

## فشرده‌سازی تصاویر دارای درجه‌ی تفکیک مکانی بالا مبتنی بر تخمین زیرباندهای جزئیات در حوزه‌ی تبدیل موجک

هادی گرایلو\*

### چکیده

درجه‌ی تفکیک مکانی مناسب در بسیاری از انواع تصاویر اهمیت بالایی دارد زیرا دربرگیرنده‌ی اطلاعات برخی جزئیات مهم می‌باشد. در عمده‌ی انواع تصاویر مانند تصاویر متنی، چهره، اثرانگشت و پزشکی کارایی روش‌های استخراج ویژگی تا حد زیادی به کیفیت تصویر وابسته است. درجه‌ی تفکیک مکانی بالا یکی از مهمترین عوامل افزایش کیفیت تصویر است اما موجب افزایش حجم حافظه‌ی ذخیره‌سازی نیز می‌شود؛ لذا اهمیت روش‌های موثر فشرده‌سازی دوچندان می‌شود.

در روش فشرده‌سازی پیشنهادی در این مقاله، ابتدا ابعاد تصویر ورودی به کمک تبدیل موجک تا حد مشخصی کاهش یافته و سپس تصویر به کمک هر روش کدگذار دلخواهی قابل فشرده‌سازی است. در مرحله‌ی بازسازی، ابتدا تصویر کاهش یافته به کمک کدگشای متناظر، بازسازی شده و سپس ابعاد آن به کمک تخمین زیرباندهای جزئیات در حوزه‌ی تبدیل موجک، افزایش می‌یابد.

در بخش شبیه‌سازی و ارزیابی، دو نوع تصویر متداول و مهم شامل تصویر متنی و تصویر چهره (نماینده‌ی تصاویر دارای طیف به ترتیب عمدتاً میان‌گذر و عمدتاً پایین‌گذر) به عنوان مطالعه موردی انتخاب و در هرکدام، کارایی فشرده‌سازی و بازشناسی تصاویر فشرده شده به کمک ترکیب روش پیشنهادی با سه روش JPEG، JPEG2000 و SPIHT بررسی می‌شود. نتایج شبیه‌سازی نشان‌دهنده‌ی تاثیر قابل توجه روش پیشنهادی در کاهش حجم ذخیره‌سازی تصویر فشرده‌شده و حفظ همزمان کارایی بازشناسی است.

### کلمات کلیدی

فشرده‌سازی تصویر؛ کاهش/افزایش ابعاد؛ تخمین زیرباندهای جزئیات؛ تصاویر متنی؛ تصاویر چهره؛ کارایی بازشناسی

### ۱ مقدمه

با تولید روزافزون تصاویر مختلف در کاربردهای مختلف (مانند تولید تصاویر متنی در کتابخانه‌های دیجیتال و تولید تصاویر چهره در موسسات مالی و سازمان‌های بایگانی و امنیتی) از یک طرف و قیمت بالا و حجم محدود حافظه‌های دیجیتال از سوی دیگر، اهمیت روش‌های جدید فشرده‌سازی تصاویر نیز رو به افزایش است.

یکی از مهمترین کاربردهای مبتنی بر این تصاویر، استخراج اطلاعات به منظور استفاده در کاربرد مورد نظر است؛ در کتابخانه‌های دیجیتال و سازمان ثبت اسناد، کتاب‌ها و اسناد کاغذی به قالب دیجیتالی تبدیل می‌شوند تا علاوه بر صرفه‌جویی در مکان فیزیکی لازم جهت نگهداری، از الگوریتم‌های پردازش تصویر نظیر بازشناسی در کاربردهای مربوطه بهره برده شود. در سازمان‌های امنیتی، مانند پلیس فرودگاه و صدور روادید، سازمان‌های اجتماعی، نظیر سازمان ثبت احوال، و موسسات مالی-اعتباری نظیر بانک‌ها نیز تصاویر پرسنلی و اثرانگشت افراد همه روزه به شکل دیجیتالی در حال تولید و یا پردازش است. دیگر انواع تصاویر نیز، مانند تصاویر پزشکی و تصاویر ماهواره‌ای، کم و بیش مشمول این نیازمندی (ذخیره‌ی دیجیتالی و

این مقاله در مردادماه سال ۱۳۹۵ دریافت، در بهمن‌ماه بازنگری و در تیرماه سال ۱۳۹۶ پذیرفته شد.

\*گروه الکترونیک، دانشکده برق و ریاضیات، دانشگاه صنعتی شاهرود.

هستند (مانند تصاویر متنی و تصاویر اثرانگشت) کمتر از تصاویر دارای طیف عمدتاً پایین‌گذر خواهد بود.

مهمترین ایده‌های مورد استفاده در روش پیشنهادی عبارتند از: (۱) کاهش ابعاد تصویر ورودی به کمک تبدیل موجک در واحد فشرده‌ساز، (۲) افزایش ابعاد تصویر فشرده‌شده در واحد بازسازی به کمک تخمین زیرباندهای جزئیات؛ اساس این ایده به تبدیل موجک معکوس و قابلیت آن در بازسازی یک تصویر با ابعاد بزرگ‌تر از روی زیرباندهای تقریب و جزئیات (افقی، عمودی، و قطری) مرحله‌ی قبل خود مربوط می‌شود. در اینجا تصویر بازسازی شده‌ی اولیه در واحد بازسازی، به عنوان زیرباند تقریب در نظر گرفته شده و سه زیرباند مورد نیاز جهت عملکرد تبدیل موجک معکوس از روی همین تصویر تقریب تخمین زده می‌شوند. واحد تخمین‌گر قبلاً به صورت برون‌خط آموزش دیده شده است تا بتواند از روی یک تصویر تقریب، سه زیرباند جزئیات مورد نیاز را تخمین بزند.

در مرحله‌ی بررسی و ارزیابی کارایی فشرده‌سازی روش پیشنهادی، با توجه به اهمیت ماهیت طیفی تصاویر ورودی، دو نوع تصویر متنی و چهره به عنوان نمایندگانی از تصاویر دارای طیف به ترتیب عمدتاً میان‌گذر و عمدتاً پایین‌گذر انتخاب و مورد استفاده قرار گرفته‌اند. در این ارزیابی، علاوه بر کارایی فشرده‌سازی، کارایی بازشناسی نیز مورد توجه قرار گرفته است. کارایی فشرده‌سازی بر طبق معیارهایی مانند منحنی متوسط PSNR برحسب نرخ بیت و کارایی بازشناسی نیز بر طبق معیارهایی مانند منحنی متوسط صحت بازشناسی برحسب نرخ بیت ارزیابی شده‌اند.

ادامه مقاله بدین صورت است که در بخش ۲ مهم‌ترین روش‌های فشرده‌سازی تصاویر متنی و چهره، در بخش‌های جداگانه، مرور خواهند شد. بخش ۳ به تشریح جزئیات روش پیشنهادی پرداخته و بخش ۴ نتایج بررسی و ارزیابی روش پیشنهادی و مقایسه با چند روش مرجع را ارائه خواهد نمود. در بخش ۵ به نتیجه‌گیری پرداخته شده است.

## ۲ مروری بر روش‌های موجود

با توجه به انتخاب دو نوع تصویر متنی و چهره جهت استفاده در بخش ارزیابی روش پیشنهادی، در این بخش مهم‌ترین روش‌های فشرده‌سازی این دو نوع تصویر را بررسی خواهیم کرد. مهم‌ترین نکاتی که از بررسی این روش‌ها می‌توان دریافت عبارتند از: (۱) کیفیت تصویر تاثیر زیادی در کارایی بازشناسی و نیز خوانایی/خوشایندی آن دارد، (۲) کارایی بازشناسی با کاهش نرخ بیت، در حالت کلی، در ابتدا تاحدی ثابت بوده و سپس با شیب نسبتاً زیادی شروع به اُفت می‌کند، (۳) تاکنون در روش‌های فشرده‌سازی موجود به ویژه جهت تصاویر چهره از درجه‌ی تفکیک مکانی پایین (طوری که فاصله‌ی بین دو چشم در حدود ۹۰ پیکسل بوده است؛ بنابراین ابعاد تصاویر در حدود ۱۸۰ الی

پردازش جهت استخراج اطلاعات) می‌باشند. در تمام یا لاقط قسمت عمده‌ای از این تصاویر و کاربردها، استخراج اطلاعات رابطه‌ی تنگاتنگی با کیفیت تصویر دارد؛ به ویژه، کارایی روش‌های بازشناسی (متن، چهره، اثرانگشت، ماموگرافی و غیره) به طور مستقیم به کیفیت تصاویر بستگی دارد. درجه‌ی تفکیک مکانی بالا از جمله عوامل تاثیرگذار بر کیفیت تصویر است؛ برای مثال در تصاویر متنی موجب افزایش کیفیت دیداری و خوانایی ماشینی/انسانی می‌شود، در تصاویر ماموگرافی (سینه‌نگاری) موجب تسهیل تشخیص غدد سرطانی و ذرات رسوبی کلسیمی می‌شود، در تصاویر چهره و اثرانگشت موجب افزایش کارایی بازشناسی می‌شود و بالاخره، در تصاویر طبیعی موجب تسهیل تشخیص عوارض طبیعی از غیرطبیعی می‌شود.

افزایش درجه‌ی تفکیک مکانی از سوی دیگر موجب افزایش حجم حافظه‌ی دیجیتال مورد نیاز جهت ذخیره‌سازی این تصاویر نیز می‌شود؛ این امر، با توجه به حجم محدود و قیمت بالای حافظه‌های موجود، در کاربردهای نامبرده شده، یک مشکل جدی محسوب می‌شود. بنابراین، استفاده از روش‌های موثر فشرده‌سازی بسیار ضروری است. این روش‌ها باید همزمان با کاهش موثر حجم حافظه‌ی ذخیره‌سازی تصاویر فشرده‌شده، کیفیت تصاویر بازسازی شده را در حدی مطلوب حفظ نمایند طوری که کارایی روش‌های استخراج و پردازش اطلاعات اُفت نکند.

روش پیشنهادی در این مقاله به منظور فشرده‌سازی انواع تصاویر دارای درجه‌ی تفکیک مکانی بالا ارائه شده است. مهم‌ترین محاسن روش پیشنهادی عبارتند از: (۱) برای انواع تصاویر قابل استفاده است، (۲) قابل استفاده با هر روش یا استاندارد فشرده‌سازی موجود می‌باشد؛ بنابراین همچنان می‌توان از استانداردهای متداول امروزی مانند JPEG و JPEG2000 استفاده نمود. اهمیت این مطلب این است که کاربر بنا به اهداف و خواسته‌های خود می‌تواند روش فشرده‌سازی مناسبی را انتخاب و استفاده نماید؛ برای مثال اگر به همه منظوره و رایج بودن روش فشرده‌سازی توجه می‌کند می‌تواند از استانداردهایی نظیر JPEG و JPEG2000 استفاده کند و اگر به پارامترهایی نظیر قابلیت فشرده‌سازی پیش‌رونده و کنترل دقیق بودجه‌ی بی‌تی توجه دارد از کدگذارهای بخش‌بندی مجموعه [۱-۲] نظیر SPIHT [۳] می‌تواند استفاده نماید، (۳) کارایی فشرده‌سازی آن از دو دیدگاهی که قبلاً اشاره شده قابل قبول است؛ یعنی علاوه بر کاهش حجم ذخیره‌سازی، اطلاعات مهم و موثر در پردازش اطلاعات، نظیر بازشناسی، را تا حد خوبی حفظ می‌کند، (۴) کارایی فشرده‌سازی آن تا حد زیادی به ماهیت طیفی تصویر ورودی بستگی دارد؛ یعنی هر قدر طیف تصویر ورودی به سمت پایین‌گذر متمایل باشد (مانند تصاویر طبیعی و تصاویر چهره) میزان فشرده‌سازی آن بیشتر خواهد بود. بنابراین در حالت کلی، میزان فشرده‌سازی برای تصاویری که دارای طیف عمدتاً میان‌گذر

H.264/AVC برای فشرده‌سازی تصاویر متنی استفاده شده است.

روشهای مبتنی بر مدل جزو روشهای تخصصی فشرده‌سازی تصاویر متنی محسوب شده و در آنها از مدل محتوای ترکیبی (MRC) استفاده می‌شود [۱۱]. در این مدل، لایه‌های مختلفی برای تصویر متنی تعریف و از یکدیگر جدا شده و هرکدام از این لایه‌ها، جداگانه کدگذاری می‌شوند. در بیشتر روشهای مبتنی بر این مدل، سه لایه‌ی پس‌زمینه، پیش‌زمینه، و ماسک (پوشش) تعریف می‌شوند. لایه‌ی پیش‌زمینه شامل رنگ یا شدت روشنایی نواحی متنی است. لایه‌ی ماسک یک تصویر دودویی بوده و نواحی متنی را از نواحی غیرمتنی جدا می‌کند. لایه‌ی پس‌زمینه نیز شامل اطلاعات نواحی غیرمتنی مانند گرافیک، شکل، و رنگ پس‌زمینه‌ی تصویر است.

عمده‌ی روش‌های مبتنی بر این مدل روی تصاویر متنی چاپی کار کرده و تعداد بسیار کمی تصاویر دستنویس را در نظر گرفته‌اند. با استفاده از این مدل می‌توان به میزان فشرده‌سازی بالا و در عین حال کیفیت مناسبی برای تصاویر فشرده شده دست یافت [۱۴-۱۲]. این مدل جزو استاندارد ITU (به نام T.44) درآمده و حتی در استاندارد JPEG2000 در بخش تصاویر متنی مرکب<sup>۲</sup> نیز استفاده شده است. این مدل دارای چند ویژگی مطلوب می‌باشد: اول، کدگذارهای مورد استفاده در این مدل مستقل از هم هستند. دوم، امکان استفاده از استانداردهای نسبتاً عمومی موجود فراهم می‌شود. سوم، بین نواحی متنی و غیرمتنی، تمایز در نظر گرفته شود؛ چهارم، می‌توان هر لایه را، بسته به میزان اهمیت آن، با نرخ بیت متفاوت و مشخصی کدگذاری نمود؛ پنجمین حسن این مدل، برخورداری از سادگی و قانون‌مندی است. و ششم، برای کدگذاری هر لایه می‌توان با استفاده از فرضیات معقول و مختص به همان لایه، به کارایی فشرده‌سازی بالاتری در مقایسه با روشهای عمومی فشرده‌سازی تصاویر دست یافت. برای مثال، برخی روشها برای کدگذاری لایه‌ی پس‌زمینه از رنگ نماینده برای کدگذاری قسمت قابل توجهی از لایه‌ی پس‌زمینه استفاده می‌کنند [۱۵]. در مورد لایه‌ی پیش‌زمینه نیز تا حدی به طور مشابه می‌توان عمل کرد. در مورد لایه‌ی ماسک نیز از ایده‌هایی مانند انطباق الگو [۱۶] و استفاده از ویژگی‌های رسم‌الخط [۱۷-۱۸] می‌توان استفاده کرد. در [۱۹] روشی مبتنی بر مدل محتوای ترکیبی ارائه شده که تصویر ورودی را به چهار لایه تجزیه کرده و هر لایه را به شیوه‌ی جداگانه‌ای کدگذاری می‌کند. تشکیل لایه‌ها وابسته به نوع کلمات در زبان عربی و جامد و مرکب بودن آنها است.

روش‌های زیادی مبتنی بر این مدل تاکنون ارائه شده‌اند که از بین آنها برخی مانند [۲۰] DjVu، [۲۱] DigiPaper و [۲۲-۲۳] LuraDocument تجاری‌سازی شده‌اند. از بین این روش‌ها می‌توان DjVu را یکی از بهترین و کارآمدترین روش‌های

۲۰۰ پیکسل بوده است) و جهت تصاویر متنی از درجه‌ی تفکیک مکانی متوسط (در حدود ۳۰۰ dpi) استفاده شده و تصاویر با درجه‌ی تفکیک مکانی بالا مورد توجه قرار نگرفته است.

## ۲-۱ فشرده‌سازی تصاویر متنی

تصاویر متنی از چند جنبه با دیگر انواع تصاویر به ویژه تصاویر طبیعی که متداول‌ترین نوع از تصاویر هستند، متفاوت می‌باشند. اول این که در تصاویر متنی، برخلاف تصاویر طبیعی، درجه‌ی تفکیک مکانی بیشتر از درجه‌ی تفکیک مقادیر شدت روشنایی در حفظ کیفیت یا اطلاعات اولیه اهمیت دارد [۴-۵]. دوم این که تصاویر متنی، برخلاف تصاویر طبیعی، شامل جزئیات دیداری مهم مانند شکل و انحنا، رنگ کاغذ، و بافت کاغذ هستند. سومین تفاوت این است که در تصاویر متنی عمده‌ی اطلاعات در لبه‌ها قرار داشته و نواحی داخلی نویسه‌ها و علائم متنی حاوی اطلاعات چندانی نمی‌باشند. در برخی انواع تصاویر مانند تصاویر طبیعی و تصاویر چهره، قسمت مهمی از اطلاعات در بدنه‌ی اشیاء موجود در تصویر قرار دارد. به همین دلیل می‌توان گفت طیف تصاویر متنی، طیفی عمدتاً میان‌گذراست اما طیف تصاویر طبیعی و چهره، عمدتاً پایین‌گذر است [۶]. 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 فایلی با نام JPM را به خود اختصاص داده است [۷].

روش‌های فشرده‌سازی 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<sup>۱</sup>) استفاده نموده و تصویر متنی را به لایه‌های مختلفی تجزیه می‌کنند. روش‌های غیرمبتنی بر مدل، عمدتاً به دنبال ارتقاء کارایی فشرده‌سازی یکی از روش‌ها یا استانداردهای نسبتاً عمومی موجود به منظور استفاده در کاربرد فشرده‌سازی تصاویر متنی می‌باشند.

روش‌های غیرمبتنی بر مدل ممکن است استانداردهای موجود را برای فشرده‌سازی تصاویر متنی به کار گرفته و عملکرد آنها را بهبود داده باشند. برای مثال، در [۸] 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 [۹] 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 [۱۰] از کدک ویدیویی

<sup>2</sup>Compound Documents

<sup>1</sup> Mixed Raster Content (MRC)

پیشنهادی پس از آشکارسازی چهره در تصویر ورودی و نرمالیزه کردن مقیاس و ابعاد هندسی آن، از یک مدل پارامتری فشرده جهت نمایش تصویر و استفاده در کاربردهایی نظیر بازشناسی و فشرده‌سازی استفاده می‌کند. این مدل مبتنی بر تبدیل KL<sup>۱</sup> می‌باشد. روش پیشنهادی در [۳۲] مبتنی بر کاربرد انتقال توسط مویایل بوده و جزو روشهای مبتنی بر مدل محسوب می‌شود زیرا به جای انتقال تصویر، برخی ویژگیها را از تصویر چهره استخراج و سپس منتقل می‌کند. این ویژگیها شامل دره‌ها<sup>۲</sup> و نقاط منفرد<sup>۳</sup> سطح تصویر چهره هستند. برای نمایش و مدل کردن تصویر نیز از مجموعه‌های سطح<sup>۴</sup> استفاده شده است.

بقیه‌ی روش‌های فشرده‌سازی تصاویر چهره از نوع روش‌های غیرمبتنی بر مدل بوده و سعی در استفاده‌ی موثر از یک روش استاندارد یا غیراستاندارد در جهت فشرده‌سازی تصاویر چهره دارند. روش‌هایی که از استانداردهای موجود در جهت فشرده‌سازی تصاویر چهره بهره گرفته‌اند، عمدتاً از استاندارد JPEG2000 بهره برده‌اند [۳۷ و ۳۵] زیرا این استاندارد حالت‌های کاری یا بخش‌های مختلفی را تعریف کرده است. یکی از این حالت‌ها امکان استفاده از کدگذاری مبتنی بر نواحی مطلوب (ROI)<sup>۵</sup> است. در این نوع کدگذاری، نواحی مختلفی بسته به میزان اهمیت آنها در تصویر تعریف و هر ناحیه با نرخ بیت مختلفی فشرده‌سازی می‌شود. تفاوت روش‌هایی که از این قابلیت استفاده کرده‌اند در پارامترهایی مانند تعریف نواحی مطلوب و چگونگی شناسایی این نواحی است. در [۳۳] روشی جهت فشرده‌سازی تصاویر چهره موجود در روادید<sup>۶</sup> مبتنی بر کدگذاری نواحی مطلوب ارائه شده است. ابتدا تصویر با استفاده از روشی مبتنی بر تفاضل پیکسلی به دو ناحیه پس‌زمینه و پیش‌زمینه قطعه‌بندی می‌شود. سپس برای فشرده‌سازی نواحی پیش‌زمینه (شامل چهره) ابتدا از یک الگوریتم قطعه‌بندی چهره جهت آشکارسازی چشم و دهان و سپس از کدگذاری ROI جهت فشرده‌سازی این نواحی با کیفیت بیشتر استفاده می‌شود. در [۳۴] برای تعیین نواحی مطلوب، از برخی روش‌های آشکارسازی چهره، یافتن چشمها، دهان و گونه استفاده شده است. در این مرجع نیز از قابلیت کدگذاری مبتنی بر نواحی مطلوب (ROI) موجود در JPEG2000 جهت فشرده‌سازی تصاویر چهره مربوط به روادید استفاده شده است.

دومین دسته از روش‌های غیرمبتنی بر مدل از استانداردهای موجود استفاده‌ای نکرده‌اند اما ممکن است از برخی فنون مورد استفاده در دسته‌ی قبل، مانند کدگذاری ROI بهره برده باشند.

مبتنی بر مدل MRC دانست زیرا بسیاری از روشها کارایی خود را با آن مقایسه کرده‌اند [۲۷-۲۴، ۱۲].

اما با وجود محاسن نام برده شده، کارایی روشهای مبتنی بر MRC از چند جهت چندان مناسب نمی‌باشد: اول، لبه‌ها معمولاً متناظر با گذر نسبتاً آرام هستند نه ناگهانی، بنابراین مدل فوق نمی‌تواند کیفیت این گذر را حفظ کند. دوم، کارایی این مدل تا حد زیادی وابسته به روش مورد استفاده جهت جداسازی لایه‌ها یا همان قطعه‌بندی تصویر متنی دارد. سوم، حجم محاسباتی این روشها در حالت کلی، نسبتاً بالا است. چهارم، نرخ بیت خروجی در این روشها در حالت کلی، قابل کنترل نمی‌باشد. پنجم، کارایی و عملکرد این روشها در برخی موارد وابسته به چاپی و یا دستنویس بودن نوع تصویر متنی و حتی نوع زبان متن می‌باشد [۲۸، ۶]. ششم در برخی از روشها از کدگذار/کدگشای مبتنی بر این مدل، از تبدیل موجک استفاده می‌شود که این امر موجب نشت رنگ بین دو ناحیه رنگی یکنواخت و مجاور می‌شود [۲۹]. و هفتم، این گونه روشها با مشکل سربار کدگذاری نیز مواجه‌اند [۱۳-۱۲].

## ۲-۲ فشرده‌سازی تصاویر چهره

در مقایسه با تعداد روشهای فشرده‌سازی تصاویر طبیعی و حتی تصاویر متنی، تعداد نسبتاً کمی روش برای فشرده‌سازی تصاویر چهره ارائه شده است؛ لذا، دسته‌بندی خاصی نیز برای این روشها ارائه نشده است. به نظر، یکی از دلایل کم بودن تعداد روشها، عدم توجه جدی به منحصربفرد بودن چنین تصاویری است [۳۰]. اما در هر حال، با بررسی منابع موجود، روشهای فشرده‌سازی تصاویر چهره را نیز به دو دسته‌ی مبتنی بر مدل و غیر مبتنی بر مدل دسته‌بندی نموده‌ایم.

روشهای مبتنی بر مدل به جای ذخیره‌ی تصویر چهره، مهمترین پارامترهای تصویر چهره که در کاربردهای بازشناسی موثر هستند را ذخیره کرده و گاهاً قادر به بازسازی تصویر چهره نیز می‌باشند. حُسن بزرگ این روشها این است که میزان فشرده‌سازی بالایی دارند؛ اما عیب بزرگ آنها این است که تصویر بازسازی شده تاحدی مصنوعی بوده و چندان به تصویر اصلی نزدیک و شبیه نمی‌باشد. این مشکل از آنجا اهمیت پیدا می‌کند که امکان استفاده از روشهای جدیدتر به منظور پردازش تصویر چهره از بین خواهد رفت زیرا اینگونه روشها قادر به بازسازی تصویر اصلی یا تصویری به حد کافی شبیه به تصویر اصلی نمی‌باشند. روشهای فشرده‌سازی غیرمبتنی بر مدل تصویر چهره را فشرده‌سازی و کدگذاری می‌کنند. این روشها مبتنی بر استفاده‌ی مستقیم یا اصلاح شده از روشها یا استانداردهای عام فشرده‌سازی موجود عمل می‌کنند.

تعداد روشهای مبتنی بر مدل بسیار اندک است. روش پیشنهاد شده در [۳۱] برای کاربردهایی نظیر تلفن ویدیویی، فشرده‌سازی تصاویر پایگاه داده، و بازشناسی چهره پیشنهاد شده است. روش

<sup>1</sup>Karhunen-Loève

<sup>2</sup>Ravine

<sup>3</sup>Image Singularities

<sup>4</sup>Level Sets

<sup>5</sup>Region of Interest (ROI)

<sup>6</sup>Passport

چهره بررسی شده است. برای فشرده‌سازی از JPEG2000 در نرخ بیت‌های مختلف استفاده شده تا بده-بستان بین میزان فشرده‌سازی و صحت بازشناسی محاسبه و بررسی شود. نتایج کارایی بازشناسی (برحسب سه معیار ROC، CMC، و صحت بازشناسی) در دو حالت انجام و عدم انجام فشرده‌سازی تصاویر چهره با یکدیگر مقایسه شده‌اند. همچنین دو حالت درجه‌ی تفکیک پایین (۳۰۰×۳۰۰) و متوسط (۶۰۰×۶۰۰) نیز در بررسی مذکور مورد بررسی قرار گرفت. نتیجه‌ی بسیار مهم و جالبی که از کار این محققان به دست آمد این بود که درجه‌ی تفکیک مکانی پایین در درجه‌ی اول و عامل فشرده‌سازی در درجه‌ی بعدی، موجب اُفت کارایی بازشناسی می‌شوند. نتیجه‌ی دیگر این بود که در یک درجه‌ی تفکیک مکانی ثابت، انجام فشرده‌سازی، در مقایسه با حالت عدم انجام فشرده‌سازی، لزوماً همواره موجب اُفت کارایی بازشناسی نمی‌شود؛ حتی گاه آن را اندکی بهبود می‌دهد! در [۴۷] اثر سه استاندارد فشرده‌سازی JPEG، JPEG2000 و WSQ [۴۸] روی کارایی بازشناسی چهره و اثرانگشت بررسی شده است. کارایی دو روش JPEG2000 و WSQ بهتر از JPEG گزارش شده است. در [۴۹] تاثیر دو روش فشرده‌سازی JPEG و JPEG2000 روی کارایی بازشناسی سه روش بازشناسی شامل PCA، LDA و ICA بررسی شده است. نتایج بیانگر این مطلب بودند که کارایی بازشناسی در حالت کلی با افزایش میزان فشرده‌سازی کاهش می‌یابد. مولفان در [۴۶] نیز روش پیشنهادی خود را بر طبق دو معیار PSNR و SSIM [۵۰] ارزیابی و با روش‌های JPEG2000 و K-SVD مقایسه کرده‌اند. در [۵۱] تاثیر سه روش فشرده‌سازی روی کارایی بازشناسی چهره بررسی شده است. سه روش فشرده‌سازی شامل JPEG، JPEG2000 و JPEG XR است. روش JPEG XR [۵۲] در سال ۲۰۰۷ از سوی شرکت میکروسافت پیشنهاد و به عنوان یک استاندارد صنعتی ثبت گردیده است. روش بازشناسی چهره مورد استفاده مبتنی بر SIFT [۵۳-۵۴] می‌باشد. این ویژگی روشی جهت توصیف ویژگی‌های محلی یک تصویر می‌باشد. یکی از نتایج این است که JPEG XR بهتر از دو روش دیگر عمل کرده است. در [۳۰] پنج روش فشرده‌سازی روی کارایی بازشناسی دو سیستم بازشناسی چهره بررسی شده است. روش‌های فشرده‌سازی عبارتند از: JPEG، JPEG2000، SPIHT، فرکتال، و چندی‌سازی برداری. برای پیاده‌سازی روش فرکتالی از روش درخت چهارتایی<sup>۶</sup> مبتنی بر کد فیشر استفاده شده است. سیستم‌های بازشناسی چهره نیز عبارتند از: یک محصول تجاری به نام VeriLook و یک ابزار تجاری مبتنی بر MATLAB به نام FaRS<sup>۷</sup>. از جمله مهمترین نتایج به دست آمده این است که دو

برای مثال، در [۳۸] از چندی‌سازی برداری دومرحله‌ای مبتنی بر تصحیح ویژگی جهت فشرده‌سازی تصاویر چهره استفاده شده است. نواحی مختلف چهره به طور متفاوتی مورد تاکید و فشرده‌سازی قرار گرفته‌اند طوری که نواحی شامل ویژگی‌های چهره، شامل چشمها و دهان (ناحیه‌ی ROI)، با کیفیت بیشتری فشرده‌سازی می‌شوند. این روش توانسته است تصاویر چهره پرسنلی ۸بیتی به ابعاد ۱۲۸×۱۲۸ را به طور متوسط تا ۳۵۰ بایت فشرده کند.

روش‌های دیگر از فنون مختلف دیگری بهره برده‌اند. در [۳۹] از تجزیه‌ی مقدار منفرد دوتایی (DSVD) که روشی جهت کاهش ابعاد داده‌ها است، جهت فشرده‌سازی و بازشناسی تصاویر چهره استفاده شده است. در [۴۰] از یک روش چندی‌سازی برداری دو مرحله‌ای مبتنی بر تصحیح ویژگی جهت فشرده‌سازی تصاویر رنگی چهره استفاده شده است. در [۴۱] روشی در چارچوب کدگذاری منبع توزیع شده<sup>۲</sup> جهت فشرده‌سازی تصاویر روادید ارائه شده است. در [۴۲] از نمایش چندلایه‌ای تصاویر در جهت فشرده‌سازی آنها استفاده شده است. در [۴۳] از K-SVD<sup>۳</sup> برای فشرده‌سازی تصاویر چهره استفاده شده است. نتایج دارای اثر بلوکی است؛ بنابراین مولفان در ادامه‌ی کار خود در [۴۴] روشی برای رفع اثر بلوکی پیشنهاد داده‌اند. این روش حذف اثر بلوکی مبتنی بر فیلترهای محلی تصویر بوده و برای هر پیکسل یک فیلتر جداگانه طراحی و اعمال می‌شود. در [۴۵] از روش نمایش تنگ فراکامل<sup>۴</sup> برای فشرده‌سازی تصاویر چهره استفاده شده است. ابتدا به کمک روشی به نام K-LMS یک دیکشنری وفقی ایجاد می‌شود. سپس به کمک الگوریتم ردگیری انطباق متعامد<sup>۵</sup>، تصویر تصویر چهره برحسب اتمهای دیکشنری به صورت تنگ نمایش داده می‌شود. در پایان، ضرایب این نمایش تنگ به کمک کدگذاری هافمن فشرده و کدگذاری می‌شوند. در [۴۶] از نوعی تبدیل موجک مبتنی بر درخت دارای تزیاید برای فشرده‌سازی تصاویر چهره استفاده شده است. این روش مبتنی بر دوباره‌چینی پیکسل‌های تصویر به گونه‌ای است که موجب تولید یک سیگنال یک بعدی هموار شده و بنابراین، کارایی تبدیل موجک را افزایش می‌دهد. ضرایب تبدیل به کمک کدگذار آنتروپی کدگذاری می‌شوند.

برخی محققان اقدام به بررسی تاثیر برخی روش‌های فشرده‌سازی روی کارایی بازشناسی نموده‌اند؛ کاری که در زمینه‌ی تصاویر متنی تقریباً دیده نشده است. در [۳۷] اثرات عوامل مختلفی مانند فشرده‌سازی باتلاف تصویر چهره، تفاوت‌های فردی (مانند عینک و مدل مو)، و محیط تهیه‌ی تصاویر چهره روی کارایی بازشناسی

<sup>۱</sup>Dyadic Singular Value Decomposition

<sup>۲</sup>Distributed Source Coding

<sup>۳</sup>Singular Value Decomposition

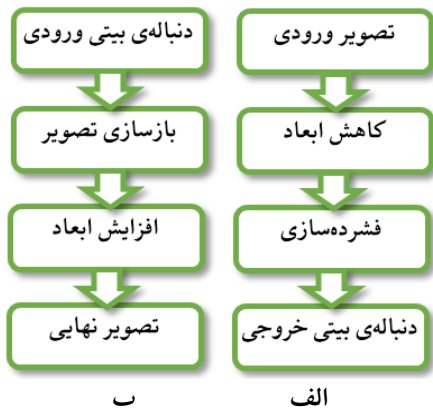
<sup>۴</sup>Over-complete Sparse Representation

<sup>۵</sup>Orthogonal Matching Pursuit

<sup>۶</sup>Quadtree

<sup>۷</sup>Face Recognition System

دست می‌آید به عنوان یک تصویر متنی مستقل در مراحل بعدی استفاده می‌شود. لذا روش‌هایی مانند JPEG، JPEG2000 و SPIHT هیچگاه لازم نیست با جزئیات روش کاهش ابعاد پیشنهادی درگیر شوند. آنها تنها یک تصویر معمولی را به عنوان ورودی دریافت می‌کنند. البته ما می‌دانیم که این تصویر ورودی، حاصل کاهش ابعاد تصویر اصلی به کمک روشی پیشنهادی مبتنی بر تبدیل موجک بوده و در واقع آخرین زیرباند تقریب تجزیه موجک در این روش پیشنهادی است. مهمترین بلوک‌های مورد استفاده در شکل ۱ شامل «کاهش ابعاد» و «افزایش ابعاد» بوده و جزئیات آنها در زیربخش‌های بعدی توضیح داده می‌شوند.



شکل ۱: دیاگرام بلوکی روش پیشنهادی مربوط به دو مسیر (الف) فشرده‌سازی، و (ب) بازسازی.

### ۳-۱ کاهش ابعاد

هر بار که تجزیه‌ی موجک روی تصویری اعمال شود، ابعاد تصویر تقریباً به اندازه‌ی یک چهارم ابعاد مرحله‌ی قبل خود کاهش می‌یابد؛ برای مثال، اگر تجزیه‌ی موجک دوسطحی روی تصویری اعمال شود، ابعاد تصویر تقریباً در سطح دوم تجزیه حدود یک شانزدهم ابعاد اولیه خواهد بود. بنابراین، تبدیل موجک راهی برای کاهش ابعاد است. در روش پیشنهادی، این که تصویر ورودی را تا چند مرحله (L) تجزیه کرده و از زیرتصویر تقریباً آن سطح در ادامه‌ی کار استفاده کنیم، برای هر تصویر به طور خودکار تعیین می‌شود.

در هر سطح از تجزیه‌ی موجک، سه زیرباند جزئیات وجود دارند که جزئیات افقی ( $D^h$ )، عمودی ( $D^v$ )، و قطری ( $D^d$ ) نامیده می‌شوند. کمیت انرژی نرمالیزه شده به مساحت،  $E_N$ ، را برای هر زیرباند به صورت تقسیم انرژی آن زیرباند بر تعداد کل پیکسل‌های آن زیرباند تعریف می‌کنیم. در روش پیشنهادی، تجزیه‌ی تصویر ورودی را تا جایی (یا «شماره‌ی سطحی») ادامه می‌دهیم که حداقل یکی از زیرباند‌های آن سطح تجزیه، شرط زیر را برآورده کند (یعنی به طور معادل، انرژی حداقل یکی از زیرباند‌ها، «قابل توجه» شود). در این صورت، مقدار پارامتر L برابر با شماره‌ی این سطح خواهد بود.

روش SPIHT و JPEG2000 به ترتیب بهترین نتایج فشرده‌سازی و نیز بازشناسی چهره را به دست آوردند. در پایان دو نکته‌ی دیگر نیز قابل توجه است: اول این که تقریباً تمامی روش‌های موجود جهت فشرده‌سازی تصاویر چهره از تصاویری با درجه‌ی تفکیک مکانی پایین (به طور معمول تصاویری با ابعاد حدودی  $200 \times 200$ ) استفاده کرده و هیچ یک مساله‌ی استفاده از تصاویر با درجه‌ی تفکیک بالا را بررسی نکرده‌اند. از جمله مهمترین دلایل این امر حجم بسیار محدود در نظر گرفته شده برای ذخیره‌ی تصویر چهره در کاربردهای مربوطه می‌باشد [۴۰]. این مطلب حتی در شبه‌استانداردهای وضع شده‌ی موجود نیز دیده می‌شود: سازمان بین‌المللی هوانوردی غیرنظامی (ICAO) در سال ۲۰۰۲ یک گروه کاری برای تدوین استاندارد جهت تهیه اسناد مسافرتی قابل خواندن توسط ماشین (MRTD<sup>۲</sup>) تشکیل داد که این گروه کاری به استفاده از تصویر چهره در زمره‌ی علائم حیاتی مورد نیاز جهت شناسایی افراد تاکید نمود. این گروه یک قالب فایل به نام M1 جهت تبادل داده‌ها و ذخیره روی چیپ الکترونیکی تعیین نمود [۵۵] که اطلاعاتی راجع به تصویر دیجیتالی چهره را مشخص می‌کند. در این قالب فایل توصیه شده است که تصاویر به کمک استاندارد JPEG یا JPEG2000 در نرخ بیت 0.44 bpp فشرده شوند تا حجم فایلی در حدود 12 KB ایجاد شود (معادل با یک تصویر با ابعاد تقریبی  $180 \times 180$  پیکسل). حداقل درجه‌ی تفکیک مکانی توصیه شده در این قالب فایل، معادل با وجود حداقل ۹۰ پیکسل در فاصله‌ی بین دو چشم فرد می‌باشد.

دومین نکته‌ی قابل تامل این است که علی‌رغم وجود تعداد نسبتاً زیادی پایگاه تصاویر استاندارد یا شناخته شده [۵۴] متأسفانه تعداد نسبتاً زیادی از روش‌های موجود از پایگاه تصاویر مخصوص به خود و با تعداد تصاویر کم استفاده نموده‌اند.

### ۳ روش پیشنهادی

دیاگرام بلوکی روش پیشنهادی مربوط به فشرده‌سازی و بازسازی تصویر به ترتیب در شکل‌های ۱-الف و ۱-ب نمایش داده شده است. در مسیر فشرده‌سازی، ابتدا ابعاد تصویر ورودی به کمک تبدیل موجک کاهش داده شده و سپس به کمک هر روش دلخواه (مانند JPEG و JPEG2000) فشرده‌سازی می‌شود. در مسیر بازسازی، ابتدا از کدگشای متناسب با مسیر فشرده‌سازی (مانند کدگشای JPEG و JPEG2000) تصویر اولیه با ابعاد کاهش‌یافته بازسازی شده و سپس ابعاد آن افزایش داده می‌شود تا به همان ابعاد اولیه‌ی خود، قبل از فشرده‌سازی، برگردد. در مسیر فشرده‌سازی (شکل ۱) پس از کاهش ابعاد تصویر ورودی به کمک تبدیل موجک، از زیرتصویر تقریبی که در انتها به

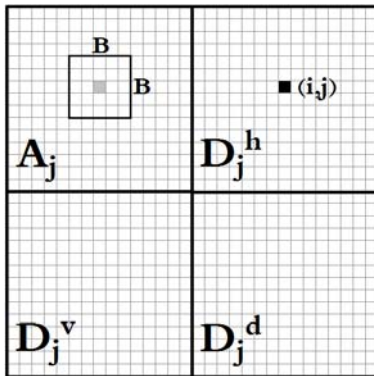
<sup>1</sup> International Civil Aviation Organization

<sup>2</sup> Machine-Readable Travel Documents

خواهد شد) باید به تعداد  $3 \times L$  بار تکرار شود تا همین تعداد مجموعه‌ی طبقات تشکیل و ذخیره شوند (هر مجموعه‌ی طبقه شامل ۲۵۶ طبقه متناظر با ۲۵۶ مقدار مختلف برای یک پیکسل فرضی از زیرتصویر جزئیات می‌باشد).

حال در مرحله‌ی بازسازی، بسته به این که بردار ویژگی  $F_{(i,j)}$  متعلق به کدام سطح تجزیه و کدام زیرباند جزئیات است، مجموعه‌ی طبقات مناسب انتخاب شده و این بردار ویژگی به یکی از طبقه‌های موجود در مجموعه تخصیص داده می‌شود. شماره‌ی طبقه‌ی تخصیص داده شده (اندیس  $k$  در نمادگذاری اخیر)، تعیین کننده‌ی مقدار پیکسل واقع در مختصات  $(i,j)$  از زیرتصویر جزئیات افقی مورد استفاده در این مثال خواهد بود.

با توجه به اهمیت دو مرحله‌ی «آموزش و تشکیل مجموعه‌ی طبقات» و «طبقه‌بندی بردارهای ویژگی»، زیربخش‌های بعدی این مراحل را توضیح خواهند داد.



شکل ۲: نمایش ساختار یک سطح از تجزیه به کمک تبدیل موجک؛ در هر سطح چهار زیرباند متناظر با هم وجود دارند: یک زیرباند (یا زیرتصویر) تقریب و سه زیرباند جزئیات (افقی، عمودی، و قطری).

### ۳-۳ آموزش و تشکیل مجموعه‌ی طبقات

مراحلی که در این بخش توضیح داده می‌شوند برای تشکیل هر مجموعه‌ی طبقات متناظر با یک سطح تجزیه‌ی خاص و یک زیرباند جزئیات خاص باید به طور مشابه تکرار شوند. تنها تفاوتی که وجود دارد این است که برای تهیه‌ی داده‌های آموزشی از زیرباندهای جزئیات خاص و سطح تجزیه‌ی خاصی باید استفاده شود.

دیاگرام بلوکی مرحله‌ی آموزش و تشکیل یک مجموعه از طبقات (متناظر با یک سطح تجزیه خاص و یک زیرباند جزئیات خاص) در شکل ۳ نشان داده شده است. این دیاگرام بلوکی نشان دهنده‌ی چهار قدم اساسی شامل «استخراج ویژگی»، «برچسب‌زنی»، «اصلاح خوشه‌ها» و «تعیین نمایندگان» است. مطابق با این دیاگرام بلوکی، نقطه‌ی شروع کار جهت تشکیل بردارهای ویژگی (یا همان داده‌های آموزشی) زیرتصویر تقریب مربوط به سطح تجزیه‌ی مورد نظر می‌باشد. این زیرتصویرهای تقریب از روی تعدادی تصویر آموزشی (مانند تصاویر چهره) و تبدیل موجک محاسبه می‌شوند. برای هر یک از پیکسل‌های واقع

$$E_N \geq T \quad (1)$$

در رابطه‌ی فوق،  $T$  آستانه‌ای است که مقدار آن برابر ضریبی ( $\alpha$ ) از انرژی نرمالیزه شده به مساحت تصویر ورودی ( $I$ ) تعیین می‌شود:

$$T = \frac{\alpha \times \sum_{i=1}^M \sum_{j=1}^N I(i,j)^2}{M \times N} \quad (2)$$

در رابطه‌ی فوق،  $M$  و  $N$  ابعاد تصویر ورودی می‌باشند. در حالت کلی، هر قدر مقدار  $L$  بزرگتر شود، میزان کاهش ابعاد تصویر ورودی بیشتر و میزان فشرده‌سازی روش پیشنهادی نیز بیشتر خواهد شد.

### ۲-۳ افزایش ابعاد

در مسیر بازسازی، شکل ۱-ب، پس از بازسازی تصویر اولیه به کمک کدگشای مناسب (مانند JPEG و JPEG2000)، این تصویر به عنوان زیرتصویر تقریب در سطح تجزیه‌ی  $L$  ام در نظر گرفته شده و به کمک عکس تبدیل موجک، به دنبال بازسازی زیرتصویرهای تقریب سطوح پایین‌تر خواهیم بود تا جایی که به تصویری با ابعاد اولیه (قبل از فشرده‌سازی) برسیم. در هر سطح، برای بازسازی هر زیرتصویر تقریب سطح پایین‌تر نیاز به داشتن اطلاعات زیرباندهای جزئیات متناظر با سطح فعلی داریم. برای مثال، مطابق با شکل ۲، برای بازسازی تصویر تقریب سطح  $1-j$  علاوه بر زیرتصویر تقریب  $A_j$  (متناظر با سطح  $j$  ام) به اطلاعات سه زیرباند جزئیات مربوط به سطح  $j$  ام (یعنی زیرباندهای  $D_j^h$ ،  $D_j^v$  و  $D_j^d$  در شکل ۲) نیز نیاز داریم. در روش پیشنهادی، این اطلاعات را به کمک یک فرآیند پیشنهادی مبتنی بر طبقه‌بندی تخمین می‌زنیم. نحوه‌ی عملکرد این فرآیند (که در زیربخش‌های بعدی بیشتر توضیح داده خواهد شد) به این صورت است که برای مثال اگر بخواهیم مقدار پیکسل واقع در مختصات  $(i,j)$  از زیرتصویر جزئیات افقی که در شکل ۲ با رنگ سیاه مشخص شده است را تخمین بزنیم، ابتدا بلوکی به ابعاد مشخص  $B$  واقع در زیرتصویر تقریب و به مرکز همان مختصات  $(i,j)$  را در نظر می‌گیریم (این بلوک در شکل ۲ به مرکز پیکسلی با رنگ خاکستری مشخص شده است). از این بلوک تعدادی ویژگی استخراج و به کمک آنها یک بردار ویژگی  $F_{(i,j)}$  تشکیل می‌دهیم. در روش پیشنهادی، قبل از شروع فرآیند فشرده‌سازی-بازسازی، چند مرحله «آموزش» به منظور تشکیل چند مجموعه از طبقات،  $\{L_{k,j,o}\}, k=0, 1, \dots, 255; j=1, \dots, L; o=\{h,v,d\}$  یا اندیس  $k$  اندیس شماره‌ی هر یک از طبقه‌های موجود در یک مجموعه‌ی فرضی،  $j$  اندیس مربوط به سطح تجزیه و  $o$  اندیس مربوط به نوع زیرباند جزئیات می‌باشد) انجام شده است؛ ملاحظه می‌شود که برای هر زیرباند جزئیات متعلق به یک سطح تجزیه‌ی خاص، یک مرحله آموزش به منظور تشکیل مجموعه‌ی طبقات لازم به انجام است؛ به عبارت دیگر، با در نظر گرفتن تعریف  $L$  به صورتی که بیان شد، مرحله‌ی آموزشی (که در زیربخش‌های بعدی بیشتر توضیح داده

طبقه‌بندی نادرست آنها در مرحله‌ی بازسازی از بین رود. مراحل انجام اصلاح به صورت زیر می‌باشد:

الف) مراکز دسته‌ها را با محاسبه‌ی میانگین نقاط هر دسته محاسبه و تعیین کنید،

ب) یک دسته را به عنوان دسته‌ی فعلی در نظر بگیرید. نام این دسته را  $G_0$ ، جمعیت آن را با  $P_0$ ، و مرکز آن را با  $C_0$  نمایش دهید،

ج) فاصله‌ی مرکز دسته‌ی فعلی از بقیه‌ی دسته‌ها را تعیین کرده و این فاصله‌ها را به ترتیب صعودی مرتب کنید تا دنباله‌ی فاصله‌های مرتب شده،  $\{d_k\}$ ، تشکیل شود،

د) دسته‌ی متناظر با  $d_1$  را در نظر بگیرید. نام این دسته را  $G_1$ ، جمعیت آن را با  $P_1$ ، و مرکز آن را با  $C_1$  نمایش دهید،

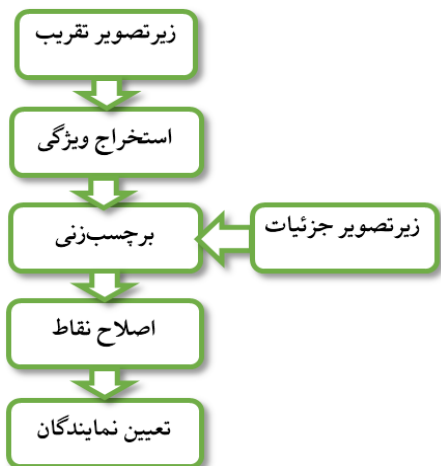
ه) تمام نقاطی از دسته‌ی فعلی که شرط زیر را برقرار می‌کنند، حذف کنید:

$$R_0 > \frac{P_0}{P_1} R_1 \quad (5)$$

در رابطه‌ی فوق،  $R_0$  و  $R_1$  فاصله‌ی نقطه‌ی مورد بررسی از به ترتیب  $C_0$  و  $C_1$  می‌باشند. مبنای طراحی شرط فوق این است که هر دسته‌ای که جمعیت بیشتری داشته باشد نقاط متنوع‌تری در فضا داشته و بنابراین، در مرحله‌ی بازسازی، قدرت جذب نقاط بیشتری را به خود پیدا خواهد کرد.

و) مراحل (د) و (ه) را برای دسته‌های متناظر با  $d_2$  و  $d_3$  تکرار کنید،

ز) مراحل (ب) تا (و) را برای بقیه‌ی دسته‌ها تکرار کنید. در آخرین قدم از دیاگرام بلوکی شکل ۳، برای کاهش حجم محاسبات در مرحله‌ی بازسازی، برای هر دسته تعدادی نماینده تعیین می‌کنیم. اگر سطح محیطی بر نقاط متعلق به یک دسته‌ی فرضی را تصور کنیم، نمایندگان به گونه‌ای تعیین می‌شوند که دو شرط را برقرار کرده باشند: اول، تا حد امکان روی سطح یا نزدیک به آن قرار داشته باشند، و دوم، به طور یکنواخت روی این سطح توزیع شده باشند؛ به عبارت دیگر تمام نواحی سطح را تا حد امکان پوشش داده باشند.



شکل ۳: دیاگرام بلوکی مربوط به روش پیشنهادی جهت آموزش و تشکیل یک مجموعه از طبقات به منظور استفاده در فرآیند بازسازی.

در این زیرتصویر به مختصات  $(i,j)$ ، مانند پیکسلی که با رنگ خاکستری در شکل ۲ مشخص شده است، یک بلوک به ابعاد  $B$  و به مرکزیت پیکسل نامبرده شده (مطابق با شکل ۲) در نظر گرفته و از آن تعدادی ویژگی استخراج می‌کنیم (قدم اول). برای محاسبه‌ی این ویژگی‌ها ابتدا روی بلوک مذکور، سه سطح تجزیه موجک انجام داده تا ۹ زیرباند جزئیات حاصل شود. واریانس هر یک از زیرباندهای جزئیات جزو بردار ویژگی خواهند بود. علاوه براین، برای هر زیرباند، نسبت واریانس زیرباند جزئیات عمودی به واریانس زیرباند جزئیات افقی را نیز محاسبه نموده و به جمع ویژگی‌ها اضافه می‌کنیم. بدین ترتیب بردار ویژگی استخراج شده از بلوک مذکور شامل ۱۲ ویژگی به صورت زیر خواهد بود:

$$x^{\text{block}} = (\sigma_{1h}^2, \sigma_{1v}^2, \sigma_{1d}^2, a_1, \dots, \sigma_{3h}^2, \sigma_{3v}^2, \sigma_{3d}^2, a_3) \quad (3)$$

که در آن  $a_j$  نسبت واریانس زیرباند جزئیات عمودی به واریانس زیرباند جزئیات افقی برای سطح تجزیه‌ی  $j$ ام می‌باشد:

$$a_j = \frac{\sigma_{jv}^2}{\sigma_{jh}^2} \quad (4)$$

مبنای انتخاب ویژگی‌ها به صورت فوق، قابلیت آنها در ایجاد تمایز بین تصاویر مختلف و مناسب بودن در کاربردهایی مانند بازیابی تصویر است [۵۶]. دلیل توجهی دیگر این است که در تبدیل موجک، زیرباندها خروجی فیلترهای بالاگذر بوده و بنابراین هرچه لبه‌های تصویر تیزتر باشند، مقادیر بزرگتری در زیرباندهای جزئیات تولید می‌شود. فیلترهای بالاگذر نوعی لبه‌یاب محسوب می‌شوند. از طرف دیگر، ویژگی واریانس نیز معرف قدرت لبه‌ها در تصویر است؛ بنابراین بین ویژگی واریانس و خروجی تبدیل موجک همبستگی وجود داشته و از واریانس برای تخمین مقادیر زیرباندهای جزئیات می‌توان استفاده کرد.

بردار ویژگی که بدین ترتیب محاسبه شود متناظر با پیکسل واقع در مختصات  $(i,j)$  از زیرتصویر تقریب است. این بردار را متعلق به طبقه‌ای در نظر می‌گیریم (قدم دوم از دیاگرام بلوکی شکل ۳) که شماره‌ی آن برابر مقدار پیکسل واقع در مختصات  $(i,j)$  از زیرتصویر جزئیات نرمالیزه شده می‌باشد. منظور از زیرتصویر جزئیات نرمالیزه شده، زیرتصویر جزئیاتی است که مقادیر پیکسل‌های آن به بازه‌ی  $[0, 255]$  نرمالیزه شده باشد. نرمالیزه کردن نسبت به مقدار بیشینه‌ی موجود در چهار زیرباند سطح تجزیه‌ی فعلی انجام می‌شود. بدین ترتیب با استفاده از تصاویر آموزشی مختلف، برای هر طبقه (یا دسته) تعدادی بردار ویژگی به دست می‌آید که هر یک از این بردارها به عنوان یک نقطه در فضای  $n$  بُعدی (در اینجا  $n = 12$ ) متعلق به دسته‌ی خود می‌باشد.

در سومین قدم از دیاگرام بلوکی شکل ۳، نقاط هر طبقه باید اصلاح شوند. در این اصلاح، نقاطی که فاصله‌ی آنها از مرکز دسته‌ی خود، در مقایسه با فاصله‌شان از نزدیک‌ترین دسته‌ی همسایه، بیش از حد تعیین شده باشد، حذف می‌شوند تا اثر



ج) تابع فاصله‌های تجمعی،  $R_k(r)$ ، را به صورت زیر برای یک مقدار مشخص و از قبل تعیین شده  $k=K$  و تمامی مقادیر ممکن  $r$  (یعنی  $r=1, \dots, 256$ ) محاسبه می‌کنیم:

$$R_k(r) = \sum_{m=1}^k D_{m,r} \quad (V)$$

د) اندیس محل وقوع کمینه‌ی تابع فوق همان شماره‌ی دسته‌ی تخصیص یافته به بردار ویژگی  $F(i,j)$  و بنابراین، مقدار تخمین زده شده برای زیرباند جزئیات می‌باشد:

$$\text{Estimated Value} = \operatorname{argmin}_r \{R(r)\} \quad (A)$$

### ۳-۵ توجیه شهودی-استدلالی روش پیشنهادی

ممکن است سوالات یا ابهاماتی از این قبیل مطرح شود که چگونه می‌توان ضرایب زیرباند جزئیات هر سطح را از پیکسل متناظرش در زیرباند تقریب (واقع در یک همسایگی) تخمین زد؟ همان گونه که می‌دانیم تبدیل موجک یک روش تجزیه‌ی تصویر به زیرباندهای مختلف و در سه جهت اساسی افقی، عمودی و قطری است. چگونه اطلاعات در زیرباند مثلاً  $H$  که محتوای فرکانسی بالاتری نسبت به زیرباند تقریب  $A$  دارد را می‌توان با آموزش تعدادی ویژگی از روی یک همسایگی در زیرباند  $A$  آن هم با گرفتن سه سطح تبدیل موجک انجام داد؟ زیرا به نظر می‌رسد محتوای فرکانسی تبدیل موجک سه سطح همسایگی  $B$  هیچ ارتباطی با محتوای فرکانسی پیکسل متناظر در زیرباند  $H$  ندارد. در پاسخ به این سوال و ابهامات مشابه می‌توان به موارد زیر اشاره کرد:

۱- مولفه‌های فرکانسی در خروجی تبدیل فوریه مستقل از هم نیستند که نتوان از روی یک مولفه، تقریبی از مولفه‌ی دیگر به دست آورد. اندازه‌ی هر مولفه فرکانسی (یعنی خروجی تبدیل فوریه در یک فرکانس خاص مانند  $w_0$ ) برابر با «متوسط سهم حضور» نوسان سینوسی دارای فرکانس  $w_0$  در کل سیگنال است. به عبارت دیگر رابطه‌ی تبدیل فوریه همان رابطه‌ی ضرب داخلی سیگنال مورد نظر در یک تابع نمایشی مختلط (همان سینوسی در حالت ساده) بوده و همان طور که می‌دانیم ضرب داخلی به معنای «میزان شباهت» یا همان «سهم حضور» است. حال یک رفتار دلخواه از سیگنال ورودی را در نظر بگیرید. این رفتار ممکن است با سینوسی‌های بسیاری (دارای فرکانسهای مختلف) (درجه‌ای از) شباهت داشته باشد. بنابراین، این رفتار موجب ایجاد مولفه‌ی فرکانسی در فرکانسهای مختلف می‌شود؛ در نتیجه، مولفه‌های فرکانسی مختلف (در طیف خروجی تبدیل فوریه) لزوماً مستقل از هم نبوده و دارای درجاتی از همبستگی هستند. همین امر این نوید را می‌دهد که با در دست داشتن مقدار تعدادی مولفه‌ی فرکانسی بتوانیم مقدار یک مولفه‌ی فرکانسی دیگر را «تخمین» بزنیم.

۲- در حوزه فرکانسی تحلیل بند قبلی ارائه شد، در حوزه مکانی میتوان به تحلیل روشن‌تر و ملموس‌تری دست یافت. در این کار

مراحل تعیین نمایندگان هر دسته، به صورت زیر است:  
الف) ماتریس فاصله‌های دویه دوی نقاط دسته‌ی فعلی را تشکیل دهید؛ اگر عنصر واقع بر سطر  $i$  و ستون  $j$  از این ماتریس را با  $D_{i,j}$  نمایش دهیم آن گاه این عنصر برابر با فاصله‌ی اقلیدسی بین نقاط  $i$  و  $j$  از این دسته خواهد بود. بدیهی است که  $D_{i,j}=D_{j,i}$  بوده و بنابراین ماتریس مذکور نسبت به قطر اصلی خود متقارن بوده و نیز دارای عناصر صفر واقع بر این قطر می‌باشد.

ب) لیست نمایندگان را برابر تھی قرار دهید:  $L=\{\}$   
ج) برای مقادیر طبیعی متغیر  $n$  اعمال زیر را تکرار کنید تا زمانی که تعداد اعضای لیست نمایندگان به حد تعیین شده،  $N$ ، برسد:

- ۱)  $n$  اُمین عدد بزرگ در ماتریس مذکور را پیدا کنید،
- ۲) فرض کنید این عدد،  $D_{i,j}$  بوده و بنابراین مربوط به فاصله‌ی دو نقطه‌ی  $i$  و  $j$  می‌باشد،
- ۳) هر یک از نقاط  $i$  و  $j$  که شرط زیر را برقرار کنند، به لیست نمایندگان اضافه کنید:

$$d_{\min} > \beta R \quad (E)$$

که در آن،  $d_{\min}$  کمینه فاصله‌ی نقطه‌ی مورد بررسی از نقاط موجود در لیست نمایندگان،  $\beta$  یک ضریب ثابت، و  $R$  شعاع دسته می‌باشند. شعاع هر دسته را به سادگی برابر میانگین فاصله‌های نقاط از مرکز دسته تعریف می‌کنیم.  
د) مراحل (الف) تا (ج) را برای تمام دسته‌ها تکرار کنید.

### ۳-۴ طبقه‌بندی بردارهای ویژگی

همان طور که در بخش ۳-۲ بیان شد، هر بردار ویژگی  $F(i,j)$  (مطابق با نمادگذاری بخش ۳-۲) می‌بایست به نحوی با دسته‌های موجود در مجموعه طبقات متعلق به یک زیرباند جزئیات مفروض و یک سطح تجزیه‌ی مفروض، مقایسه و به یکی از آن‌ها تخصیص یابد. تعداد دسته‌ها در حالت کلی ۲۵۶ دسته است. هر دسته شامل تعدادی ( $N$ ) نماینده است که در بخش ۳-۳ تعیین شدند؛ بنابراین، برای تعیین میزان تعلق هر بردار ویژگی به دسته‌ها از نمایندگان دسته‌ها استفاده می‌کنیم. مراحل انجام طبقه‌بندی و تخصیص یکی از دسته‌های موجود به بردار ویژگی ورودی،  $F(i,j)$ ، به صورت زیر است:

الف) فاصله‌ی اقلیدسی بردار ورودی را از تمام نمایندگان تمام دسته‌ها محاسبه کرده و فاصله‌های مربوط به هر دسته را به ترتیب صعودی مرتب می‌کنیم.  
ب) فاصله‌های مرتب شده را داخل ماتریس فاصله‌ها،  $[D_{p,q}]$ ، قرار می‌دهیم (منظور از  $D_{p,q}$  عنصر واقع در سطر  $p$  و ستون  $q$  از این ماتریس است). این ماتریس دارای  $N$  سطر و ۲۵۶ ستون بوده و عناصر موجود در ستون  $q$  اُم آن همان فاصله‌های مرتب شده مربوط به دسته‌ی  $q$  اُم می‌باشند.

ناهمواری) افقی است. همان طور که در زیرباند افقی نیز با شکل مربع مشخص شده است، این زیرباند در ناحیه‌ی متناظر، دارای ناهمواری قوی و بزرگی است اما در دیگر زیرباند‌ها چنین رفتاری وجود ندارد. (۴) ناحیه‌ی مشخص شده با علامت دایره در زیرباند تقریب، مربوط به یک ناهمواری یا لبه در هر دو جهت افقی و عمودی (یعنی راستای قطری) است؛ همان طور که در سه زیرباند جزئیات نیز با علامت دایره مشخص شده این زیرباند‌ها نیز دارای رفتار مشابهی هستند که البته شدت این رفتار در زیرباند قطری قوی‌تر از دو زیرباند دیگر است.

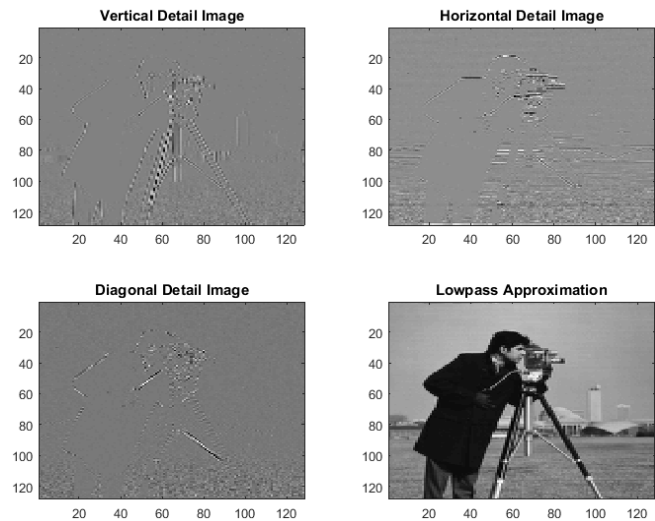
با توجه به مثال و بررسی فوق، ملاحظه می‌شود که از روی رفتارها و ویژگی‌های موجود در زیرباند تقریب می‌توان «انتظار داشت» یک تخمین نسبتاً مناسب (یعنی با درجه‌ای از دقت) از هر یک از زیرباند‌های سه‌گانه‌ی جزئیات به دست آورد.

در ادامه توجه به این تحلیل نیز ممکن است مفید باشد: «زیرباند تقریب، خروجی فیلترهای پایین‌گذر و زیرباند‌های جزئیات، خروجی (حداقل یک بار اعمال) فیلتر بالاگذر هستند که روی «یک» سیگنال اصلی «مشترک» اعمال شده‌اند. لذا معقول است که انتظار داشته باشیم بین این «دو نوع» زیرباند درجاتی از همبستگی وجود داشته باشد (البته تا زمانی که به نقطه‌ی بحرانی اشاره شده در بند ۴ نرسیده باشیم). برای توضیح بیشتر سه حالت را در نظر بگیرید؛ حالت اول، اگر زیرباند تقریب در هر دو جهت افقی و عمودی تا حد زیادی صاف و هموار باشد، احتمال «زیادی» وجود دارد که زیرباند‌های جزئیات و در نتیجه به تبع آنها زیرباند قطری، دارای انرژی کوچک و ناچیز باشند. همچنین، هر قدر زیرباند تقریب هموارتر و صاف‌تر باشد، انرژی زیرباند‌های جزئیات نیز کمتر خواهد شد.

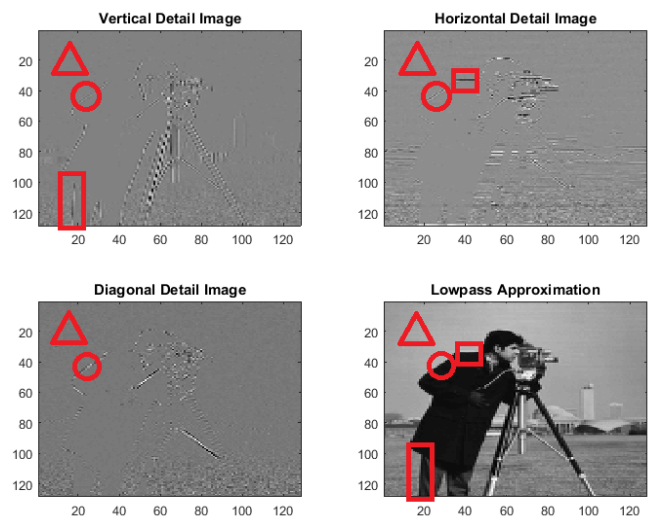
حالت دوم، اگر زیرباند تقریب در یک راستای خاص (مثلاً افقی) دارای ناهمواریهایی باشد اما در راستای دیگر (یعنی عمودی) ناهمواری قابل توجهی نداشته باشد، «به احتمال زیاد» زیرباند متناظر با راستای افقی (یعنی در اینجا زیرباند H) دارای انرژی بالا و زیرباند راستای دیگر (یعنی زیرباند V) انرژی پایینی دارد؛ در همین حال، زیرباند قطری دارای انرژی متوسط (بین دو مقدار قبلی) خواهد بود. یک نکته دیگر در این حالت این است که هر قدر شدت «ناهمواری/همواری» در زیرباند تقریب در یک راستای خاص بیشتر شود انرژی زیرباند متناظر نیز «بیشتر/کمتر» می‌شود. به طریق مشابه، حالت سوم، اگر سیگنال تقریب در هر دو جهت افقی و عمودی دارای ناهمواریهایی باشد هر دو زیرباند افقی و عمودی و به تبع آنها، زیرباند قطری، (به احتمال زیاد) دارای انرژی بیشتری (نسبت به حالت قبل) خواهند بود. در این حالت نیز هر قدر شدت ناهمواری در زیرباند تقریب بیشتر باشد، انرژی زیرباند‌ها بیشتر خواهد شد.»

۳- هر قدر در تجزیه‌ی موجک جلوتر برویم، در زیرباند‌های جزئیات وارد فرکانسهای پایین و پایین‌تر می‌شویم حال در عمل تجزیه از حدی که بیشتر جلوتر برویم وارد طیف مهم مربوط به

ابتدا یک نمونه تجزیه نمایش داده می‌شود. با توجه به شکل ۴ که مربوط به چهار زیرباند حاصل از تجزیه‌ی سطح اول یک تصویر است، ملاحظه می‌شود که بین زیرباند‌های جزئیات و زیرباند تقریب همبستگی وجود دارد.



شکل ۴: یک نمونه تجزیه تک سطحی یک تصویر



شکل ۵: مشخص کردن وجود تناظر بین برخی نواحی واقع در زیرباند تقریب و زیرباند‌های جزئیات

برای نمایش بهتر، شکل ۴ در شکل ۵ مجدداً در حالی نمایش داده شده است که چهار ناحیه‌ی مختلف در زیرباند تقریب مشخص شده است. (۱) ناحیه‌ی مشخص شده با مستطیل در زیرباند تقریب مربوط به ناحیه‌ای شامل لبه‌ی عمودی است. همان طور که در زیرباند عمودی نیز با شکل مستطیل مشخص شده است، این زیرباند در ناحیه‌ی متناظر، دارای ناهمواری عمودی قوی است اما در دیگر زیرباند‌ها چنین ناهمواری‌ای وجود ندارد. (۲) ناحیه‌ی مشخص شده با علامت مثلث در زیرباند تقریب مربوط به یک ناحیه‌ی هموار در هر دو جهت افقی و عمودی است. لذا همان طور که در سه زیرباند جزئیات نیز با علامت مثلث مشخص شده است، این سه زیرباند نیز دارای بخش هموار در ناحیه‌ی متناظر هستند. (۳) ناحیه‌ی مشخص شده با علامت مربع در زیرباند تقریب مربوط به ناحیه‌ای شامل لبه‌ی (یا

مفید بوده و اطلاعاتی که قبلاً مفید بود حالا ممکن است چندان مفید نباشند. این گونه رابطه‌ها را انتظار داریم که طبقه‌بند فراگرفته باشد تا در مرحله‌ی «آزمایش» بتواند اطلاعات مفید و «نه چندان مفید» را از هم تمیز دهد.

#### ۴ نتایج شبیه‌سازی

گرچه روش پیشنهادی روی هر نوع تصویری قابل اعمال است، اما جهت نمایش کارایی فشرده‌سازی روش پیشنهادی و نیز با توجه به تاثیر ماهیت طیفی تصویر روی این کارایی، دو نوع تصویر متنی و چهره برای انجام این ارزیابی انتخاب شدند. تصاویر متنی دارای ماهیت طیفی میان‌گذر و تصاویر چهره دارای ماهیت طیفی پایین‌گذر می‌باشند. برای هر یک از این دو نوع تصویر از یک پایگاه تصاویر مناسب دارای دو درجه‌ی تفکیک مکانی مختلف بالا و متداول استفاده شد. در تصاویر متنی درجات تفکیک مکانی بالا و متداول به ترتیب عبارتند از dpi ۶۰۰ و ۳۰۰. برای تصاویر چهره، درجه‌های تفکیک انتخاب شده عبارتند از ۲۰۴۸×۱۵۳۶ (بالا)، ۶۴۰×۴۸۰ (نسبتاً بالا) و ۲۰۰×۱۵۰ (پایین). برای تهیه تصاویر متنی از روبش ۱۲۴ سند کاغذی دارای زبان متن انگلیسی و با کیفیت معمولی به کمک دستگاه روبش Avison مدل AV210D2+ استفاده گردید. برای تهیه تصاویر چهره نیز از پایگاه تصاویر MIT CBCL [۵۷] مربوط به مرکز آموزش محاسباتی و بیولوژیکی دانشگاه MIT به همراه پایگاه تصاویر دانشگاه جورجیا [۵۸] استفاده شده است. این پایگاه‌ها دارای تصاویری با درجه‌ی تفکیک مکانی به ترتیب بالا و نسبتاً بالا بوده و برای تهیه‌ی یک نسخه‌ی با درجه‌ی تفکیک مکانی متداول (پایین) از روی تصاویر این پایگاه‌ها، تصاویر را در محیط MATLAB کاهش ابعاد داده‌ایم<sup>۲</sup>. نمونه‌هایی از تصاویر متنی و تصاویر چهره مورد استفاده در کار این مقاله، در شکل ۶ نمایش داده شده‌اند.

از آن جا که روش پیشنهادی قابلیت ترکیب با روش‌های مختلف فشرده‌سازی را دارد، روش پیشنهادی را با سه روش JPEG، JPEG2000 و کدگذاری SPIHT مورد استفاده قرار داده تا عملکرد روش پیشنهادی با/بدون استفاده از این روش‌ها ارزیابی شود. کدگذاری SPIHT نوعی کدگذار بخش‌بندی مجموعه (SPC<sup>۳</sup>) است که ضرایب تبدیل موجک را فشرده‌سازی و کدگذاری می‌کند. کدگذارهای SPC عموماً از ویژگی وجود شباهت در زیرباندهای هم‌راستا در تجزیه‌ی هرمی موجک بهره می‌برند. عملکرد این روشها مبتنی بر روشی قانون‌مند برای جستجوی مهمترین ضرایب موجک و تصمیم‌گیری در مورد

سیگنال اصلی می‌شویم و لذا محتوای فرکانسی زیرباندهای جزئیات قوی و دارای انرژی بیشتری می‌شوند. این همان نقطه‌ای است که باید در روش پیشنهادی دیگر جلوتر نرویم تا عمل تقریب زیرباندهای جزئیات از روی اطلاعات زیرباند تقریب با دقت و موفقیت قابل قبولی انجام شود. می‌توانیم این نقطه را «نقطه بحرانی» بنامیم. از طرف دیگر هر قدر سیگنال ماهیت پایین‌گذرتر داشته باشد و نیز هر قدر درجه‌ی تفکیک مکانی تصویر بالاتر باشد، در عمل تجزیه، دیرتر به این نقطه (یا سطح تجزیه‌ی) بحرانی می‌رسیم. مشخص است که هر قدر دیرتر به این نقطه بحرانی برسیم عمل فشرده‌سازی موثرتر خواهد بود زیرا با هر بار تجزیه، مساحت (و یا اطلاعات لازم جهت فشرده‌سازی) یک چهارم می‌شود.

۴- می‌توان گفت حتی اطلاعات موجود در سطوح مختلف تجزیه اما مربوط به «یک زیرباند» نیز با یکدیگر تزیاید نسبتاً قابل توجهی دارند. اصولاً علت اصلی موفقیت کدگذارهای بخش‌بندی مجموعه (SPC) به ویژه کدگذاری SPIHT و یا کدگذاری EBCOT (که در استاندارد JPEG2000 استفاده شده است) وجود چنین تزیادهایی و استفاده از آن در جهت پیش‌بینی ضرایب موجک به منظور کاهش حجم اطلاعات ضروری جهت ذخیره‌سازی است.

۵- برای این که در انجام تخمین موفق‌تر عمل کنیم، ما برای هر نوع تصویر (مثلاً متنی یا چهره) عمل آموزش را جداگانه انجام می‌دهیم.

۶- در استدلالی که در بند ۲ مطرح شد، از مفاهیم «انرژی و شدت ناهم‌واری» و «راستا یا جهت زیرباندها» به عنوان مفاهیم کلیدی در روش پیشنهادی نام برده شد. به همین دلیل در روش پیشنهادی از ویژگی‌های واریانس (زیرباندهای حاصل از تجزیه‌ی زیرتصویر برگرفته از زیرباند تقریب) که مبین انرژی سیگنال است در جهت آموزش استفاده کرده‌ایم. همچنین ویژگی نسبت واریانس‌های افقی و عمودی نیز در جهت کمک به فرآیند آموزش و افزایش دقت طبقه‌بندی استفاده شده است. این ویژگی‌ها در حوزه‌ی بازبانی تصویر جزو ویژگی‌های مفید در نمایش و نمایندگی یک تصویر به جهت یافتن موارد مشابه می‌باشند.

۷- در ارتباط با این سوال که چگونه اطلاعات جهتی موجود در یک پیکسل از زیرباندی مانند H را می‌توان از آموزش واریانس یک زیرباند دیگر استخراج کرد، قابل ذکر است که تا جایی که مرتبط با اطلاعات جهتی یک زیرباند خاص (مثلاً H) است، اطلاعات واریانس یک زیرباند دیگر «ممکن است» بی‌تاثیر و بی‌اهمیت باشد. این که آیا اصولاً رابطه‌ای بین این دو وجود دارد یا ندارد را خود طبقه‌بند در حین فرآیند آموزش یاد خواهد گرفت. بنابراین، تا جایی که مربوط به یادگیری اطلاعات زیرباند H می‌باشد، ممکن است از برخی اطلاعات تهیه شده در بردارهای ویژگی استفاده شده و از بقیه استفاده نشود، اما برای یک زیرباند دیگر (مثلاً V) همین اطلاعات که قبلاً مفید نبود، حالا کاملاً

<sup>۱</sup>MIT Center for Biological and Computation Learning

<sup>۲</sup>در محیط MATLAB با دستوراتی نظیر imresize که مبتنی بر تکنیک‌های درون‌بانی عمل می‌کنند، می‌توان ابعاد یک تصویر را تغییر داد.

<sup>۳</sup>Set Partitioning Coder

تصاویر برای استفاده در مرحله‌ی آموزش استفاده شده است. این ۱۰ درصد به صورت تصادفی انتخاب نشده‌اند بلکه تصاویر آموزشی به گونه‌ای انتخاب شده‌اند که بیشترین تنوع را در بین خود داشته باشند؛ این کار به این دلیل انجام شده است که موجب تولید داده‌های آموزشی مناسب جهت نمایندگی دسته‌ی خود شود؛ به عبارت دیگر، داده‌های آموزشی تا حد امکان به بهترین نحو نمایش‌گر شکل فضای دسته‌ی خود باشند.

#### ۴-۱ مقادیر عددی پارامترها

روش پیشنهادی مستلزم بهره‌گیری از چند پارامتر و تنظیم مناسب آنها است. پارامترهای تاثیرگذار بر روش پیشنهادی به همراه مقدار عددی آنها و نیز روش محاسبه‌ی این مقادیر در جدول ۴-۱ آمده است.

جدول ۴-۱: پارامترهای موثر بر عملکرد روش پیشنهادی و مقادیر عددی آنها

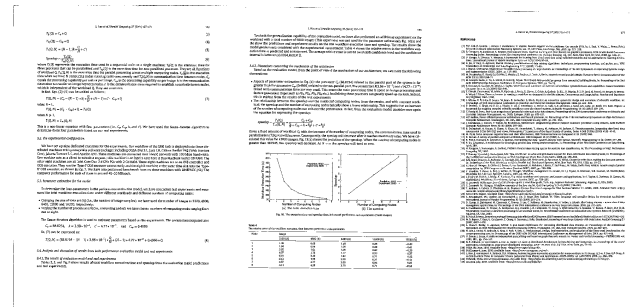
پارامتر	مقدار عددی	بخش مربوطه	توضیحات
L	-	۱-۳	مقدار این پارامتر به طور خودکار توسط الگوریتم پیشنهادی (مرتبط با روابط ۱ و ۲) تعیین می‌شود.
$\alpha$	۰/۱	۱-۳	هر قدر مقدار این ضریب بزرگتر انتخاب شود میزان فشرده‌سازی بیشتر اما میزان اعوجاج هم بیشتر می‌شود؛ همچنین، مقادیر بزرگتر موجب کاهش دقت تخمین زیرباندهای جزئیات می‌شود.
B	۱۷	۲-۳ و ۳-۳	مقدار این پارامتر متناسب با طول پاسخ ضربه‌ی فیلترهای موجک مورد استفاده در ساختار فیلتر بانک دارد؛ هرچه طول این فیلترها بزرگتر باشد، مقدار این پارامتر نیز بزرگتر انتخاب می‌شود. موجک مورد استفاده در این مقاله، bior4.4 می‌باشد که خواص خوبی در فشرده‌سازی دارد.
N	۱۶	۳-۳	تعداد نمایندگان (N) می‌بایست نه خیلی بزرگ و نه خیلی کوچک انتخاب شود؛ زیرا مقادیر کوچک منجر به کاهش دقت تخمین و مقادیر بزرگ نیز منجر به افزایش حجم محاسبات می‌شود.
	۰/۱	۳-۳	اگر مقدار این پارامتر خیلی کوچک‌تر انتخاب شود احتمال این که نمایندگان به خوبی تمام نقاط سطح محیطی دسته‌ی خود را پوشش دهند کمتر می‌شود؛ در مقابل اگر مقدار این پارامتر خیلی بزرگتر انتخاب شود، ممکن است قبل از این که تعداد نمایندگان به N برسد، دیگر هیچ‌گاه شرط (۶) برقرار نشود.
K	۲	۴-۳	هر قدر مقدار این پارامتر بزرگتر انتخاب شود به معنای آن است که در مورد تصمیم‌گیری تعلق بردار ورودی، $F(i,j)$ حساسیت بیشتری اعمال کرده‌ایم.

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 [۵۹].

معمولاً نتایج این جستجو به کمک کدگذارهایی مانند کدگذار حسابی<sup>۱</sup>، کدگذاری شده تا دنباله‌ی بیتی خروجی تولید شود. از جمله مهمترین روشهای کدگذاری SPC، روش کدگذاری SPIHT است. این کدگذار یکی از بهترین کدگذارهای SPC است؛ کارایی کدگذاری آن در حدود کارایی کدگذاری EBCOT<sup>۲</sup> (مورد استفاده در استاندارد JPEG2000) بوده و در عین حال، پیچیدگی محاسباتی آن کمتر است [۵۹]. تاکنون از این کدگذار برای تصاویر مختلفی مانند تصاویر هوایی، تصاویر پزشکی، تصاویر اثرانگشت و تصاویر عمومی استفاده شده است [۶۱-۵۹] اما طبق دانسته‌ی ما، در مورد تصاویر متنی و چهره استفاده‌ی چندانی از آن نشده است.



الف



ب

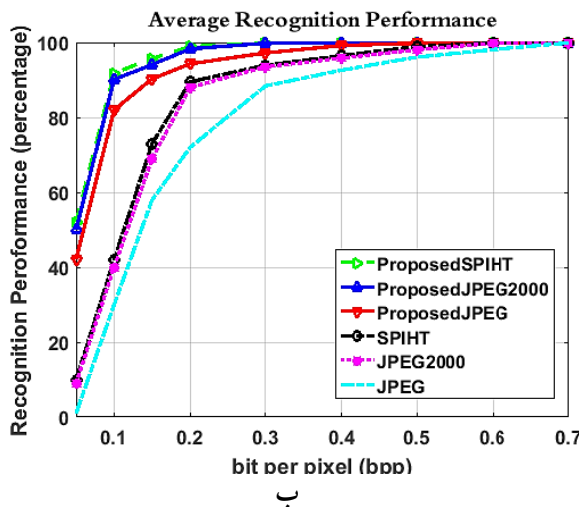
شکل ۶: چند نمونه از (الف) تصاویر چهره و (ب) تصاویر متنی، مورد استفاده در کار ارزیابی کارایی فشرده‌سازی روش پیشنهادی.

دومین رویکرد ارزیابی روش پیشنهادی، مقایسه‌ی حالت استفاده از روش پیشنهادی در ترکیب با روش‌های سه‌گانه‌ی نامبرده شده روی تصاویر درجه‌ی تفکیک مکانی بالا با حالت استفاده از روش‌های سه‌گانه به تنهایی روی تصاویر با درجه‌ی تفکیک مکانی متداول است. این مقایسه برطبق دو معیار PSNR و صحت بازشناسی انجام خواهد گرفت. برای انجام بازشناسی روی تصاویر متنی از نرم افزار ABBYY Fine Reader [۶۲] و برای انجام بازشناسی تصاویر چهره از نرم افزار VeriLook [۶۳] استفاده شده است. از آنجا که روش پیشنهادی قبل از استفاده، نیاز به آموزش و تشکیل مجموعه‌ی طبقات دارد، برای هر یک از انواع تصاویر مورد استفاده در این مقاله، تنها از حدود ۱۰ درصد

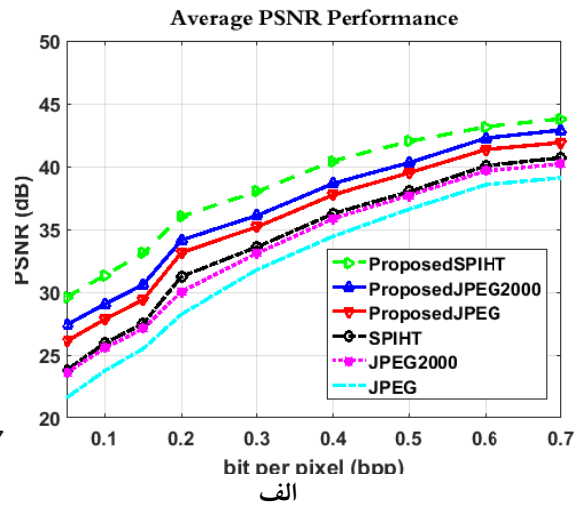
<sup>1</sup> Arithmetic Coding

<sup>2</sup> Embedded Block Coding with Optimized Truncation

PSNR حاصل از فشرده‌سازی تصاویر موجود در پایگاه تصاویر متنی به دست می‌آید. در سومین مرحله از ارزیابی، مجدداً روش‌های سه‌گانه‌ی نامبرده شده را (به تنهایی) روی تصاویر متنی با درجات تفکیک متداول (پایین) و ترکیب آنها با روش پیشنهادی را روی تصاویر با درجات تفکیک بالا اعمال می‌کنیم. حال صحت بازشناسی هر تصویر فشرده‌شده در هر گروه را محاسبه و ثبت می‌کنیم تا منحنی «متوسط صحت بازشناسی» را برحسب نرخ بیتی که فشرده‌سازی در آن انجام شده است، به دست آوریم. این منحنی در شکل ۷-۷ ب نمایش داده شده است.



۴-۲ ارزیابی کارایی فشرده‌سازی روی تصاویر متنی در اولین مرحله از ارزیابی کارایی فشرده‌سازی روش پیشنهادی روی تصاویر متنی، هر یک از روش‌های سه‌گانه‌ی JPEG، JPEG2000 و ترکیب تبدیل موجک و کدگذاری SPIHT را یک بار به صورت جداگانه و یک بار در ترکیب با روش پیشنهادی روی تصاویر متنی با درجه‌ی تفکیک مکانی بالا اعمال می‌کنیم تا کارآمدی روش پیشنهادی در استفاده از تزیاید موجود در درجات تفکیک بالا ارزیابی شود. معیار مورد استفاده در این مرحله، PSNR بوده و منحنی «متوسط PSNR» برحسب نرخ بیت محاسبه و نمایش داده می‌شود. این منحنی که در شکل ۷-۸ الف نمایش داده شده است، از طریق محاسبه‌ی متوسط مقادیر



شکل ۷: (الف) منحنی «متوسط PSNR» بر حسب نرخ بیت فشرده‌سازی برای دو گروه از روش‌ها و در درجه‌ی تفکیک مکانی بالا، (ب) منحنی «متوسط صحت بازشناسی» برحسب نرخ بیت فشرده‌سازی برای همان دو گروه از روش‌ها و همان دو درجه‌ی تفکیک مکانی مختلف.

در ارزیابی دوم از منحنی «صحت بازشناسی» برحسب نرخ بیتی که تصاویر در آن نرخ فشرده شده‌اند، استفاده می‌کنیم. مشابه با ارزیابی قبلی، در این جا نیز روش‌های سه‌گانه‌ی فشرده‌سازی روی تصاویر با درجه‌ی تفکیک مکانی متداول (پایین) و ترکیب این روش‌ها با روش پیشنهادی روی تصاویر چهره با درجه‌ی تفکیک مکانی بالا اعمال شده‌اند.

در کاربرد تصدیق چهره، معیار صحت بازشناسی به صورت زیر تعریف می‌شود:

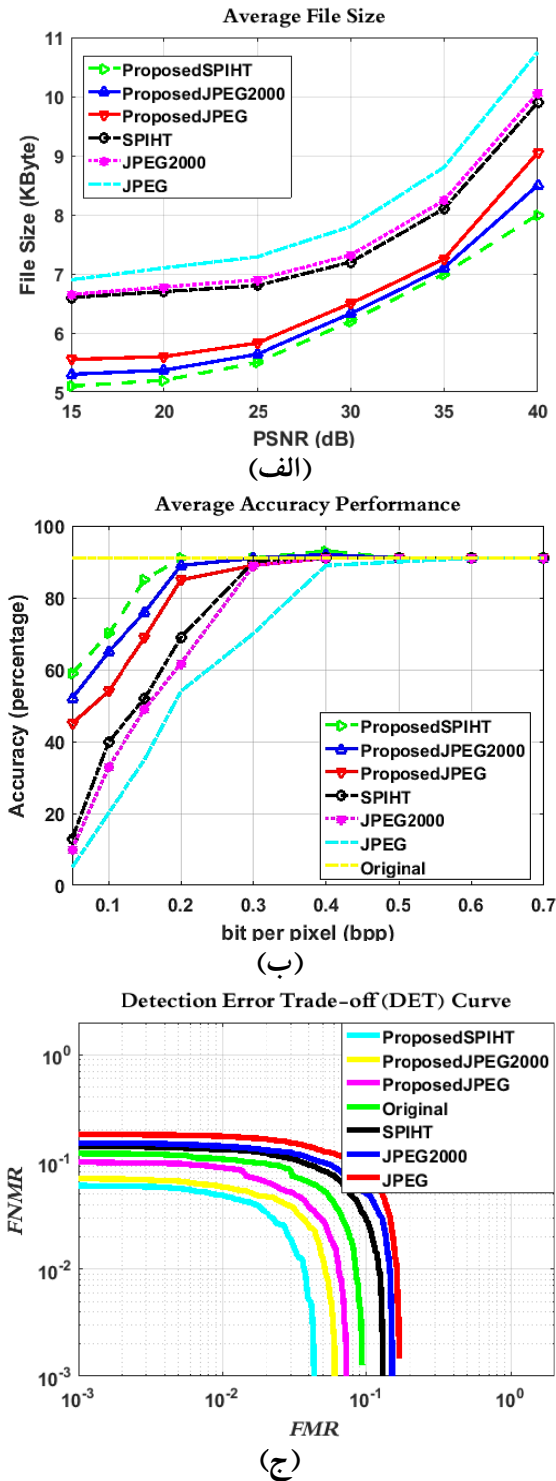
$$\text{Accuracy} = \frac{\sum \text{true positive} + \sum \text{true negative}}{\sum \text{total population}} \quad (9)$$

که در آن منظور از «مثبت صحیح» تعداد تصاویری است که به دسته‌ی «مثبت» تعلق داشته و طبقه‌بند نیز به درستی این تعلق را اعلام کرده است. منظور از «منفی صحیح» نیز تعداد تصاویری است که به دسته‌ی «منفی» تعلق داشته و طبقه‌بند به درستی این تعلق را اعلام کرده است. مخرج کسر فوق نیز تعداد کل تصاویری است که عمل تصدیق در مورد آنها انجام شده است. منحنی‌های صحت بازشناسی، متناظر با هر روش فشرده‌سازی و

از مجموع ارزیابی‌های انجام شده در روشی برتری روش پیشنهادی در رسیدن همزمان به اهداف کاهش حجم فایل فشرده شده و افزایش کارایی بازشناسی متن مشخص می‌شود. یکی از مهمترین دلایل چنین موفقیتی، یکی مقادیر نسبتاً بزرگ L و دیگری توانایی روش پیشنهادی در تخمین زیرباندهای جزئیات می‌باشد.

#### ۴-۳ ارزیابی کارایی فشرده‌سازی روی تصاویر چهره

با توجه به اهمیت کاربردهای ذخیره‌سازی و بازشناسی چهره، در ارزیابی‌های انجام شده در این بخش روی این دو کاربرد توجه شده است. در ارزیابی اول، مشابه با ارزیابی دوم در بخش قبلی عمل می‌کنیم تا منحنی «متوسط حجم فایل فشرده‌شده» برحسب PSNR به دست آید. مجدداً تاکید می‌شود که در این جا روش‌های سه‌گانه‌ی فشرده‌سازی روی تصاویر با درجه‌ی تفکیک مکانی متداول (پایین) و ترکیب این روش‌ها با روش پیشنهادی روی تصاویر چهره با درجه‌ی تفکیک مکانی بالا اعمال شده‌اند. این منحنی‌ها در شکل ۸-۸ الف نمایش داده شده‌اند.



شکل ۸: (الف) منحنی‌های «متوسط حجم فایل» برحسب مقدار مشخصی از PSNR برای دو گروه از روش‌ها و هر گروه در یک درجه‌ی تفکیک مکانی مشخص، (ب) منحنی‌های «صحت بازشناسی» برحسب نرخ بیت فشرده‌سازی برای همان دو گروه از روش‌ها و همان دو درجه‌ی تفکیک مکانی مختلف، (ج) منحنی‌های «بده-بستان خطای آشکارسازی» برای همان دو گروه از روش‌های فشرده‌سازی و در حالتی که تصاویر چهره با نرخ بیت 0.3 فشرده شده‌اند

نیز در حالت استفاده از تصاویر اصلی (فشرده‌نشده)، در شکل ۸-ب نمایش داده شده‌اند.

در سومین ارزیابی، تصاویر را در نرخ بیت مشخصی (در اینجا  $\text{bpp}=0.3$ ) فشرده کرده و منحنی بده-بستان خطای آشکارسازی ( $\text{DET}^1$ ) را محاسبه می‌کنیم. این منحنی تغییرات کمیت «اعلام اشتباه عدم انطباق» یا  $\text{FNMR}^2$  را بر حسب «اعلام اشتباه انطباق» یا  $\text{FMR}$  وقتی که مقدار آستانه‌ی مورد استفاده (در هر روش طبقه‌بندی) تغییر می‌کند، نمایش می‌دهد. منحنی  $\text{DET}$  یکی از منحنی‌هایی است که در کاربردهای بازشناسی چهره مورد استفاده قرار می‌گیرد. منحنی دیگری که ممکن است مورد استفاده قرار بگیرد، منحنی  $\text{ROC}^3$  است که تغییرات «نرخ اعلام صحیح انطباق» را بر حسب «نرخ اعلام اشتباه انطباق» نمایش می‌دهد. مشابه با ارزیابی‌های قبلی انجام شده در این بخش، روش‌های سه‌گانه‌ی فشرده‌سازی روی تصاویر با درجه‌ی تفکیک مکانی متداول (پایین) و ترکیب این روش‌ها با روش پیشنهادی روی تصاویر چهره با درجه‌ی تفکیک مکانی بالا اعمال شده‌اند. منحنی‌های  $\text{DET}$  برای هر یک از روش‌های فشرده‌سازی و نیز برای حالت استفاده از تصاویر اصلی چهره در شکل ۸-ج نمایش داده شده است.

از ارزیابی‌های انجام شده در این بخش نیز کارایی روش پیشنهادی در رسیدن همزمان به اهداف کاهش حجم فایل فشرده شده و افزایش کارایی تصدیق چهره مشخص می‌شود.

از مجموع ارزیابی‌های انجام شده در دو بخش اخیر و نیز با توجه به نتایجی که در حین آزمایشات به دست آمد چند نکته‌ی مهم قابل توجه و ذکر است. نکته اول این است که درجه‌ی تفکیک مکانی تاثیر قابل توجهی روی کارایی بازشناسی تصاویر دارد. نکته دوم این است که عامل فشرده‌سازی، نسبت به عامل کاهش درجه‌ی تفکیک مکانی، چندان موجب افت کارایی بازشناسی نمی‌شود. حتی در برخی موارد مشاهده گردید که انجام فشرده‌سازی موجب افزایش اندکی در کارایی بازشناسی شده است. این مطلب در دو منحنی مربوط به ترکیب روش پیشنهادی با روش‌های SPIHT و JPEG2000 در شکل ۸-ب قابل مشاهده است. این دو منحنی در نرخ بیت حدود 0.4  $\text{bpp}$  یک افزایش نسبی در کارایی بازشناسی نشان می‌دهند. از این رفتار یک نکته‌ی دیگر نیز نتیجه گرفته می‌شود که انجام فشرده‌سازی در نرخ بیت پایین‌تر لزوماً منجر به کاهش یا ثابت ماندن کارایی بازشناسی نمی‌شود؛ به عبارت دیگر، ممکن است با کاهش نرخ فشرده‌سازی کارایی بازشناسی اندکی افزایش یابد!

<sup>1</sup> Detection Error Trade-off

<sup>2</sup> False Non-Match rate

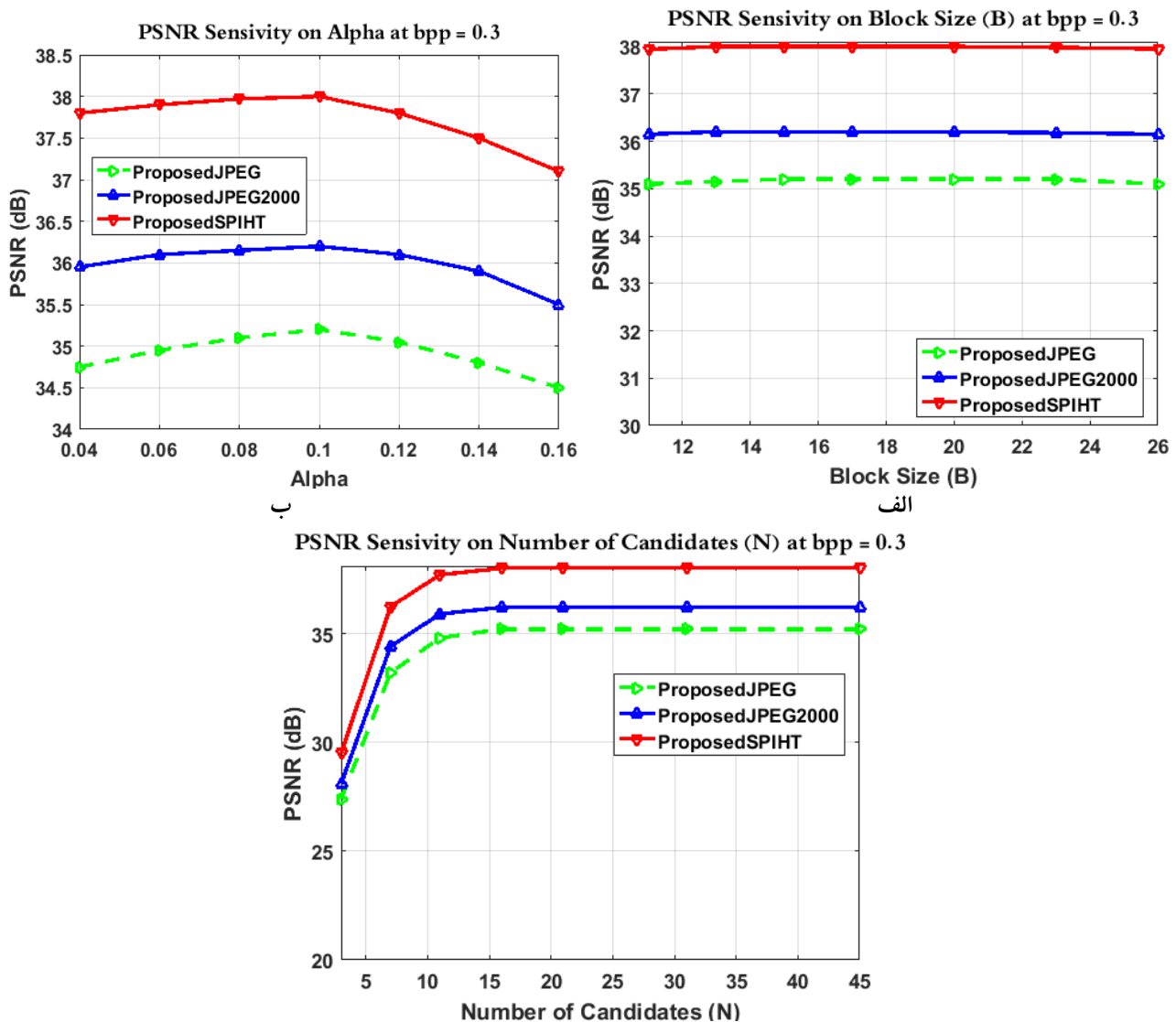
<sup>3</sup> Receiver Operating Characteristics

آزمایش‌هایی و با در نظر گرفتن اطلاعاتی مانند آن چه در جدول ۴-۱ آمده است، تعیین شده‌اند. تعیین میزان حساسیت روش پیشنهادی به مقادیر این پارامترها از اهمیت خاصی برخوردار است. در این بخش حساسیت کارایی الگوریتم پیشنهادی به برخی پارامترهای مهم که به صورت تجربی تعیین شده‌اند، بررسی می‌شود. در این کار ابتدا تمامی پارامترها طبق مقادیر تعیین شده در جدول ۴-۱ تنظیم می‌شوند. حال پارامتر مورد نظر را مقدار خود تغییر داده و هر بار مقدار (متوسط) معیار مورد نظر را محاسبه کرده‌ایم. برای تصاویر متنی معیار PSNR را در نرخ بیت 0.3 و برای تصاویر چهره معیار سحت را در نرخ بیت ثابت 0.2 در نظر گرفته‌ایم. از میان پارامترهای معرفی شده در جدول ۴-۱، سه پارامتر مهم  $\alpha$ ، B و N جهت بررسی حساسیت انتخاب شده‌اند. نتایج مربوط به تصاویر متنی در شکل ۹ و نتایج مربوط به تصاویر چهره در شکل ۱۰ نمایش داده شده‌اند.

به نظر علت اصلی این امر مربوط به اثر فشرده‌سازی در حذف برخی اطلاعات گمراه‌کننده می‌باشد. آخرین نکته، با توجه به نمودارهای سحت بازشناسی در دو بخش تصاویر متنی و تصاویر چهره، این است که این منحنی‌ها در حالت کلی با کاهش نرخ بیت دارای رفتاری تقریباً ثابت (با شیب نزدیک به صفر) بوده و در یک نرخ بیت (به نام نقطه‌ی شکست) با شیب ناگهانی و تند شروع به افت می‌کنند. هر قدر نقطه‌ی شکست کوچکتر باشد، آن روش فشرده‌سازی کارتر خواهد بود. تاثیر روش‌های پیشنهادی، در مقایسه با روش‌های سه‌گانه‌ی مورد بررسی، روی این رفتار این است که نقطه‌ی شکست را کاهش داده‌اند و این امر یک مزیت برای روش پیشنهادی محسوب می‌شود.

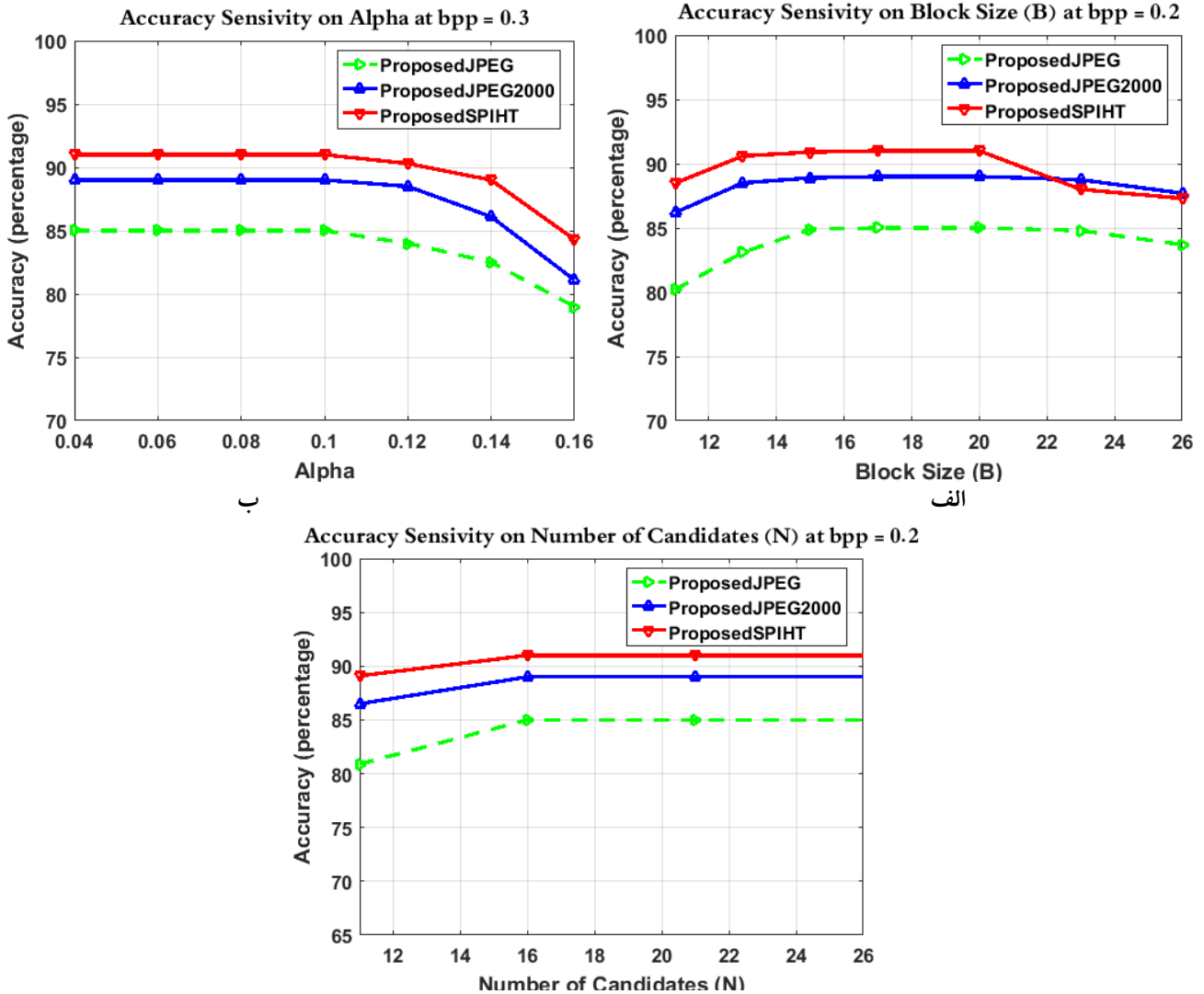
#### ۴-۴ حساسیت روش پیشنهادی به مقادیر پارامترها

در روش پیشنهادی برخی پارامترها مانند T و L به صورت وقتی تعیین می‌شوند. برخی دیگر نیز به صورت تجربی و با انجام



ج

شکل ۹: (الف) حساسیت معیار PSNR به پارامتر آلفا ( $\alpha$ )، (ب) حساسیت معیار PSNR به پارامتر B، (ج) حساسیت معیار PSNR به پارامتر N. تمامی منحنی‌ها در نرخ بیت ثابت 0.3 و برای تصاویر متنی محاسبه شده‌اند.



شکل ۱۰: (الف) حساسیت معیار صحت بازشناسی به پارامتر آلفا ( $\alpha$ )، (ب) حساسیت معیار صحت بازشناسی به پارامتر B، (ج) حساسیت معیار صحت بازشناسی به پارامتر N. تمامی منحنی‌ها در نرخ بیت ثابت 0.2 bpp و برای تصاویر چهره محاسبه شده‌اند.

قبل مشخص باشد تا متناسب با نوع تصویر ورودی، از داده‌های آموزشی مناسب استفاده کند.

نتایج شبیه‌سازی حاصل از ارزیابی کارایی فشرده‌سازی ترکیب روش پیشنهادی با هر یک از سه روش JPEG، JPEG2000 و کدگذاری SPIHT، بر طبق معیار منحنی «متوسط PSNR» بر حسب نرخ بیت، نشان دهنده‌ی قابلیت قابل توجه روش پیشنهادی در حذف تزیاید ناشی از درجه‌ی تفکیک مکانی بالا بود. همچنین ارزیابی کارایی بازشناسی روش پیشنهادی، بر طبق معیار منحنی «صحت بازشناسی» بر حسب نرخ بیت، در ترکیب با روش‌های سه‌گانه نامبرده شده نشان دهنده‌ی قابلیت این روش در حفظ ویژگی‌های مهم و تاثیرگذار در بازشناسی بوده است.

## مراجع

- [1] W. A. Pearlman and A. Said, "Set Partitioning Coding: Part I of Set Partition Coding and Image Wavelet Coding

## نتیجه‌گیری

با توجه به اهمیت درجه‌ی تفکیک بالا در افزایش کیفیت، خوانایی و یا خوشایندی تصاویر، در این مقاله روشی جهت فشرده‌سازی تصاویر دارای درجه‌ی تفکیک بالا مبتنی بر کاهش/افزایش ابعاد و تخمین زیرباندهای جزئیات در حوزه‌ی تبدیل موجک پیشنهاد شد. جهت تخمین زیرباندهای جزئیات، نیاز به یک مرحله آموزش با تصاویر مناسب می‌باشد. تصاویر مورد نظر می‌توانند از هر نوعی، مانند تصاویر متنی، چهره، اثرانگشت و طبیعی باشند. البته در این مقاله، دو نوع متنی و چهره مورد بررسی و ارزیابی قرار گرفتند. روش پیشنهادی دارای قابلیت ترکیب با هر روش فشرده‌سازی استاندارد/غیر استاندارد موجود بوده و قادر به کاهش قابل توجه تزیاید ناشی از درجه‌ی تفکیک مکانی بالا می‌باشد. یک نقطه‌ی ضعف روش پیشنهادی این است که برای داشتن بهترین کارایی، می‌بایست نوع تصاویر از



- [14] Z. Fan and T. Jacobs, "Segmentation for Mixed Raster Contents with Multiple Extracted Constant Color Areas," Proc. of SPIE-IS&T Electronic Imaging, Vol. 5667, pp. 251-262, 2005.
- [15] D. P. Huttenlocher, P. F. Felzenswalb, W. Rucklidge, "DigiPaper: a Versatile Color Document Image Representation," International Conference on Image Processing (ICIP), pp. 219-223, 1999.
- [16] H. Grailu, M. Lotfizad, and H. S. Yazdi, "1-D Chaincode Pattern Matching for Compression of Bi-level Printed Farsi and Arabic Textual Images," Image and Vision Computing (IVC), Vol. 27, No. 10, pp. 1615-1625, 2009.
- [17] H. Grailu, M. Lotfizad, and H. S. Yazdi, "An Improved Pattern Matching Technique for Lossy/Lossless Compression of Binary Printed Farsi and Arabic Textual Images," International Journal of Intelligent Computing and Cybernetics, Vol. 2, No. 1, pp. 120-147, 2009.
- [18] H. Grailu, M. Lotfizad, and H. S. Yazdi, "A Lossy/Lossless Compression Method for Printed Typeset Bi-level Text Images Based on Improved Pattern Matching," International Journal on Document Analysis and Recognition (IJ DAR), Vol. 11, No. 4, pp. 159-182, 2009.
- [19] A. Awajan, "Multilayer Model for Arabic Text Compression," The International Arab Journal of Information Technology, Vol. 8, No. 2, pp. 188-196, 2011.
- [20] L. Bottou, P. Haffner, P. G. Howard, P. Simard, Y. Bengio, and Y. LeCun, "High Quality Document Image Compression with DjVu," Journal of Electronic Imaging, Vol. 7, No. 3, pp. 410-425, 1998.
- [21] D. Huttenlocher, P. Felzenswalb and W. Rucklidge, "DigiPaper: A Versatile Color Document Image Representation," Proc. of IEEE Int'l Conf. on Image Proc., pp. 219-223, Kobe, Japan, October 1999.
- [22] K. U. Barthel, S. M. Partlin, and M. Thierschmann, "New Technology for Raster Document Image Compression," Part of the IS&T/SPIE Conference on Document Recognition and Retrieval VII, San Jose, CA, Vol. 3967, pp. 286-290, 2000.
- [23] M. Thierschmann, K. -U. Barthel, and U. -E. Martin, "A Scalable DSP-Architecture For High-Speed Color Document Compression," Document Recognition and Retrieval VIII, Paul B. Kantor, Daniel Lopresti P., Jiangying Z., Editors, Proceedings of SPIE, Vol. 4307, pp. 158-166, 2001.
- [24] B. -F. Wu, C. -C. Chiu, and Y. -L. Chen, "Algorithms for Compressing Compound Document Images with Large Text/Background Overlap," IEE Proc.-Vis. Image Signal Process., Vol. 151, No. 6, pp. 453-459, 2004.
- [26] S. Mohan and S. S. Kumar (eds.), Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012), Lecture Notes in Electrical Engineering 221, Springer India, pp. 59-68, 2013.
- Systems," Foundations and Trends in Signal Processing, Vol. 2, No. 2, pp. 95-180, 2008.
- [2] 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.
- [3] A. Said and W. 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.
- [4] G. Feng and C. A. Bouman, "High-Quality MRC Document Coding," IEEE Trans. Image Proc., Vol. 15, No. 10, pp. 3152-3169, 2006.
- [5] 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.
- [6] 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.
- [7] L. H. Sharpe and B. Manns, "JPEG2000 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.
- [8] 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.
- [9] 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.
- [10] A. Zaghetto and R. L. de Queiroz, "Scanned Document Compression Using a Block-Based Hybrid Video Codec," IEEE Transactions on Image Processing, Vol. 22, No. 6, pp. 2420-2428, 2010.
- [11] E. Y. Lam, "Compound Document Compression with Model-Based Biased Reconstruction," Journal of Electronic Imaging (JEI), Vol. 13, No. 1, pp. 191-197, 2004.
- [12] G. Feng, C. A. Bouman, and H. Cheng, "High Quality MRC Document Coding," Image Processing, Image Quality, Image Capture Systems Conference, pp. 320-325, 2001.
- [13] R. L. de Queiroz., R. R. Buckley and M. Xu, "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. 1106-1117, 1999.

- [40] J. E. Vila-Forcen, S. Voloshynovskiy, O. Koval, and T. Pun, "Facial Image Compression Based on Structured Codebooks in Overcomplete Domain," *EURASIP Journal on Applied Signal Processing*, Vol. 2006, No. 69042, pp. 1–11, 2006.
- [41] Sohrab Ferdowsi, Sviatoslav Voloshynovskiy, Dimche Kostadinov, Marcin Korytkowski, and Rafal Scherer, "Secure Representation of Images Using Multi-layer Compression," *Proceedings of 14th International Conference, ICAISC 2015, Part I*, pp. 696–705, Zakopane, Poland, June 14–18, 2015.
- [42] Ori Bryt and Michael Elad, "Compression of facial images using the K-SVD algorithm," *Journal of Visual Communication and Image Representation*, Vol. 19, pp. 270–282, 2008.
- [43] Ori Bryt and Michael Elad, "Improving the K-SVD Facial Image Compression Using a Linear Deblocking Method," *IEEE 25th Convention of Electrical and Electronics Engineers in Israel*, pp. 533 – 537 , 3–5 Dec. 2008.
- [44] Ding Xing hao, Qian Kun, Xiao Quan, Liao Ying, hao, Guo Dong hui, and Wang Shou jue, "Low Bit Rate Compression of Facial Images Based on Adaptive Over-complete Sparse Representation," *2nd International Congress on Image and Signal Processing (CISP '09)*, pp. 1 – 3 , 17–19 Oct. 2009.
- [45] Idan Ram, Israel Cohen, and Michael Elad, "Facial Image Compression using Patch-Ordering-Based Adaptive Wavelet Transform," *IEEE Signal Processing Letters*, Vol. 21, No. 10, pp. 1270–1274, 2014.
- [46] Wolfgang Funk, Michael Arnold, Christoph Busch, Axel Munde, "Evaluation of Image Compression Algorithms for Fingerprint and Face Recognition Systems," *Proceedings from the Sixth Annual IEEE SMC Information Assurance Workshop*, pp. 72 – 78, 15–17 June 2005.
- [47] J. N. Bradley, C. M. Brislawn, and T. Hopper, "The FBI Wavelet/Scalar Quantization Standard for Gray-Scale Fingerprint Image Compression," *SPIE Proceedings, Visual Information Processing II*, 1961, pp. 293–304, Apr. 1993.
- [48] Kresimir Delac, Mislav Grgic and Sonja Grgic, *Face recognition*, Chapter 5: "Image Compression Effects in Face Recognition Systems," I-Tech Education and Publishing, 2007.
- [49] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600–612, 2004.
- [50] Peilong Zhao, Jiwen Dong, and Hengjian Li, "Effect of Severe Image Compression on Face Recognition Algorithms," *Proceedings of SPIE: Image Processing and Analysis*, Vol. 9675, 96750E, 8 October 2015.
- [51] ITU-T and ISO/IEC, *JPEG XR Image Coding System — Part 2 : Image Coding Specification*. ISO/IEC Standard 29199-2, 2011.
- [27] M. Barni (Editor), *Document and Image Compression*, CRC Press, Chapter 13: "Compressing Compound Documents" by R. L. de Queiroz , 2006.
- [28] K. Hu, Z. Tang, L. Gao, and Y. Mu, "MC-JBIG2: an Improved Algorithm for Chinese Textual Image Compression," *International Journal on Document Analysis and Recognition (IJ DAR)*, Vol. 13, No. 4, pp. 271–284, 2010.
- [29] R. L. de Queiroz, Z. Fan, and T. D. Tran, "Optimizing Block-Thresholding Segmentation for Multilayer Compression of Compound Images," *IEEE Trans. Image Proc.*, Vol. 9, No. 9, pp. 1461–1471, 2000.
- [30] A. Mascher-Kampfer, Herbert Stogner, and Andreas Uhl, "Comparison of Compression Algorithms' Impact on Fingerprint and Face Recognition Accuracy," *Proc. SPIE 6508, Visual Communications and Image Processing 2007*, 650810, 12 pages, 29 January 2007.
- [31] Baback Moghaddam and Alex Pentland, "An Automatic System for Model-Based Coding of Faces", *IEEE Data Compression Conference (DCC)*, pp. 362–370, 1995.
- [32] Aysegul Gunduz, Humid Krim, and P. Allan Sadowski, "Compression and transmission of facial images over very narrowband wireless channels," *Processing of IEEE International Conference on Acoustics, Speech (ICASSP '03)*, Vol. 5, pp. V-704–7, 6–10 April 2003.
- [33] Alexander Tropf and Douglas Chai, "Region Segmentation for Facial Image Compression," *5th International Conference on Information Communications & Signal Processing*, pp. 1556 – 1560, Bangkok, 2005.
- [34] Jun Hou, Wen Chen, and Jing Yang, "Formal Photograph Compression In JPEG2000 Coding," *International Symposium on Intelligence Information Processing and Trusted Computing*, pp. 491–494, 2010.
- [35] Jun Hou, Ran Li, Yan Cheng, Haojie Shi, "Passport Photo Compression Technique with JPEG2000," *IEEE International Conference on Mechatronics and Automation*, pp. 1389 – 1393, 4–7 Aug. 2013.
- [36] Delia P. McGarry, Craig M. Arndt, Steven A. McCabe and Donald P. D'Amato, "Effects of compression and individual variability on face recognition performance," *Proceeding of SPIE, Biometric Technology for Human Identification*, 25 August 2004.
- [37] Jian-Hong Hu, Ru-Shang Wang, and Yao Wang, "Compression of Personal Identification Pictures Using Vector Quantization with Facial Feature Correction," *Optical Engineering*, Vol. 35, No. 1, pp. 198– 203, 1996.
- [38] Kohei Inoue and Kiichi Urahama, "DSVD: A Tensor-Based Image Compression and Recognition Method", *IEEE International Symposium on Circuits and Systems*, Vol. 6, pp. 6308 – 6311, 23–26 May 2005.
- [39] Jincheng Huang and Yao Wang, "Compression of Color Facial Images Using Feature Correction Two-Stage Vector Quantization," *IEEE Transactions on Image Processing*, Vol. 8, No. 1, pp. 102–109, 1999.

- [52] D. G. Lowe, "Object Recognition from Local Scale-Invariant Features," The Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, pp. 1150-1157, 1999.
- [53] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [54] Stan Z. Li, and Anil K. Jain, Handbook of Face Recognition, 2<sup>nd</sup> edition, Springer-Verlag London, 2011.
- [55] Griffin, P., Face Recognition Format for Data Interchange," INCITS M1/02-0228.
- [56] Hans-Georg Stark, Wavelets and Signal Processing: An Application-Based Introduction, Springer-Verlag Berlin Heidelberg, 2005.
- [57] <http://cbcl.mit.edu>, last visited 10 August 2016.
- [58] [http://www.anejian.com/research/face\\_reco.htm](http://www.anejian.com/research/face_reco.htm), last visited 10 August 2016.
- [59] W. A. Pearlman and A. Said, Digital Signal Compression: Principles and Practice, Cambridge University Press, New York, 2011.
- [60] <http://www.cipr.rpi.edu/research/SPIHT/>, last visited 10 August 2016.
- [۶۱] هادی گرایلو، «فشرده‌سازی تصاویر اثرانگشت در نرخ بیت‌های پایین با رویکرد حفظ یا بهبود کارایی بازشناسی»، مجله‌ی ماشین‌بینایی و پردازش تصویر، سال دوم، شماره‌ی دو، زمستان ۱۳۹۴.
- [62] <https://www.abbyy.com/finereader/>, last visited 10 August 2016.
- [63] <http://www.neurotechnology.com/verilook.html>, last visited 10 August 2016.



هادی گرایلو تحصیلات خود را در مقطع کارشناسی در سال ۱۳۸۲ در دانشگاه صنعتی شریف به پایان برد. وی همچنین، تحصیلات خود در مقاطع کارشناسی ارشد و دکتری را در دانشگاه تربیت مدرس تهران در سالهای به ترتیب ۱۳۸۴ و ۱۳۸۸ به پایان رساند. زمینه‌های

تحقیقاتی ایشان در مقاطع تحصیلات تکمیلی شامل پردازش تصاویر متنی بوده است. وی هم‌اکنون عضو هیات علمی دانشگاه صنعتی شاهرود بوده و حوزه‌های تحقیقاتی مورد علاقه‌ی ایشان شامل پردازش تصاویر متنی، پردازش سیگنال‌های حیاتی، و پیاده‌سازی سخت‌افزاری الگوریتم‌های پردازش سیگنال در بستر سخت‌افزاری به ویژه پردازشگرهای سیگنال و FPGA می‌باشد.