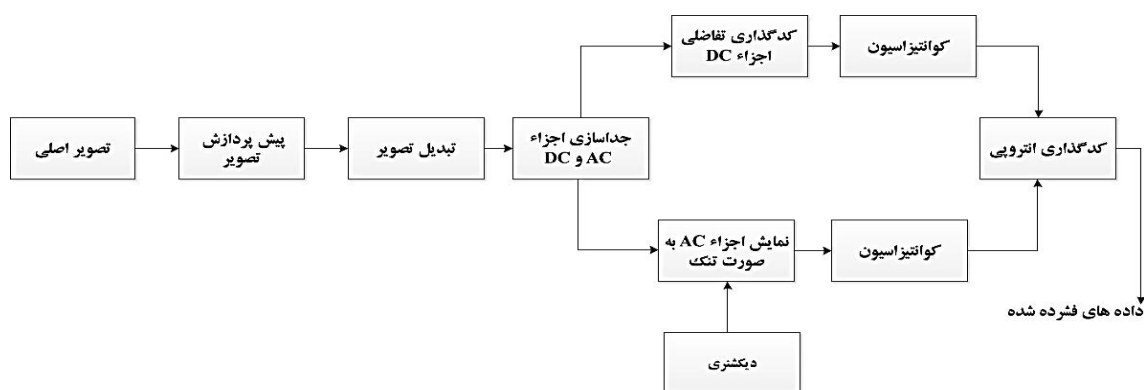
<sup>2</sup> Online

<sup>3</sup> Preprocessing

<sup>4</sup> Subimages

کوانتایزر *midtread* بیان شده در [۳۱]، [۳۲] استفاده کرد. سپس ضرایب کوانتیزه به صورت یک رشته درآمده و توسط یک کدگذاری انتروپی برای فشرده سازی کد می گردند که برای این منظور در ابتدا عناصر غیر صفر و همچنین اطلاعات موقعیت آنها در رشته به دست آمده مشخص و در رشته دیگر گذاشته می شوند. سپس هر دو رشته به دست آمده توسط یک کدگذاری انتروپی بهینه مانند هافمن کد شده و به صورت رشته های بیت درمی آیند که به برای کدبردار ارسال شده یا در یک فایل ذخیره می شوند. بدیهی است که در کدبردار دقیقاً عکس فرآیند مذکور انجام می گیرد تا تصویر اصلی بازایی شود.

است. در کدگذاری تنک از خطای نمایش به عنوان معیار توقف الگوریتم استفاده می شود. بدین صورت که با استفاده از یک PSNR هدف داده شده، پارامتر  $\epsilon$  یا معیار توقف تعیین می شود. پس از اعمال کدگذاری تنک به بردارهای ضرایب AC بردارها به صورت بردارهایی با کمترین تعداد عناصر غیر صفر درمی آیند. روشن است که هر چه بردارها تنک تر باشند عمل فشرده سازی تصویر به گونه ای بهتر انجام خواهد شد. پس از این که ضرایب AC و DC به طور جداگانه کد شدند، کدهای به دست آمده برای هر کدام به صورت جداگانه کوانتیزه می شوند. در ساده ترین حالت کوانتیزاسیون مقادیر ضرایب به نزدیک ترین مقدار صحیح گرد می شوند. همچنین می توان از



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

وابستگی به دیکشنری اولیه با استفاده از یک ضریب فراموشی<sup>۱</sup>  $\lambda$  است. در این الگوریتم از یک گام زمانی<sup>۲</sup>  $i$  برای معرفی ماتریس داده های آموزشی استفاده می شود که به صورت  $X_i = [x_1, x_2, \dots, x_i]$  با ابعاد  $m \times i$  تعریف می شود. هر گام زمانی  $i$  در واقع بیان کننده یک مرحله است. در هر حلقه تکرار یکی از بردارهای آموزشی به ماتریس افزوده می شود از این رو هر زمان که بخواهیم یک بردار آموزشی جدید به ماتریس داده ها اضافه کنیم کافی است که الگوریتم را با یک مرحله یا گام اضافی اجرا کنیم. به این دلیل این روش را یک

## ۵- روش پیشنهادی

در شکل (۳) سیستم پیشنهادی برای فشرده سازی 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 می شود. در بخش آموزش، دیکشنری متناسب با تصاویر چهره ای که باید فرآیند فشرده سازی بر روی آنها صورت گیرد ساخته می شود.

الگوریتمی که برای آموزش دیکشنری به کار برده ایم الگوریتم RLS-DLA است که یکی از مزیت های مهم آن کاهش

<sup>1</sup> Forgetting Factor

<sup>2</sup> Time Step

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  $8 \times 8$  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 زیرتصویرهایی با ابعاد  $8 \times 8$  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 وقتی از چند دیکشنری که هر

روش یادگیری دیکشنری برخط می‌نامند [۳۳]. در این الگوریتم ماتریس ضرایب به‌صورت  $A_i = [\alpha_1, \alpha_2, \dots, \alpha_i]$  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  $\lambda$  به‌صورت 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  $\lambda$  در صورت ثابت در نظر گرفته شدن مقداری نزدیک به 1 مثلاً حدود  $\lambda \approx 1 - \frac{1}{L}$  در نظر گرفته‌شده و برای  $\lambda_i$  تطبیقی مقدار آن به به‌صورت بازگشتی تغییر می‌کند.

الگوریتم RLS-DLA در کاربردهایی که در آن‌ها از داده‌های آموزشی بزرگ استفاده می‌شود، روش بسیار مناسبی است. برای مقادردهی 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  $\epsilon$  برای تقریب خطا استفاده می‌شود که بر اساس آن می‌توان یک PSNR هدف<sup>۱</sup> را در آغاز الگوریتم یادگیری دیکشنری محاسبه کرد. در این الگوریتم به‌کل بازگشت‌هایی که به تعداد بردارهای آموزش داریم یک آزمایش<sup>۲</sup> 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

<sup>1</sup> Target PSNR

<sup>2</sup> Trial

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 ارائه شده فشرده‌سازی تصاویر چهره

به منظور پیاده‌سازی و شبیه‌سازی روش ارائه شده از نرم‌افزار MATLAB 2015a و یک کامپیوتر شخصی با توان پردازشی 2 GHz و مقدار حافظه رم 8 GB و سیستم عامل Windows 7 استفاده شده است. به منظور 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.<sup>۱</sup> 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<sup>۲</sup> انتخاب شده‌اند، بکار رفته‌اند. هرکدام 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

دیکشنری مطابق با کیفیت فشرده‌سازی تصویر چهره مورد نظر 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. در واقع برای فشرده‌سازی هر تصویر با کیفیت موردنظر با استفاده از یک الگوریتم تصمیم‌گیری یک دیکشنری از میان چند دیکشنری ساخته شده 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 ساخته شده با الگوریتم یادگیری دیکشنری RLS-DLA به طور تجربی دریافتیم که PSNR تصاویر فشرده شده، متناسب با Target PSNR دیکشنری ساخته شده در برخی بازه‌های کیفیت موردنظر در گیرنده (کدبردار) عملکرد نرخ-عوجاج 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 ساخته شده به صورت زیر تعیین گردید:

$$\begin{aligned} L \\ < (18) \\ \leq \end{aligned}$$

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 فشرده شده یا مقادیر نوعی PSNR های تصویر در فشرده‌سازی با اتلاف تصاویر چهره برای نرخ‌های فشرده‌سازی 0.10 تا 2.00 بیت بر پیکسل، بین 35 تا 55 دسی‌بل است. از این رو 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 (۴ 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 یک بار

<sup>۱</sup> 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>

<sup>۲</sup> تصاویر چهره پایگاه داده FEI از طریق آدرس زیر موجود و قابل دسترسی می‌باشند:

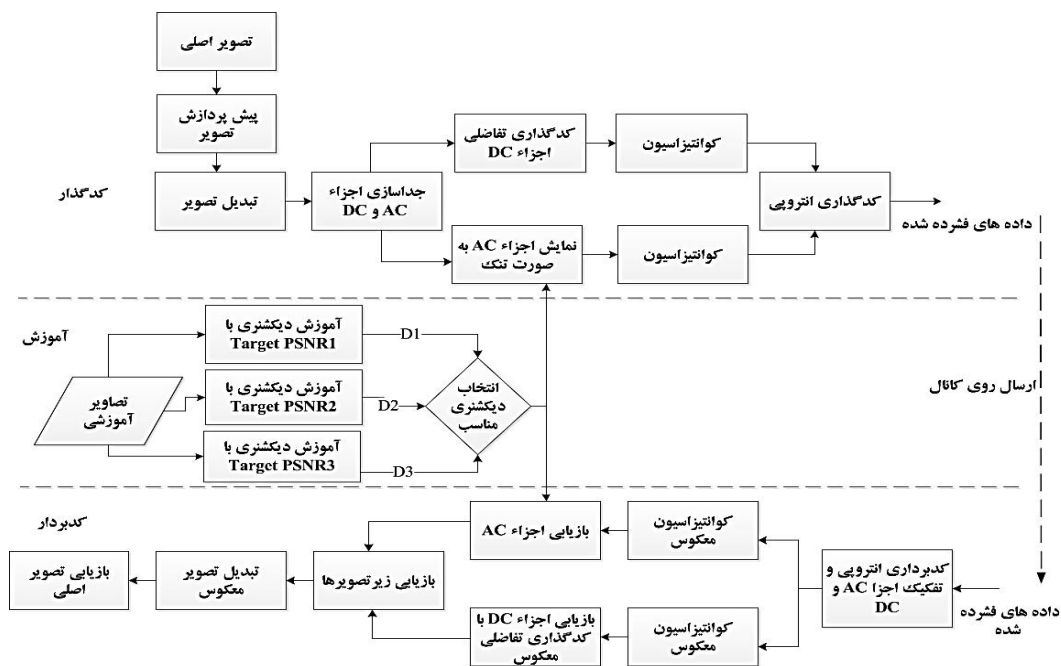
<http://fei.edu.br/~cet/facedatabase.html>



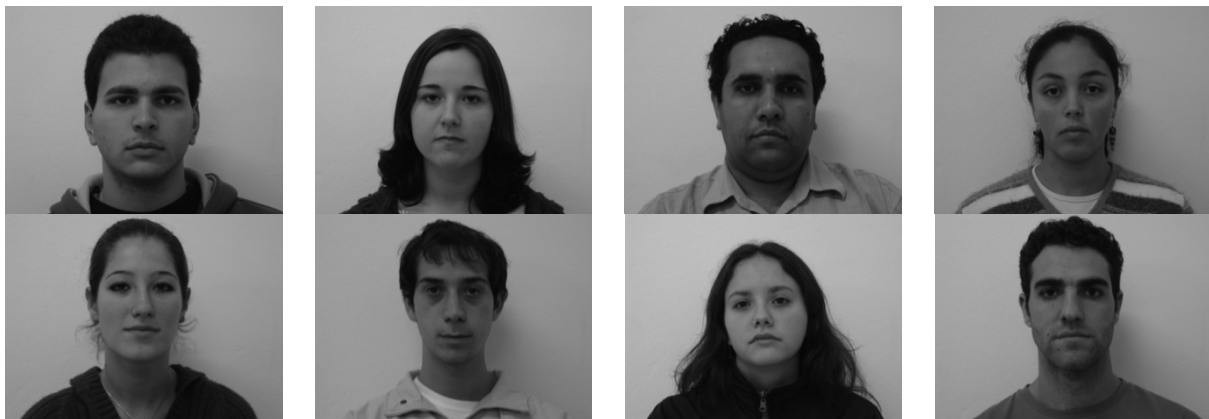
ارائه یک روش جدید فشرده سازی تصاویر با استفاده از نمایش ...

135.7 دقیقه است. دیکشنری‌های ساخته شده شامل 440 اتم با ابعاد  $8 \times 8$  پیکسل می‌باشند (با ستون‌های به طول 64 درایه). هر دیکشنری از هر تصویر آموزشی 3000 تکه تصویر با ابعاد  $8 \times 8$  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.

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 45 و 50 و 55 دسی‌بل ساخته شدند و برای ساخت هر دیکشنری از 1000 حلقه تکرار به‌منظور دستیابی به کیفیت موردنظر (Target PSNR) استفاده شده است. زمانی که برای آموزش هر دیکشنری طی شده به ترتیب برابر 315.6 و 399.1 و 440.7



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



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

کیفیت تصاویر بازیابی شده با استفاده از الگوریتم پیشنهادی از لحاظ بصری بهتر از بقیه الگوریتم‌ها است.

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 مطرح‌شده فشرده شدند. نتایج به‌دست‌آمده حاکی از آن است که روش فشرده‌سازی ارائه‌شده 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) به‌صورت 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 به‌طوری‌که نتایج به‌دست‌آمده تنها با ۸ تصویر آموزشی حاصل شده است. با به‌کارگیری روش‌های فشرده‌سازی انترپوی 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.

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.

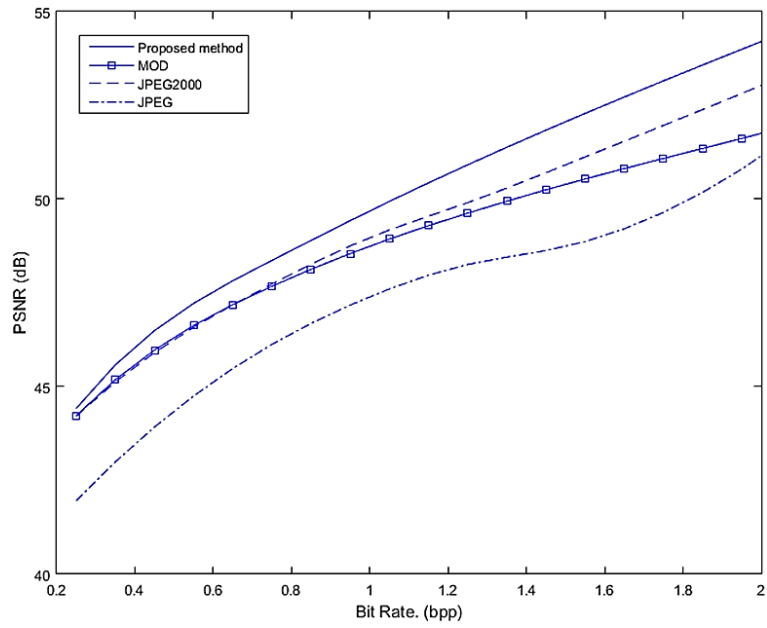
پس از ساخته‌شدن دیکشنری‌های موردنظر، از دو تصویر چهره متفاوت از تصاویر آموزشی برای آزمایش الگوریتم فشرده‌سازی استفاده شد و فرآیند فشرده‌سازی و بازیابی تصویر بر روی آن‌ها با نرخ بیت‌های متفاوت صورت گرفته است و پس‌از آن کیفیت تصاویر بازیابی شده اندازه‌گیری شده‌اند. در شکل (۵) تصاویر آزمایشی مذکور ارائه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



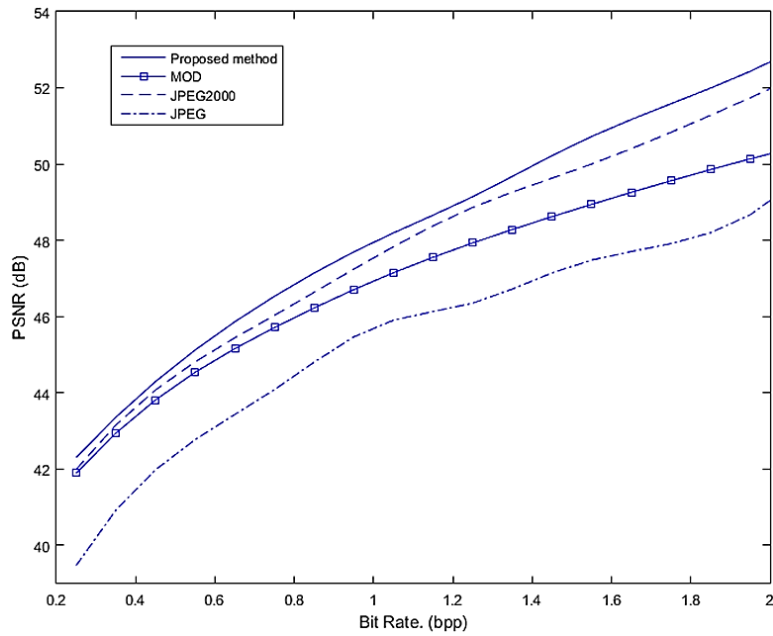
شکل ۵: دو تصویر مورد آزمایش در فشرده‌سازی با الگوریتم‌های مختلف، راست: تصویر شماره ۱، چپ: تصویر شماره ۲.

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

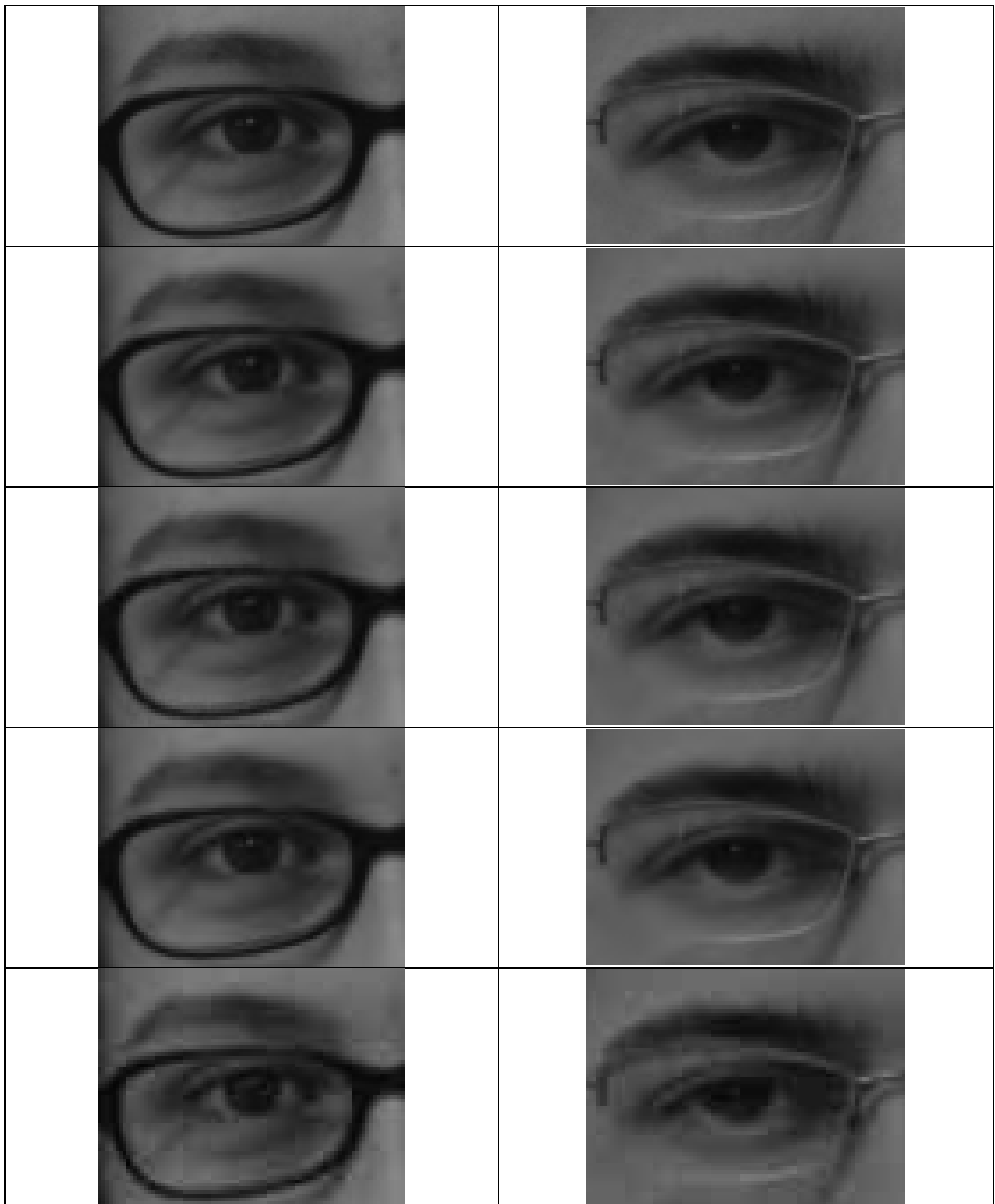
ارائه یک روش جدید فشرده سازی تصاویر با استفاده از نمایش ...



شکل ۶- مقایسه فشرده سازی تصویر آزمایشی شماره ۱ توسط روش های مختلف



شکل ۷: مقایسه فشرده سازی تصویر آزمایشی شماره ۲ توسط روش های مختلف



شکل ۸: کیفیت دو تصویر آزمایشی فشرده شده در ناحیه چشمها (راست: تصویر آزمایشی شماره ۱، چپ: تصویر آزمایشی شماره ۲) با نرخ  $0.25 \text{ bpp}$  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 فشرده سازی (PSNR) تصویر آزمایشی ۱ با استفاده از الگوریتم های مختلف برحسب dB با نرخ بیت های مختلف

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

جدول (۲): نتایج کیفیت فشرده سازی (PSNR) تصویر آزمایشی ۲ با استفاده از الگوریتم های مختلف برحسب dB با نرخ بیت های مختلف

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

#### ۷- نتیجه گیری

عملاً بالاترین کیفیت را در بین الگوریتم های فشرده سازی ارائه کرده و به عنوان یک استاندارد در فشرده سازی تصویر شناخته شده است، در حدود 0.5 dB تا 1.2 dB برتر است. در روش ارائه شده استفاده از کدگذاری انتروپی پیشرفته تری همچون کدگذاری حسابی یا هافمن تطبیقی می تواند مورد بررسی قرار گرفته و کیفیت و نرخ فشرده سازی را ارتقا دهد. همچنین می توان تأثیر دیکشنری های مختلف در فشرده سازی تصاویر با ابعاد متغیر را مورد بررسی قرار داد. از الگوریتم مطرح شده می توان در فشرده سازی تصاویر پایگاه های داده بزرگ با چند میلیون تصویر مانند داده های سازمان ها، دانشگاه ها و بیمارستان ها بهره جست و به میزان قابل توجهی در فضای ذخیره سازی صرفه جویی نمود که این امر نیز به نوبه خود باعث کاهش پهنای باند اشغالی جهت تبادل داده در شبکه های کامپیوتری و مخابرات تصویری می شود.

در این تحقیق با بررسی روش ها و الگوریتم های موجود برای فشرده سازی تصویر با استفاده از حوزه تنک روش جدیدی برای فشرده سازی تصاویر چهره پیشنهاد گردید که باعث بهبود کیفیت تصاویر فشرده شده نسبت به دیگر الگوریتم های فشرده سازی در نرخ بیت یکسان می شود. در این الگوریتم با بررسی و تجزیه و تحلیل نتایج متوجه شدیم که کیفیت تصاویر فشرده شده با کیفیت دیکشنری (مقدار خطای دیکشنری) ساخته شده مرتبط است و همین امر را اساس ساخت الگوریتم پیشنهادی قرار دادیم. بدین صورت که با استفاده از الگوریتم آموزش RLS سه دیکشنری با کیفیت های متفاوت ساخته شدند و هنگام فشرده سازی دیکشنری منطبق با کیفیت تصویر مورد نظر در فشرده سازی انتخاب گردید. نتایج شبیه سازی حاصل از الگوریتم پیشنهادی نشان دهنده بهبود کیفیت تصاویر فشرده شده و برتری آن نسبت به دیگر روش های فشرده سازی است، به طوری که نسبت به الگوریتم پیشرفته JPEG2000 که

## مراجع

- [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
- [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.

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.
- 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.*,