





دانشگاه صنعتی شاهرود

دانشکده برق و رباتیک
گروه الکترونیک

فشرده سازی تصاویر چهره مبتنی بر نواحی مطلوب

محسن اکرمی ذکرا باد

اساتید راهنما:

دکتر هادی گرایلو

دکتر امیدرضا معروضی

استاد مشاور:

دکتر حسین مروی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

بهمن ۱۳۹۱



مدیریت تحصیلات تکمیلی
فرم شماره (۶)

شماره : / آ.ت.ب
تاریخ : ۹۱/۱۱/۱۷
ویرایش : -----

بسمه تعالی

فرم صورتجلسه دفاع پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای :

گرایش : الکترونیک

رشته : برق

محسن اکرمی ذکرآباد

تحت عنوان : فشرده سازی تصاویر چهره مبتنی بر نواحی مطلوب

که در تاریخ ۱۳۹۱/۱۱/۱۷ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح زیر است :

☐ مردود	☐ دفاع مجدد	☒ قبول (با درجه : <u>بسیار خوب</u> امتیاز <u>۱۸/۸۷</u>)
--------------------------------	------------------------------------	--

۲- بسیار خوب (۱۸ - ۱۸/۹۹)

۱- عالی (۱۹ - ۲۰)

۴- قابل قبول (۱۴ - ۱۵/۹۹)

۳- خوب (۱۶ - ۱۷/۹۹)

۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنما	هادی کریمو امیررضا سرزنده	استادیار استادیار	
۲- استاد مشاور	میلر	استادیار	
۳- نماینده شورای تحصیلات تکمیلی	ساسان ناصح	استادیار	
۴- استاد ممتحن	رضا زاهد	استادیار	
۵- استاد ممتحن	مصطفی حسینی	استادیار	

رئیس دانشکده :

تقدیم به:

پدر، مادر، همسر عزیزم

و به تمام آزاد مردانی که نیک می اندیشند و عقل و منطق را پیشه خود نموده و جز رضای الهی و
پیشرفت و سعادت جامعه، هدی ندارند.

دانشمندان، بزرگان، و جوانمردانی که جان و مال خود را در حفظ و اعتلای این مرز و بوم فدا
نموده و میمانند

سپاس بی کران پروردگار یکتا را که هستی مان بخشید و به طریق علم و دانش رهنمونمان شد و به همنشینی رهروان علم و دانش مفتخرمان نمود و خوشه چینی از علم و معرفت را روزیمان ساخت.

بدون شک جایگاه و منزلت معلم اجل از آن است که در مقام قدردانی از زحمات بی شائبه‌ی او، با زبان قاصر و دست ناتوان چیزی بنگاریم. اما بر حسب وظیفه بر خود واجب می‌دانم از پدر و مادر عزیزم... این دو معلم بزرگوارم... که همواره بر کوتاهی و درشتی من قلم عفو کشیده و کریمانه از کنار غفلت‌هایم گذشته‌اند و در تمام عرصه‌های زندگی یار و یآوری بی‌چشم داشت برای من بوده‌اند؛ از اساتید شایسته و فرهیخته، جناب آقای دکتر هادی گرایلو و آقای دکتر امیدرضا معروضی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این رساله را بر عهده گرفتند؛ از استاد گرامی و ارجمند، جناب آقای دکتر حسین مروی، که زحمت مشاوره این رساله را متقبل شدند؛ و نیز سایر اساتید فرهیخته و دلسوز دانشکده برق و رباتیک دانشگاه صنعتی شاهرود کمال تشکر و قدردانی را بجا آورم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

همچنین در انتها از همسر عزیزم، مادر بزرگ دلسوز و مهربانم و سایر اعضای خانواده که در دوران تحصیل همواره مشوق و پشتیبان من بوده‌اند، و نیز کلیه کسانی که به نوعی در موفقیت و پیشرفت من در زندگی نقش داشته‌اند کمال تشکر را دارم.

تعهد نامه

اینجانب دانشجوی دوره کارشناسی ارشد رشته مهندسی برق و ... دانشگاه صنعتی شاهرود

ریاست دانشگاه صنعتی شاهرود نویسنده پایان نامه با عنوان :

.....

موضوعی
تحت راهنمایی آقای دکتر می شوم :

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت های آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ: ۹۱/۱۱/۱۷

امضاء دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

* متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد .

چکیده

امروزه کاربرد سیستم‌های شناسایی چهره به منظور تشخیص هویت، کنترل تردد و حفاظت اطلاعات افزایش چشمگیری داشته است. این سیستم‌ها معمولاً نیازمند پایگاه بزرگی از تصاویر چهره هستند. محدودیت حافظه و افزایش سرعت بازیابی از جمله عواملی هستند که ما را به سوی فشرده‌سازی این تصاویر سوق می‌دهند. نکته مهمی که در فشرده‌سازی این تصاویر باید در نظر گرفته شود، حفظ کیفیت ویژگی‌های مهم چهره در حد مناسب، مخصوصاً در نرخ‌بیت‌های پایین است.

در این پایان‌نامه طی یک سیر تکاملی سه روش برای فشرده‌سازی تصاویر چهره مبتنی بر نواحی مطلوب در نرخ‌بیت‌های پایین پیشنهاد شده است. تبدیل موجک و کدگذاری WBTC پایه و اساس هر سه روش پیشنهادی را تشکیل می‌دهند. در روش‌های پیشنهادی با بازآرایی بلوک‌های ضرایب موجک بر حسب اهمیت، کارایی الگوریتم کدگذاری WBTC افزایش یافته است؛ چرا که در این حالت زمان و نرخ‌بیت کمتری صرف یافتن ضرایب مهم می‌گردد. در روش اول با محاسبه یک تصویر باقیمانده در حوزه موجک و کدگذاری نواحی مطلوب در این تصویر به کمک روش Maxshift، کیفیت دیداری این نواحی در نرخ‌بیت‌های پایین بهبود یافته است. در روش دوم با استفاده از تبدیل RWT، امکان فشرده‌سازی مجزای نواحی مطلوب و غیرمطلوب در نرخ‌بیت دلخواه فراهم شده است. در روش سوم با استفاده تلفیقی از تبدیل RWT و تکنیک کدگذاری تصویر باقیمانده در روش اول، به کارایی بالاتری نسبت به دو روش قبل دست یافته‌ایم. نتایج شبیه‌سازی بیانگر کارایی بالاتر روش‌های پیشنهادی در مقایسه با JPEG2000، SPIHT و یک روش مبتنی بر ناحیه مطلوب است. این کارایی بالا در ناحیه چهره و مخصوصاً در نرخ‌بیت‌های پایین‌تر خود را بهتر نشان می‌دهد، چرا که با کاهش نرخ‌بیت کارایی الگوریتم WBTC و تکنیک بازآرایی بلوک‌های ضرایب موجک افزایش می‌یابد.

کلمات کلیدی: تبدیل موجک، تصویر چهره، فشرده‌سازی تصویر، کدگذار WBTC، ناحیه مطلوب.

لیست مقالات مستخرج از پایان نامه

مقالات پذیرفته شده

[۱] اکرمی م.، گرایلو ه.، معروضی ا.ر. و سلطانی ر.، (۱۳۹۱)، "روشی کارا برای فشرده سازی تصاویر

چهره پرسنلی در نرخ بیت های پایین به منظور استفاده در تشخیص هویت بلادرنگ"، اولین

کنفرانس بازشناسی الگو و تحلیل تصویر ایران، بیرجند، ایران.

[۲] اکرمی م.، گرایلو ه. و معروضی ا.ر.، (۱۳۹۱)، "استفاده از تبدیل RWT و کدگذاری تصویر

باقیمانده در فشرده سازی تصاویر چهره پرسنلی به منظور بهبود کیفیت ویژگی های چهره در

نرخ بیت های پایین"، یازدهمین کنفرانس سیستم های هوشمند ایران، تهران، ایران.

[3] Grailu H., Soltani R. and Akrami M., (2013), "ROI Soccer Video Compression Using Wavelet-Domain Block Matching and EBCOT Coding Techniques In Very Low Bit Rate Communication", The First Iranian Conference on Pattern Recognition and Image Analysis, Birjand, Iran.

[4] Grailu H., Soltani R. and Akrami M., (2012), " Improving the Compression Efficiency of SPC Coders Based on Spectral Reordering and Block Matching Techniques", The 5th International Conference on Computer and Electrical Engineering, Hong Kong, China.

مقالات ارسال شده

[۱] اکرمی م.، گرایلو ه. و معروضی ا.ر.، (۱۳۹۲)، "روشی کارآمد برای فشرده سازی تصاویر چهره

پرسنلی در نرخ بیت های خیلی پایین براساس تبدیل موجک مبتنی بر ناحیه و بهبود کارایی

کدگذار WBTC"، ارائه شده به بیست و یکمین کنفرانس مهندسی برق ایران، مشهد، ایران.

چکیده.....ز

لیست مقالات مستخرج از پایان نامه.....ح

فصل اول: مقدمه.....۱

۱-۱ تعریف مسئله.....۲

۱-۱-۱ ضرورت فشرده سازی مبتنی بر ناحیه مطلوب تصاویر چهره.....۳

۲-۱ مروری بر کارهای گذشته.....۵

۱-۲-۱ روش های مبتنی بر فرهنگ داده.....۵

۲-۲-۱ روش های مبتنی بر تبدیل.....۹

۳-۱ انتخاب طرح کلی.....۱۲

۴-۱ اهداف پایان نامه.....۱۶

۵-۱ ساختار پایان نامه.....۲۳

فصل دوم: مفاهیم پایه.....۲۴

۱-۲ تبدیل موجک.....۲۵

۱-۱-۲ کدگذاری زیرباند.....۲۶

۲-۱-۲ تبدیل موجک.....۲۷

۳-۱-۲ تبدیل موجک گسسته.....۳۰

۳۳	۴-۱-۲ تبدیل موجک مبتنی بر ناحیه.....
۳۸	۱-۴-۱-۲ پیچیدگی محاسباتی تبدیل RWT.....
۳۹	۲-۲ کدگذاری تصویر مبتنی بر موجک.....
۴۱	۱-۲-۲ کدگذاری مبتنی بر صفحه بیتی ضرایب موجک.....
۴۴	۲-۲-۲ کدگذاری مبتنی بر درخت صفر.....
۴۸	۳-۲-۲ کدگذاری مبتنی بر بلوک صفر.....
۵۰	۴-۲-۲ کدگذاری درخت بلوکی موجک.....
۵۷	۳-۲ کدگذاری ROI در حوزه موجک.....
۵۸	۱-۳-۲ مقیاس گذاری کلی.....
۶۰	۲-۳-۲ انتقال ماکزیمم.....
۶۱	فصل سوم: روش های پیشنهادی.....
۶۲	۱-۳ روش پیشنهادی اول.....
۶۵	۱-۱-۳ فرآیند فشرده سازی.....
۶۸	۲-۱-۳ فرآیند بازسازی.....
۶۸	۲-۲ روش پیشنهادی دوم.....
۷۱	۱-۲-۳ فرآیند فشرده سازی.....
۷۲	۲-۲-۳ فرآیند بازسازی.....
۷۳	۳-۳ روش پیشنهادی سوم.....
۷۴	۱-۳-۳ فرآیند فشرده سازی.....

۷۶	۲-۳-۳ فرآیند بازسازی.....
۷۷	۴-۳ جداسازی نواحی مطلوب و غیر مطلوب.....
۸۱	۵-۳ کدگذاری مرز نواحی مطلوب.....
۸۶	۶-۳ بهبود کارایی کدگذاری WBTC.....
۸۹	۱-۶-۳ فاز شبه-آموزش.....
۹۲	فصل چهارم: نتایج شبیه‌سازی.....
۹۳	۱-۴ معرفی پایگاه داده.....
۹۴	۲-۴ معیارهای ارزیابی.....
۹۶	۳-۴ نتایج شبیه‌سازی.....
۱۱۴	فصل پنجم: نتیجه‌گیری.....
۱۱۵	۱-۵ نتیجه‌گیری.....
۱۱۸	۲-۵ پیشنهادات.....
۱۲۰	مراجع.....

شکل ۱-۱: طرح کلی مبتنی بر تبدیل.....	۱۸
شکل ۱-۲: بانکی از فیلترهای میان‌گذر.....	۲۷
شکل ۲-۲: تاثیر گسترش و انتقال زمانی روی موجک مادر.....	۲۹
شکل ۳-۲: تبدیل موجک گسسته دوبعدی یک تصویر چهره پرسنلی با سه سطح تجزیه.....	۳۳
شکل ۴-۲: گام‌های پردازشی تبدیل RWT.....	۳۶
شکل ۵-۲: تصویر فروشنده.....	۳۷
شکل ۶-۲: یک سیستم کدگذاری تصویر مبتنی بر کدگذاری صفحه بیتی ضرایب موجک.....	۴۱
شکل ۷-۲: توضیح نحوه نمایش، انتقال و چندی‌سازی ضرایب در کدگذاری صفحه بیتی ضرایب موجک.....	۴۲
شکل ۸-۲: درخت‌های جهت‌دار مکانی و بخش‌بندی مجموعه در SPIHT.....	۴۷
شکل ۹-۲: نمایش بخش‌بندی درخت چهارگانه و کدگذاری در یک ناحیه بلوکی ۱۶×۱۶.....	۴۹
شکل ۱۰-۲: درخت‌های جهت‌دار مکانی.....	۵۳
شکل ۱۱-۲: بخش‌بندی درخت چهارگانه و درخت بلوکی در WBTC.....	۵۴
شکل ۱۲-۲: روش‌های کدگذاری ROI در استاندارد JPEG2000.....	۵۹
شکل ۱-۳: تکنیک کدگذاری ROI ارائه شده در [۱۱].....	۶۳
شکل ۲-۳: نمودار بلوکی روش پیشنهادی اول.....	۶۴
شکل ۳-۳: ماسک ROI.....	۶۶
شکل ۴-۳: نمودار بلوکی روش پیشنهادی دوم.....	۷۰
شکل ۵-۳: نمودار بلوکی روش پیشنهادی سوم.....	۷۵
شکل ۶-۳: مراحل مختلف تکنیک پیشنهادی برای جداسازی نواحی مطلوب و غیرمطلوب تصاویر چهره پرسنلی.....	۷۹
شکل ۷-۳: نتایج حاصل از تکنیک پیشنهادی برای جداسازی ناحیه چهره.....	۸۰
شکل ۸-۳: تکنیک پیشنهادی برای کدگذاری مرز ناحیه پیش‌زمینه.....	۸۲
شکل ۹-۳: نمایش نقطه شروع و میزان جابجایی‌ها نسبت به این نقطه برای کدگذاری مرز ناحیه چهره.....	۸۶
شکل ۱۰-۳: الگوی پیمایش Z.....	۸۸
شکل ۱۱-۳: ترتیب قرارگیری ضرایب موجک مطابق الگوی Z در یک ماتریس ضرایب ۱۶×۱۶ با دو سطح تجزیه.....	۸۸
شکل ۱۲-۳: منحنی‌های سطح تجزیه-انرژی تجمعی میانگین ضرایب موجک ناحیه چهره پیش و پس از بازآرایی.....	۹۱

- شکل ۴-۱: چند نمونه تصویر چهره پرسنلی از پایگاه داده..... ۹۳
- شکل ۴-۲: منحنی‌های نرخ-اعوجاج سرتاسر تصویر برای روش‌های مختلف..... ۱۰۰
- شکل ۴-۳: منحنی‌های نرخ-اعوجاج ناحیه چهره برای روش‌های مختلف..... ۱۰۱
- شکل ۴-۴: منحنی‌های نرخ-زمان اجرای فشرده‌سازی برای روش‌های مختلف..... ۱۱۱
- شکل ۴-۵: منحنی‌های نرخ-زمان اجرای بازسازی برای روش‌های مختلف..... ۱۱۲

جدول ۱-۲: مقایسه پیچیدگی محاسباتی.....	۳۹
جدول ۱-۴: مقایسه کمی و کیفی یک تصویر بازسازی شده با استفاده از روش های مختلف در چندین نرخ بیت.....	۱۰۴
جدول ۲-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۱.....	۱۰۷
جدول ۳-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۸.....	۱۰۸
جدول ۴-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۶.....	۱۰۹
جدول ۵-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۴.....	۱۱۰
جدول ۶-۴: متوسط زمان اجرای تکنیک های استفاده شده در روش های پیشنهادی روی ۴۰ تصویر پایگاه داده.....	۱۱۳



مقدمه

۱-۱ تعریف مسئله

امروزه حجم داده‌ها و اطلاعات از جمله تصاویر چه در فضای اینترنت، چه در رایانه‌های شخصی و چه در سازمان‌ها و شرکت‌ها روز به روز در حال افزایش است. از طرفی فضای حافظه کافی برای ذخیره‌سازی این حجم عظیم از داده‌ها در دسترس نیست. اینجاست که فشرده‌سازی نقش مثبت خود را نشان می‌دهد. البته شاید برخی بگویند که، درست است که امروزه حجم داده‌ها روز به روز افزایش می‌یابد، اما از آنطرف با پیشرفت تکنولوژی، ظرفیت حافظه‌ها نیز روز به روز در حال افزایش است. در پاسخ به این افراد می‌توان سه دلیل برای ضرورت فشرده‌سازی ذکر کرد:

(۱) سرعت افزایش حجم حافظه‌ها به اندازه سرعت افزایش حجم داده‌ها نیست.

(۲) خرید حافظه‌های گران قیمت مقرون به صرفه نیست.

(۳) هرچه حجم داده‌ها بیشتر باشد، سرعت دسترسی به آن‌ها و نیز سرعت انتقالشان، مخصوصاً در شبکه‌های با پهنای باند کم کاهش می‌یابد.

روش‌های فشرده‌سازی را بطور کلی می‌توان به دو دسته فشرده‌سازی بی‌اتلاف^۱ و فشرده‌سازی بااتلاف^۲ تقسیم‌بندی کرد. نرخ فشرده‌سازی^۳ نوع بااتلاف بسیار بیشتر از نوع بی‌اتلاف است، اما این مزیت فشرده‌سازی بااتلاف به قیمت از دست دادن بخشی از داده‌ها تمام می‌شود.

¹ Lossless Compression

² Lossy Compression

³ Compression Rate

در حوزه داده‌های تصویری همانند سایر حوزه‌های داده تاکنون روش‌های زیادی برای فشرده‌سازی ارائه شده است. حتی برخی از این روش‌ها به عنوان روش‌های استاندارد پذیرفته شده‌اند، که از بهترین آن‌ها می‌توان به استاندارد JPEG [۱] و نسخه جدیدتر و کارآمدتر آن یعنی JPEG-2000 [۲] اشاره کرد. اکثر روش‌هایی که تاکنون برای فشرده‌سازی تصاویر ارائه شده‌اند، همه‌منظوره^۱ هستند، بدین معنی که برای انواع و اقسام تصاویر کاربرد دارند. درست است که همه‌منظوره بودن یک مزیت محسوب می‌شود اما در عین حال عیب نیز هست. مزیت می‌باشد چرا که قابل استفاده برای هر نوع تصویری است، و عیب محسوب می‌شود چرا که تمام تصاویر را به یک روش فشرده کرده و به محتوای تصاویر توجهی ندارد. در سال‌های اخیر توجه خاصی به این موضوع شده و لذا باب جدیدی از تحقیقات در حوزه فشرده‌سازی تصاویر بازگشته است. در حقیقت می‌توان گفت که روش‌های همه‌منظوره قادر نیستند که از تزیاید^۲ موجود در یک خانواده محدود از تصاویر، به عنوان مثال تصاویر پزشکی، تصاویر متنی و یا تصاویر چهره بطور بهینه استفاده نمایند. همچنین این روش‌ها در نرخ‌های فشرده‌سازی بالا، کیفیت نواحی و ویژگی‌های مهم تصاویر را بشدت کاهش می‌دهند.

۱-۱-۱ ضرورت فشرده‌سازی مبتنی بر ناحیه مطلوب تصاویر چهره

در میان خانواده‌های تصاویر، می‌توان گفت که تصاویر چهره جزء محبوبترین و پرکاربردترین تصاویر هستند. مهمترین کاربرد تصاویر چهره در سیستم‌های تشخیص و شناسایی چهره^۳ و تصدیق

¹ General Purpose

² Redundancy

³ Face Detection and Face Recognition

هویت^۱ است که در امور نظارتی، کنترلی و امنیتی مورد استفاده قرار می‌گیرند. این سیستم‌ها معمولاً نیازمند پایگاه بزرگی از تصاویر چهره افراد مختلف هستند، مخصوصاً اگر در مراکز و سازمان‌های بزرگ، مهم، ملی و بین‌المللی مورد استفاده قرار گیرند. از طرفی کارایی^۲ این سیستم‌ها بسیار وابسته به تشخیص صحیح ناحیه چهره و شناسایی صحیح ویژگی‌های چهره^۳ است. تحقیقات گوناگون نشان داده است که برخی ویژگی‌های چهره نقش مهمی را در شناسایی چهره ایفا می‌کنند. به عنوان مثال در [۳] به این نکته تاکید شده است که در روش‌های کنونی، ویژگی‌هایی از قبیل چشم‌ها، بینی و دهان برای شناسایی چهره‌ها مورد استفاده قرار می‌گیرند. محدودیت حافظه، کاهش هزینه حافظه و نیز افزایش سرعت ذخیره‌سازی و بازیابی^۴ از جمله دلایلی هستند که ما را ملزم به فشردن تصاویر چهره می‌کنند. همانطور که قبلاً گفته شد، اکثر روش‌های فشردن تصاویر متداول، کیفیت تصاویر را بطور سراسری کاهش می‌دهند، چرا که عمل فشردن تصاویر در این روش‌ها در سرتاسر تصویر با یک نرخ بیت یکسان صورت می‌گیرد. بنابراین استفاده از روش‌های متداول برای فشردن تصاویر چهره مخصوصاً در نرخ بیت‌های پایین، نتیجه‌ای جز کاهش شدید کیفیت نواحی و ویژگی‌های مهم چهره و در نهایت کاهش دقت و کارایی سیستم شناسایی چهره نخواهد داشت. برای رفع این مشکل می‌توان به جای فشردن تصاویر سرتاسر تصویر در یک نرخ بیت یکسان، به نواحی مهم چهره نسبت به سایر نواحی اهمیت و نرخ بیت بیشتری اختصاص داد. این نوع فشردن تصاویر را مبتنی بر ناحیه مطلوب (ROI)^۵ می‌نامند.

¹ Identification

² Performance

³ Facial Features

⁴ Retrieval

⁵ Region Of Interest

۱-۲ مروری بر کارهای گذشته

مطالعات و تحقیقات انجام شده تاکنون در زمینه فشرده‌سازی تصاویر چهره با تکیه بر نواحی مطلوب بسیار محدود است. در این بخش بطور مختصر به مرور و بررسی کارهای انجام شده در این زمینه خواهیم پرداخت.

بطور کلی راهکارهای ارائه شده برای فشرده‌سازی مبتنی بر ROI تصاویر چهره را می‌توان به دو دسته تقسیم‌بندی کرد: فشرده‌سازی مبتنی بر فرهنگ داده^۱ یا کتاب رمز^۲ و فشرده‌سازی مبتنی بر تبدیل^۳. البته این دسته‌بندی خیلی کلی بوده و در مورد سایر حوزه‌های فشرده‌سازی تصاویر نیز صادق است.

۱-۲-۱ روش‌های مبتنی بر فرهنگ داده

در روش‌های مبتنی بر فرهنگ داده، تصاویر به بلوک‌هایی مجزا تقسیم شده و هر بلوک بطور مجزا براساس فرهنگ داده بدست آمده در فاز آموزش^۴ کدگذاری^۵ می‌شود. در فاز آموزش برای هر موقعیت بلوک یک فرهنگ داده با اندازه مشخص بدست می‌آید. این فرهنگ داده‌ها معمولاً با استفاده از

^۱ Dictionary-Based Compression

^۲ Code Book

^۳ Transform-Based Compression

^۴ Training Phase

^۵ Encoding

الگوریتم‌های خوشه‌بندی^۱ یا تبدیلات ماتریسی روی مجموعه بزرگی از تصاویر با نام مجموعه آموزش بدست می‌آیند. هر بلوک را می‌توان بصورت یک بردار فرض کرد، در اینصورت هر فرهنگ داده نیز مجموعه‌ای از بردارها خواهد بود که هر یک از این بردارها را اصطلاحاً کلمه رمز^۲ می‌نامند. در حالتی که از الگوریتم‌های خوشه‌بندی در فاز آموزش استفاده شود، به ازای یک موقعیت بلوک مشخص، بردارهای متناظر با این موقعیت بلوکی در مجموعه آموزش به تعدادی خوشه دسته‌بندی می‌شوند و از هر خوشه یک بردار به عنوان نماینده آن خوشه در فرهنگ داده مربوطه قرار خواهد گرفت. بردار نماینده در هر خوشه برداریست که بطور میانگین بیشترین شباهت را به تمامی اعضای خوشه داشته باشد. این میزان شباهت معمولاً توسط یک معیار فاصله سنجیده می‌شود. در فاز آزمایش^۳ با سنجش میزان شباهت هر بلوک با تک تک اعضای فرهنگ داده مربوطه و سپس ارسال اندیس^۴ متناظر با شبیه‌ترین عضو فرهنگ داده به واحد کدگشا^۵ یا ذخیره آن در حافظه عمل کدگذاری صورت می‌گیرد. در صورت استفاده از تبدیلات ماتریسی در فاز آموزش، به ازای یک موقعیت بلوک مشخص، فرهنگ داده مربوطه مجموعه‌ای از بردارهای یکامتعامل^۶ خواهد بود که یک سیستم پایه یکامتعامل را برای نمایش بردارهای متناظر با این موقعیت بلوکی در مجموعه آموزش تشکیل می‌دهند. در این حالت در فاز آزمایش، هر بلوک بصورت یک ترکیب خطی از اعضای فرهنگ داده مربوطه نشان داده شده، و

¹ Clustering

² Code Word

³ Testing Phase

⁴ Index

⁵ Decoder

⁶ Orthonormal Vectors

سپس با ارسال ضرایب چندی سازی شده^۱ این ترکیب خطی به واحد کدگشا یا ذخیره آن‌ها در حافظه عمل کدگذاری صورت می‌گیرد. البته ممکن است به منظور کاهش تزايد اندیس‌های حاصل از حالت اول و یا ضرایب ترکیب‌های خطی حاصل از حالت دوم، یک مرحله اضافی به نام کدگذاری آنتروپی^۲ نیز بر روی این داده‌ها صورت گیرد تا در نهایت بتوان به نرخ فشرده‌سازی بالاتری دست یافت. در روش‌های فشرده‌سازی مبتنی بر فرهنگ داده، دقت بازسازی هر بلوک وابسته به اندازه فرهنگ داده مربوط به آن است. بدین معنی که با افزایش تعداد اعضای یک فرهنگ داده، دقت بازسازی بلوک مربوطه نیز افزایش می‌یابد. البته این به معنای افزایش حافظه جانبی مورد نیاز برای ذخیره‌سازی این فرهنگ داده‌ها نیز هست و لذا بایستی بین این دو مسئله مصالحه‌ای صورت گیرد. در این روش‌ها اندازه فرهنگ داده‌ها با توجه به اهمیت بلوک‌های مربوطه مشخص می‌شود، بطوریکه به بلوک‌های مهمتر فرهنگ داده‌های بزرگتر و در نتیجه نرخ بیت بیشتری اختصاص داده می‌شود. در تصاویر چهره، این بلوک‌های مهم بلوک‌هایی هستند که ویژگی‌های مهم چهره را در بر دارند. در ادامه به مروری بر چند روش ارائه شده در این دسته می‌پردازیم.

در [۴] با تغییر شکل هندسی تصاویر چهره به یک فرم معیار^۳، تقسیم تصویر حاصل به بلوک‌های مجزا و در نهایت کدگذاری هر بلوک به روش چندی سازی برداری (VQ)^۴ عمل فشرده‌سازی صورت می‌گیرد. تبدیل تصاویر به فرم معیار و نیز استفاده از VQ این روش را قادر به بهره‌گیری از تزايد

¹ Quantized Coefficients

² Entropy Coding

³ Canonical

⁴ Vector Quantization

موجود در ویژگی‌های چهره یکسان افراد مختلف، و در نتیجه افزایش کارایی فشرده‌سازی مخصوصا در نرخ‌بیت‌های پایین ساخته است. از معایب این روش می‌توان به زمان طولانی مورد نیاز برای آموزش فرهنگ داده‌های VQ و نیز حافظه جانبی نسبتا زیاد ۴۰ مگابایتی برای ذخیره‌سازی این فرهنگ داده‌ها در هر دو واحد کدگذار و کدگشا اشاره کرد. عیب دیگر این روش عناصر مزاحم بصری^۱ ایجاد شده در تصویر، ناشی از تکنیک بلوکی کردن و کدگذاری مجزای هر بلوک است. در [۵] یک روش مبتنی بر فرهنگ داده برای فشرده‌سازی تصاویر چهره ارائه شده است که کارایی خوبی مخصوصا در نرخ‌بیت‌های پایین دارد. در این روش ابتدا مشابه [۴] شکل هندسی تصاویر به یک فرم معیار تغییر داده شده، و سپس بلوک‌های تصویر حاصل بصورت مجزا با استفاده از نمایش پراکنده^۲ کدگذاری می‌شوند. نمایش پراکنده نیازمند فرهنگ داده‌های فوق کامل^۳ است، که این فرهنگ داده‌ها در فاز آموزش با استفاده از الگوریتم تجزیه برداری منفرد (SVD)^۴ تولید می‌شوند. استفاده از نمایش پراکنده باعث افزایش کارایی فشرده‌سازی این روش نسبت به [۴] شده است، اما در عین حال همچنان همان معایب روش [۴] را با خود یدک می‌کشد. این روش در ادامه با حذف عناصر مزاحم بصری، با استفاده از یک تکنیک بلوک‌شکنی در [۶] توسعه داده شد. این تکنیک بلوک‌شکنی مبتنی بر فیلترهای محلی^۵ بوده و به عنوان یک مرحله پس‌پردازش انجام می‌گیرد. با استفاده از این تکنیک یکی از معایب روش [۵] تقریبا برطرف شده و کارایی آن تا حدودی افزایش یافته است.

¹ Visually Disturbing Artifacts

² Sparse Representation

³ Overcomplete

⁴ Singular Value Decomposition

⁵ Local Filters

۱-۲-۲ روش‌های مبتنی بر تبدیل

در این روش‌ها هدف آن است که با اعمال تبدیلی مناسب روی تصویر، نمایش جدیدی از تصویر بدست آید بطوریکه تزايد و همبستگی^۱ میان داده‌های تصویر در نمایش جدید حذف شده و یا کاهش یابد. این هدف عموماً با متراکم‌سازی و بسته‌بندی بیشترین اطلاعات تصویر در کمترین تعداد ضرایب تبدیل تحقق می‌یابد. تبدیل مناسب تبدیلی است که بتواند آنتروپی یا همان بی‌نظمی داده‌های تصویر را هرچه بیشتر کاهش داده و در نتیجه امکان فشرده‌سازی بیشتری را فراهم سازد. یک تبدیل ممکن است داده‌های تصویر را در حوزه فرکانس به نمایش گذارد مانند تبدیل کسینوسی گسسته (DCT)^۲ و تبدیل والش-هادامارد^۳، یا داده‌های تصویر را به فضای مکانی دیگری منتقل کرده و توصیف نماید مانند تحلیل مولفه اصلی (PCA)^۴. علاوه بر این تبدلاتی وجود دارند که قادرند اطلاعات تصویر را هم در حوزه مکان و هم در حوزه فرکانس به نمایش گذارند، که معروفترین و پرکاربردترین آن‌ها تبدیل موجک گسسته (DWT)^۵ است. تبدیل DWT با قابلیت بالای خود در متمرکزسازی انرژی، در اختیار داشتن اطلاعات مکانی و فرکانسی بصورت همزمان و دارا بودن خاصیت تحلیل محلی و چندمقیاسی^۶ نه تنها توانسته است خود را به عنوان تبدیلی مناسب در زمینه فشرده‌سازی مطرح سازد، بلکه در سایر کاربردهای پردازش تصویر، ویدئو، صوت و سیگنال نیز قابلیت‌های خود را به رخ کشیده است. اکثر

¹ Correlation

² Discrete Cosine Transform

³ Walsh-Hadamard Transform

⁴ Principal Component Analysis

⁵ Discrete Wavelet Transform

⁶ Local and Multiscale Analysis

روش‌های مبتنی بر تبدیل ارائه شده در زمینه فشرده‌سازی تصاویر چهره نیز تبدیل DWT را به عنوان پایه و اساس فشرده‌سازی برگزیده‌اند، که با توجه به ویژگی‌های قابل توجه این تبدیل دور از انتظار نیست. در این روش‌ها برای بهبود کیفیت نواحی مهم چهره در نرخ‌بیت‌های پایین، از تکنیک‌های کدگذاری ROI در حوزه موجک استفاده شده است. در ادامه به مروری بر چند روش ارائه شده در این دسته می‌پردازیم.

در [۷] یک روش فشرده‌سازی تصاویر چهره مبتنی بر ناحیه مطلوب ارائه شده است که از تبدیل DWT بهره می‌گیرد. در این روش پس از تشخیص ناحیه مستطیلی ویژگی‌های چهره شامل چشم‌ها، بینی و دهان به عنوان ناحیه مطلوب، و نیز محاسبه ماسک ROI در حوزه موجک، از تکنیک کدگذاری MSBshift برای کدگذاری ضرایب موجک ROI استفاده می‌شود. این تکنیک کدگذاری کارایی بالاتر و البته پیچیدگی بیشتری نسبت به دو روش انتقال ماکزیمم (Maxshift) و مقیاس‌گذاری کلی^۱ دارد. از طرفی تعیین دقیق ماسک ضرایب موجک ROI یکی از مشکلات این روش به حساب می‌آید. روش کیفیت مدرج^۲ معرفی شده در [۸] نیز یک روش مبتنی بر DWT است. در این روش تصویر چهره پرسنلی به سه ناحیه پس‌زمینه، سر و شانه‌ها و چهره تقسیم شده، و سپس ضرایب موجک مربوط به این نواحی با توجه به اهمیت مشخص شده در تصویر کیفیت کدگذاری می‌شوند. کدگذاری در این روش توسط الگوریتم کدگذاری بلوک جاسازی‌شده با کوتاه‌سازی بهینه (EBCOT)^۳ [۹]، مبتنی بر JPEG2000 صورت می‌گیرد. تصویر کیفیت، تصویری هم بعد با تصویر

¹ General Scaling

² Quality Graded

³ Embedded Block Coding with Optimised Truncation

اصلی بوده و دارای سه ناحیه با سطوح خاکستری متفاوت می‌باشد، که این سطوح بیان کننده کیفیت مورد نیاز هر ناحیه می‌باشند. براساس این تصویر کیفیت، تمامی ضرایب پایین‌ترین زیرباند فرکانسی در اولویت اول کدگذاری قرار داشته، تمامی ضرایب فرکانس بالای ناحیه پس‌زمینه صفر شده، تمامی ضرایب فرکانس بالای ناحیه چهره یک بیت به بالا منتقل شده و تمامی ضرایب ناحیه سر و شانه‌ها بدون تغییر باقی می‌مانند. به منظور کاهش اندازه ضرایب موجک ناشی از اختلاف زیاد روشنایی مرز ناحیه پس‌زمینه و سر و شانه‌ها، در این روش یک مرحله لایه‌گذاری^۱ پس زمینه نیز انجام گرفته است. مشکل عمده این روش را نیز می‌توان مشابه [۷] عدم امکان تعیین دقیق ضرایب موجک متناظر با هر ناحیه دانست. در [۱۰] یک روش سازگار با استاندارد JPEG2000 برای فشرده‌سازی تصاویر چهره ارائه شده است. این روش نیز مشابه [۸] پس از تقسیم‌بندی تصویر چهره به سه ناحیه پس‌زمینه، بدن و چهره آن‌ها را بر حسب اولویت به روش JPEG2000 کدگذاری می‌کند. در این روش ناحیه چهره با استفاده از روش Maxshift در بالاترین اولویت کدگذاری قرار می‌گیرد. از طرفی با صفر کردن مقادیر پیکسل‌های ناحیه پس‌زمینه، ناحیه بدن در اولویت بالاتری نسبت به پس‌زمینه قرار خواهد گرفت. میزان روشنایی پس‌زمینه در صورت نیاز می‌تواند به عنوان اطلاعات جانبی به واحد کدگشا ارسال شود. مشکل دو روش قبلی در مورد این روش نیز صادق است. در [۱۱] روش جدیدی مبتنی بر DWT برای کدگذاری ویژگی‌های چهره در نرخ‌بیت‌های پایین به منظور فشرده‌سازی تصاویر چهره پرسنلی معرفی شده است. در این روش ابتدا با کدگذاری ماتریس ضرایب موجک در یک نرخ‌بیت مشخص، و سپس کدگشایی و بازسازی آن یک نسخه کلی و تقریبی از تصویر چهره تولید می‌شود. سپس ضرایب موجک باقیمانده ناحیه ویژگی‌های چهره با استفاده از روش Maxshift با اولویت بالا در

¹ Padding

نرخ بیت دیگری کدگذاری می‌شوند. کدگشایی و بازسازی دنباله بیت حاصل از مرحله قبل منجر به تولید یک تصویر باقیمانده شامل ویژگی‌های چهره خواهد شد. با افزودن این تصویر به تصویر تقریبی بدست آمده از چهره در مرحله نخست، کیفیت ویژگی‌های چهره بهبود خواهد یافت. در این روش از الگوریتم بخش‌بندی مجموعه بصورت درخت‌های سلسله مراتبی (SPIHT)^۱ [۱۲]، به منظور کدگذاری بهره گرفته شده است. از مزایای این روش می‌توان به پیچیدگی کم و سرعت نسبتاً بالای فرآیند کدگذاری و کدگشایی اشاره کرد. در اینجا نیز تعیین دقیق ضرایب موجک ویژگی‌های چهره یک مشکل محسوب می‌شود. علاوه بر این امکان اختصاص نرخ بیت‌هایی مخصوص ویژگی‌های چهره و غیرچهره، که کدگذاری دقیقاً طبق این نرخ بیت‌ها صورت گیرد، وجود ندارد.

۳-۱ انتخاب طرح کلی

در بخش قبل به بررسی انواع روش‌های موجود برای فشرده‌سازی مبتنی بر ROI تصاویر چهره پرداختیم و مشاهده کردیم که هر دو دسته روش‌های مبتنی بر فرهنگ داده و مبتنی بر تبدیل دارای مزایا و معایبی هستند. بطور کلی می‌توان گفت روش‌های مبتنی بر فرهنگ داده در صورتی که به درستی طراحی شده باشند، در فشرده‌سازی پایگاه‌های بسیار بزرگ تصاویر چهره نسبت به روش‌های مبتنی بر تبدیل کاراتر خواهند بود. در حقیقت این روش‌ها قادرند با بهره‌گیری از شباهت موجود میان تصاویر چهره در موقعیت‌های مکانی مشابه، تزیاید اطلاعات میان تصاویر مجموعه‌ای بزرگ را به مقدار زیادی کاهش دهند. البته مسائل چالش برانگیزی همچون انتخاب الگوریتم مناسب برای آموزش

¹ Set Partitioning In Hierarchical Trees

فرهنگ داده‌ها و نیز تعیین اندازه مناسب بلوک‌ها، معمولاً طراحی این روش‌ها را مشکل و پیچیده می‌سازد. علاوه بر این ممکن است ساختار چهره و محل قرارگیری ویژگی‌های چهره در تصویر برای افراد مختلف یکسان نباشد، و در نتیجه کارایی فشرده‌سازی کاهش یابد. برخی از این روش‌ها از جمله [۴-۶] برای حل این مشکل به سراغ یک مرحله هم‌ترازی^۱ و انطباق تصاویر رفته‌اند. این مرحله هم‌ترازی در قالب یک مرحله پیش‌پردازش هم در فاز آموزش و هم در فاز آزمایش بر روی تصاویر اعمال گشته و با نگاشت ویژگی‌های چهره به موقعیت‌های مکانی مشابه، آن‌ها را از نظر هندسی به یک فرم معیار تغییر شکل می‌دهد. تبدیل تصاویر به یک فرم معیار این روش‌ها را قادر ساخته است که به خوبی از شباهت تصاویر چهره در موقعیت‌های مکانی مشابه بهره گرفته و تزايد اطلاعات میان آن‌ها را تا حد زیادی کاهش دهند. اما در عین حال استفاده از یک مرحله پیش‌پردازش روی تصاویر، پیچیدگی این روش‌ها را چه در فاز آموزش و چه در فاز آزمایش افزایش داده و در نتیجه منجر به طولانی‌تر شدن فاز آموزش و نیز کاهش سرعت فشرده‌سازی و بازسازی تصاویر شده است. بطور کلی روش‌های مبتنی بر فرهنگ داده نیازمند زمانی نسبتاً طولانی برای آموزش فرهنگ داده‌ها روی مجموعه‌ای بزرگی از تصاویر، و نیز حافظه جانبی نسبتاً زیادی برای ذخیره‌سازی این فرهنگ داده‌ها در هر دو واحد کدگذار و کدگشا هستند. از طرفی در این روش‌ها فشرده‌سازی تصاویر در نرخ‌بیت‌های مختلف و به تبع آن بازسازی تصاویر با کیفیت‌های مختلف به سادگی امکان‌پذیر نیست و کاری بس زمان‌بر و هزینه‌بر است، چرا که نیازمند محاسبه دقیق نرخ‌بیت مورد نیاز هر یک از بلوک‌ها و سپس تکرار مجدد فاز آموزش برای دستیابی به فرهنگ داده‌های متناسب با نرخ‌بیت و دقت مورد نیاز است. می‌توان گفت روش‌های مبتنی بر فرهنگ داده روش‌هایی تطبیقی هستند، بدین معنی که در فاز

¹ Alignment

آموزش با فراگیری دانش و اطلاعاتی راجع به یک جامعه آماری محدود از تصاویر چهره، قادرند تصاویری که متعلق به این جامعه آماری هستند و یا شباهت زیادی به اعضای آن دارند را با کارایی بالایی فشرده و بازسازی کنند. بنابراین این روش‌ها تنها روی مجموعه‌ای محدود از تصاویر چهره تطبیق یافته‌اند و روی تصاویر خارج از این مجموعه کارایی مناسب را نخواهند داشت. به عنوان مثال ممکن است در ابتدا تنها تصاویر چهره سفیدپوستان درون پایگاه داده موجود بوده و لذا فرهنگ داده‌ها روی نژاد سفید تطبیق یافته باشند، اما به تدریج تصاویر جدیدی از نژادهای مختلف به پایگاه داده اضافه شده و کارایی فشرده‌سازی کاهش یابد. در چنین مواردی لازم است که مجدداً فاز آموزش روی جامعه جدید تکرار شده و فرهنگ داده‌های جدید و مناسب بدست آیند که مسلماً فرآیند زمان‌بری خواهد بود.

با کمی سبک سنگین کردن مزایا و معایب روش‌های مبتنی بر فرهنگ داده متوجه خواهیم شد که معایب این روش‌ها وزن بیشتری دارند. همین امر باعث شده است که در این پایان‌نامه به سراغ روش‌های مبتنی بر تبدیل رویم و توجه و تمرکز خود را معطوف آن‌ها نماییم. البته بطور واضح روش‌های مبتنی بر تبدیل نیز دارای معایبی هستند، اما این معایب در کنار مزایای زیاد این روش‌ها کمتر به چشم می‌آیند. یک نقص این روش‌ها را می‌توان مشکل کار با نواحی مطلوب دانست، چرا که پس از تبدیل تصویر و انتقال آن به حوزه جدید تشخیص نواحی مطلوب به درستی امکانپذیر نیست. البته تبدیل DWT با قابلیت تحلیل مکانی خود و نمایش همزمان اطلاعات مکانی و فرکانسی تصویر این مشکل را تا حد زیادی برطرف نموده است. از طرفی ممکن است در نرخ‌بیت‌های خیلی پایین روش‌های مبتنی بر تبدیل کارایی روش‌های مبتنی بر فرهنگ داده را نداشته باشند، چرا که روش‌های مبتنی بر تبدیل تنها سعی در کاهش تزیید داده‌های درون یک تصویر را دارند، در حالیکه روش‌های مبتنی بر فرهنگ داده تزیید داده‌ها را میان مجموعه‌ای بزرگ از تصاویر کاهش می‌دهند. در مقابل این

نقایص، روش‌های مبتنی بر تبدیل برخلاف روش‌های مبتنی بر فرهنگ داده نه محتاج به فاز آموزش و فرآیند زمان‌بر آموزش فرهنگ داده‌ها هستند، و نه به حافظه جانبی زیاد برای ذخیره‌سازی فرهنگ داده‌ها نیاز دارند، چرا که این روش‌ها تنها متکی به اطلاعات درون تصویری هستند که باید فشرده شود. از طرفی همین دلیل باعث شده است که این روش‌ها محدود به استفاده روی مجموعه‌ای خاص از تصاویر چهره نباشند، و بتوان با استفاده از آن‌ها تصاویر چهره گوناگون را با کارایی مناسب فشرده و بازسازی کرد. البته با توجه به اینکه ممکن است ساختار تصاویر چهره در پایگاه‌های داده مختلف متفاوت باشد، شاید نیاز به اصلاحاتی در تکنیک جداسازی نواحی مطلوب برای تطبیق با این تفاوت‌ها احساس شود. همانطور که در بخش ۱-۲-۲ گفته شد، در میان تبدیلات مورد استفاده در این روش‌ها، استفاده از تبدیل DWT به دلیل ویژگی‌های کم‌نظیر آن رایج‌تر است. روش‌های مبتنی بر تبدیل DWT علاوه بر مزایای ذکر شده برای کلیه روش‌های مبتنی بر تبدیل، دارای مزایای دیگری نیز هستند. قابلیت بالای تبدیل DWT در متراکم‌سازی انرژی تصویر باعث شده است که این روش‌ها کارایی بالایی در فشرده‌سازی با اتلاف تصاویر داشته باشند. از طرفی استفاده از تبدیل DWT و الگوریتم‌های کدگذاری ضرایب در حوزه موجک این روش‌ها را قادر ساخته است که فشرده‌سازی تصاویر را در هر نرخ بیت دلخواهی به سهولت انجام دهند. در حقیقت الگوریتم‌های کدگذاری ضرایب موجک به گونه‌ای طراحی شده‌اند که بتوان در هر نرخ بیت دلخواهی عمل کدگذاری را متوقف کرده و دقیقاً نرخ بیت مورد نیاز تامین شود. همچنین در این الگوریتم‌ها این امکان فراهم شده است که پس از توقف کدگذاری تصویر در یک نرخ بیت دلخواه R و تولید دنباله بیت B ، بتوان تصویر را در هر نرخ بیت دلخواه کمتر از R با استفاده از همین دنباله بیت B بازسازی کرد، و یا با ادامه کدگذاری از نقطه توقف و چسباندن دنباله بیت حاصل به دنباله بیت B به نرخ بیت‌های بالاتری از تصویر دست یافت. بنابراین تنها با داشتن یک دنباله بیت می‌توان همزمان نرخ بیت‌ها و کیفیت‌های مختلفی از تصویر را در دست

داشت، بدون آنکه نیاز به تکرار چند باره کدگذاری تصویر باشد. حتی برخی از این الگوریتم‌های کدگذاری علاوه بر موارد ذکر شده، این امکان را نیز فراهم کردند که تنها با در اختیار داشتن یک دنباله بیت بتوان تصویر را در مقیاس‌های مختلف بازسازی کرد. چنین الگوریتم‌هایی را اصطلاحاً الگوریتم‌های کدگذاری جاسازی‌شده^۱ می‌نامند. با استفاده از تکنیک‌های ساده کدگذاری ROI در حوزه موجک می‌توان به سادگی به نواحی مطلوب نرخ‌بیت و اهمیت بیشتری اختصاص داد. البته مسئله‌ای که در اینجا تا حدودی چالش‌برانگیز است، عدم امکان تعیین دقیق ماسک ضرایب موجک ROI، و در نتیجه عدم امکان تخصیص نرخ‌بیتی دقیق به هر یک از نواحی مطلوب و غیر مطلوب است. روش‌های مبتنی بر تبدیل DWT معمولاً دارای پیچیدگی نسبتاً کم و در نتیجه سرعت فشرده‌سازی و بازسازی نسبتاً بالایی هستند. علت این امر را می‌توان پیچیدگی کم تبدیل DWT و نیز الگوریتم‌های کدگذاری ضرایب موجک دانست، چرا که تبدیل DWT تنها متشکل از یک سری عملیات ضرب و جمع مربوط به فیلترها بوده و الگوریتم‌های کدگذاری ضرایب موجک نیز معمولاً تنها متشکل از یک سری عملیات مقایسه‌ای و تصمیم‌گیری‌های منطقی هستند.

۴-۱ اهداف پایان‌نامه

در این پایان‌نامه بطور کلی هدف آن است که با ارائه روش یا روش‌هایی کارآمد برای فشرده‌سازی تصاویر چهره در نرخ‌بیت‌های پایین، برخی از نقایص و معایب روش‌های پیشین تا حدودی جبران شده و کاهش یابند. دستیابی به این هدف کلی در طی دو مرحله زیر صورت گرفته است.

^۱ Embedded Coding

(۱) بررسی روش‌های موجود در زمینه فشردن سازی تصاویر چهره، دسته‌بندی روش‌ها و انتخاب

طرح کلی فشردن سازی به عنوان پایه و مبنای روش‌های پیشنهادی.

(۲) انتخاب روش‌های مناسب و کارآمد برای تحقق هریک از اجزاء و بخش‌های طرح کلی و در

صورت نیاز اصلاح طرح کلی برای تطبیق با مسئله فشردن سازی مبتنی بر ناحیه مطلوب

تصاویر چهره.

در بخش ۱-۲ یک دسته‌بندی کلی برای روش‌های موجود در زمینه فشردن سازی تصاویر چهره

ارائه دادیم، که قابل تعمیم به سایر حوزه‌های فشردن سازی تصاویر نیز هست. سپس براساس این

دسته‌بندی به بررسی مزایا و معایب برخی از روش‌های فشردن سازی تصاویر چهره پرداختیم. در بخش

۱-۳ نیز با ارائه یک مقایسه کلی میان روش‌های مبتنی بر فرهنگ داده و روش‌های مبتنی بر تبدیل و

بررسی مزایا و معایب هر دسته، روش‌های مبتنی بر تبدیل را به عنوان طرح کلی فشردن سازی و پایه و

مبنای روش‌های پیشنهادی برگزیدیم. شکل ۱-۱ طرح کلی مبتنی بر تبدیل فرآیند فشردن سازی و

بازسازی تصاویر چهره را نشان می‌دهد. البته در این طرح کلی به جهت سادگی فرض شده است که

ماسک نواحی مطلوب در هر دو واحد کدگذار و کدگشا در اختیار و معلوم باشند. در غیر اینصورت باید

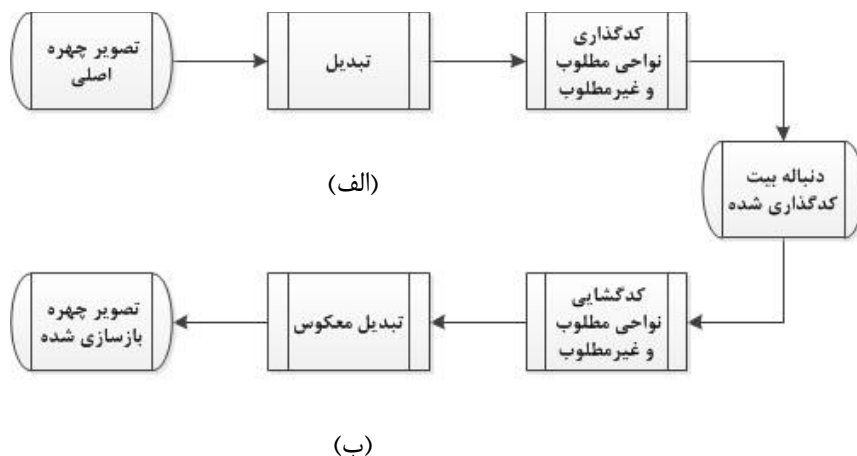
یک مرحله جداسازی نواحی مطلوب نیز در فاز فشردن سازی انجام گیرد. در این حالت چنانچه واحد

کدگشا برای بازسازی تصویر به ماسک نواحی مطلوب نیازمند باشد، یک مرحله دیگر نیز به فرآیند

فشردن سازی اضافه خواهد شد و آن کدگذاری مرز نواحی مطلوب و ارسال دنباله بیت حاصل به عنوان

دنباله بیت سرآیند^۱ به واحد کدگشا است. در روش‌های مبتنی بر تبدیل عموماً اولین گزینه‌ای که

¹ Header Bitstream



شکل ۱-۱: طرح کلی مبتنی بر تبدیل. (الف) فرآیند فشرده‌سازی و (ب) فرآیند بازسازی تصاویر چهره.

برای انتخاب تبدیل به ذهن‌خاطر می‌کند، تبدیل DWT است که معمولاً بهترین و مناسب‌ترین انتخاب نیز هست. برخی از مزایای این تبدیل در بخش ۱-۲-۲ به اختصار توضیح داده شده و توضیحات تکمیلی را نیز می‌توان در فصل دوم یافت. یک قابلیت و ویژگی مهم تبدیل DWT که معمولاً سایر تبدیلات فاقد آن هستند، امکان کار با نواحی مطلوب در حوزه تصویر تبدیل‌یافته است. با استفاده از این قابلیت تحلیل محلی تبدیل DWT و تکنیک‌های کدگذاری ROI در حوزه موجک می‌توان به سادگی عمل فشرده‌سازی و بازسازی تصاویر را مبتنی بر نواحی مطلوب انجام داد، حتی بدون اینکه نیازی به کدگذاری مرز نواحی مطلوب و ارسال آن به واحد کدگشا باشد. الگوریتم‌های متنوعی برای کدگذاری ضرایب در حوزه موجک معرفی شده است. این الگوریتم‌ها را می‌توان بطور کلی به دو گروه الگوریتم‌های مبتنی بر درخت صفر^۱، و مبتنی بر بلوک صفر^۲ دسته‌بندی کرد. نکته مهم در مورد این روش‌ها این است که، هدف همگی آن‌ها کدگذاری و ارسال بیشترین ضرایب موجک

^۱ Zerotree-Based

^۲ Zeroblock-Based

مهم^۱ با کمترین هزینه نرخ بیت و زمان است. الگوریتم‌های مبتنی بر درخت صفر با گروه‌بندی ضرایب موجک متناظر با موقعیت و جهت مکانی مشابه، تحت عنوان درخت جهت‌دار مکانی (SOT)^۲، سعی در بهره‌گیری از همبستگی بین‌باندی^۳ ضرایب موجک را دارند. از معروفترین الگوریتم‌های این دسته می‌توان به SPIHT اشاره کرد. در مقابل الگوریتم‌های مبتنی بر بلوک صفر با تقسیم تصویر تبدیل‌یافته به بلوک‌های مجاور هم، به دنبال بهره‌گیری از همبستگی درون‌باندی^۴ ضرایب موجک هستند. الگوریتم EBCOT از معروفترین الگوریتم‌های این دسته به شمار می‌رود که در استاندارد JPEG2000 نیز به کار گرفته شده است. در این پایان‌نامه از الگوریتم کدگذاری نسبتاً جدید ارائه شده در [۱۳] برای کدگذاری ضرایب موجک استفاده شده است. علت این انتخاب بهره‌گیری همزمان این الگوریتم از همبستگی درون‌باندی و بین‌باندی ضرایب است، که در نتیجه ترکیب ویژگی‌های الگوریتم‌های مبتنی بر درخت صفر و بلوک صفر حاصل شده است. این الگوریتم در اصطلاح کدگذاری درخت بلوکی موجک (WBTC)^۵ نامیده شده است. از طرفی در این پایان‌نامه با استفاده از یک تکنیک پیشنهادی توانسته‌ایم کارایی کدگذاری ضرایب موجک و در نتیجه کارایی فشرده‌سازی را بهبود دهیم. در این تکنیک پیشنهادی ماتریس ضرایب موجک یا همان تصویر تبدیل‌یافته قبل از کدگذاری وارد یک مرحله پیش‌پردازش می‌شود. در مرحله پیش‌پردازش آرایش بلوک‌های ضرایب موجک به نحوی تغییر می‌یابد که ضرایب مهم در الگوریتم WBTC با جستجوی کمتری یافت شده و در نتیجه زمان و

¹ Significant Wavelet Coefficients

² Spatial Orientation Tree

³ Inter-Subband Correlation

⁴ Intra-Subband Correlation

⁵ Wavelet Block Tree Coding

نرخ‌بیت کمتری هزینه شود. البته این روش نیازمند ذخیره‌سازی آرایه‌های ترتیب اهمیت^۱ بدست آمده در فاز شبه-آموزش، در هر دو واحد کدگذار و کدگشا است. فاز شبه-آموزش فازیست که در آن برای هر ناحیه مطلوب تصویر یک آرایه ترتیب اهمیت میانگین روی تصاویر پایگاه داده بدست می‌آید. در این پایان‌نامه برای سادگی فرض شده است که با تصاویر چهره پرسنلی سطح خاکستری سر و کار داریم. علت این فرض آن است که معمولاً تصاویر چهره پرسنلی دارای ساختار و طرح کلی نسبتاً یکسانی هستند که نه تنها باعث کارآمدتر شدن تکنیک بازآرایی بلوک‌های ضرایب مطابق آرایه‌های ترتیب اهمیت شده، بلکه در استخراج نواحی مطلوب نیز کار ساده‌تری را پیش رو خواهیم داشت. تصاویر پرسنلی عموماً تنها شامل ناحیه سر، چهره و بخشی از شانه‌ها بوده و دارای پس‌زمینه‌ای یکنواخت می‌باشند. همچنین معمولاً حدود تقریبی ناحیه چهره و حتی ویژگی‌های چهره در این تصاویر مشخص بوده و یا قابل تشخیص است. این اطلاعات به ما اجازه می‌دهند که نواحی مطلوب خود را بصورت تقریبی تعریف کرده و آن‌ها را به عنوان پیش‌فرض و معلومات مسئله فشرده‌سازی در نظر بگیریم. البته در سطحی بالاتر نیز این اطلاعات به ما کمک می‌کنند که با بکارگیری روش‌هایی ساده و با پیچیدگی کم بتوانیم نواحی مطلوب را با دقت بالاتری استخراج کنیم. در این پایان‌نامه برای جداسازی نواحی مطلوب از یک تکنیک پیشنهادی سریع مبتنی بر عملیات لبه‌گیری و شکل‌شناسی^۲ استفاده شده است.

در این پایان‌نامه طی یک سیر تکاملی سه روش برای فشرده‌سازی مبتنی بر ROI تصاویر چهره پرسنلی در نرخ‌بیت‌های پایین پیشنهاد شده است، که تبدیل DWT و الگوریتم کدگذاری WBTC

¹ Significance Order Arrays

² Morphology

پایه و اساس هر سه روش پیشنهادی را تشکیل می‌دهند. در روش اول از تکنیک کدگذاری ROI ارائه شده در [۱۱] برای بهبود کیفیت ناحیه چهره در نرخ‌بیت‌های پایین استفاده شده است. تلفیق این تکنیک با تکنیک بازآرایی بلوک‌های ضرایب موجک و الگوریتم کدگذاری WBTC در روش پیشنهادی اول باعث شده است که کیفیت ناحیه چهره در نرخ‌بیت‌های پایین نسبت به روش [۱۱] بهبود یابد، و همچنین کارایی آن در نرخ‌بیت‌های پایین از روش‌هایی همچون استاندارد JPEG2000 و SPIHT بالاتر باشد. در این روش بازسازی تصاویر در نرخ‌بیت‌های پایین با سرعت بالایی انجام گرفته و سرعت فشرده‌سازی نیز در حد مناسبی است. بنابراین این روش مناسب کاربردهایی است که نیازمند فشرده‌سازی و بازسازی سریع و نیز ارسال و دریافت سریع تصاویر پرسنلی در نرخ‌بیت‌های پایین هستند، که از آن جمله می‌توان سیستم‌های تشخیص هویت بلادرنگ را مثال زد.

در روش دوم از تکنیک جدیدی برای فشرده‌سازی مبتنی بر ROI تصاویر چهره پرسنلی بهره گرفته‌ایم. در این روش با استفاده از تبدیل موجک مبتنی بر ناحیه (RWT)^۱ [۱۴]، فشرده‌سازی نواحی مطلوب و غیرمطلوب بطور کاملاً مجزا صورت گرفته است. بنابراین در این روش می‌توان هر ناحیه را دقیقاً با نرخ‌بیت دلخواه و مورد نیاز فشرده و بازسازی کرد. یک عیب این روش نیاز تبدیل RWT و معکوس آن به ماسک نواحی مطلوب است، که در نتیجه کدگذاری مرز نواحی مطلوب و ارسال آن به واحد کدگشا ضروری خواهد بود. در این روش برای کدگذاری مرز نواحی چهره و سر و شانه به ترتیب از تکنیک‌هایی مبتنی بر اختلاف فاصله رئوس چندضلعی و چندجمله‌ای درونیاب^۲ بهره گرفته‌ایم. از طرفی سرعت فشرده‌سازی و بازسازی تصاویر در این روش نسبت به روش اول کمتر است،

^۱ Region-based Wavelet Transform

^۲ Interpolating Polynomial

چرا که پیچیدگی محاسباتی در این روش افزایش یافته است، و علت آن را می‌توان فشرده‌سازی نواحی بصورت مجزا، اضافه شدن تعداد مراحل بازآرایی بلوک‌های ضرایب بسته به تعداد نواحی مطلوب و نیز اضافه شدن مراحل کدگذاری و کدگشایی مرز نواحی مطلوب دانست. البته در این روش امکان فشرده‌سازی و بازسازی نواحی مختلف تصویر بصورت موازی وجود دارد، که در صورت استفاده از این قابلیت سرعت فشرده‌سازی و بازسازی بسته به تعداد نواحی چند برابر خواهد شد. نتایج شبیه‌سازی بیانگر کارایی بالاتر این روش در ناحیه چهره و در نرخ‌بیت‌های پایین در مقایسه با روش پیشنهادی اول، روش [۱۱]، استاندارد JPEG2000 و نیز SPIHT است.

روش سوم در حقیقت تلفیقی از روش‌های اول و دوم است. در این روش با بهره‌گیری توأم از تبدیل RWT و ایده کدگذاری ROI ارائه شده در [۱۱]، توانسته‌ایم علاوه بر دستیابی به کارایی بالاتر نرخ-عوجاج^۱ در نرخ‌بیت‌های پایین نسبت به روش‌های اول و دوم، سرعت فشرده‌سازی و بازسازی تصاویر را نیز نسبت به روش دوم افزایش دهیم. در این روش نیز مشابه روش اول، عمل فشرده‌سازی/بازسازی سرتاسر تصویر تا یک نرخ‌بیت مشخص بطور همزمان صورت می‌گیرد. سپس برای افزایش کیفیت نواحی مطلوب، با اعمال تبدیل RWT روی این نواحی در تصویر باقیمانده و کدگذاری آن‌ها با نرخ‌بیتی دیگر، اهمیت و نرخ‌بیت بیشتری به این نواحی اختصاص داده می‌شود. این روش در مقایسه با روش‌های پیشنهادی اول و دوم و به تبع آن روش [۱۱]، JPEG2000 و SPIHT کارایی بالاتری را در ناحیه چهره و بخصوص در نرخ‌بیت‌های پایین، در شبیه‌سازی‌های انجام شده نشان داده است.

¹ Rate-Distortion

۱-۵ ساختار پایان نامه

در فصل دوم این پایان نامه با مفاهیم پایه مورد نیاز شامل مبانی تبدیل موجک، تبدیل موجک مبتنی بر ناحیه، الگوریتم‌های کدگذاری تصویر مبتنی بر موجک و بطور خاص الگوریتم WBTC و نیز روش‌های کدگذاری ROI در حوزه موجک به اختصار آشنا خواهیم شد. در فصل سوم به تشریح روش‌های پیشنهادی در این پایان نامه برای فشرده‌سازی مبتنی بر ROI تصاویر چهره پرسنلی خواهیم پرداخت. در فصل چهارم نتایج شبیه‌سازی‌های انجام شده به نمایش گذاشته می‌شوند و مورد تحلیل و بررسی قرار می‌گیرند. در فصل پنجم و انتهایی پایان نامه نیز به نتیجه‌گیری و ارائه پیشنهادات خواهیم پرداخت.



مفاهیم پایه

۲-۱ تبدیل موجک

گرچه پایه‌های ریاضی موجک‌ها ریشه در مباحث قدیمی ریاضی و فیزیک دارد، ولی استفاده از آن‌ها به عنوان ابزاری برای تحلیل و پردازش سیگنال‌ها به اواسط دهه ۸۰ میلادی باز می‌گردد. تحول عظیم و اساسی در استفاده و کاربرد تبدیل موجک در مباحث پردازش سیگنال با مقاله ملت^۱ [۱۵] شروع شد. او با ارائه دیدگاهی جدید از تبدیل موجک، آن را به عنوان ابزار و راهکاری جدید و قدرتمند برای پردازش سیگنال و تحلیلی موسوم به تحلیل چند درجه تفکیک^۲ معرفی کرد. کاربرد تبدیل موجک در زمینه فشرده‌سازی تصویر تا قبل از معرفی استاندارد JPEG بسیار محدود بود که علت آن را می‌توان نبود روشی کارآمد برای کدگذاری ضرایب تبدیل موجک دانست [۱۶]. با معرفی روش موجک درخت صفر جاسازی‌شده (EZW)^۳ توسط جوزف شاپیرو^۴ در [۱۷]، انقلابی در زمینه کدگذاری ضرایب موجک و فشرده‌سازی تصویر مبتنی بر تبدیل موجک بوجود آمد، به گونه‌ای که از آن زمان تاکنون روش‌های متنوع و متعددی چه در زمینه فشرده‌سازی تصاویر مبتنی بر تبدیل موجک، و چه در زمینه کدگذاری ضرایب موجک معرفی شده‌اند و این روند همچنان ادامه دارد.

¹ Mallat

² Multiresolution Analysis

³ Embedded Zerotree Wavelet

⁴ Jussef Shapiro

۲-۱-۱ کدگذاری زیرباند

قبل از پرداختن به تبدیل موجک، بهتر است ابتدا با کدگذاری زیرباند^۱ که به نوعی از اجداد تبدیل موجک به شمار رفته و گاهی به آن تبدیل موجک اولیه نیز گفته می‌شود به اختصار آشنا شویم [۱۸]. در حقیقت این دو اصطلاح یکی هستند و فقط از دیدگاه‌های مختلف به مسئله نگاه می‌کنند، اصطلاح زیرباند از دید مهندسی و اصطلاح تبدیل موجک از دید ریاضی.

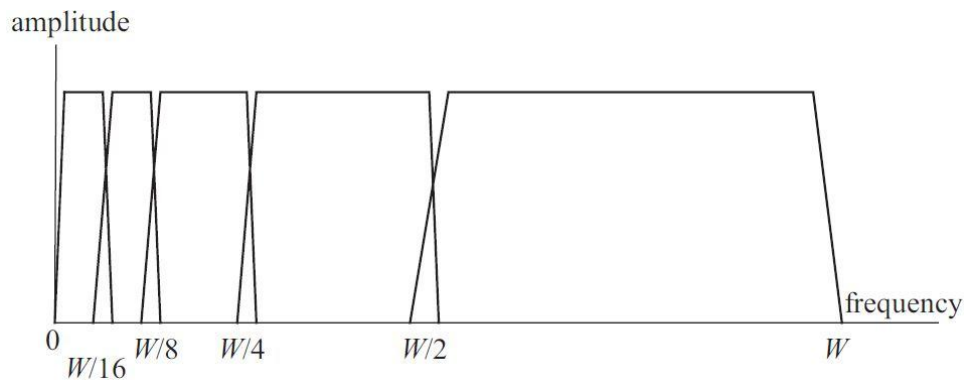
ایده اصلی تکنیک کدگذاری زیرباند، بخش‌بندی طیف سیگنال به چندین محدوده یا باند فرکانسی و سپس کدگذاری و انتقال هر باند به طور مجزا می‌باشد [۱۹]. در این تکنیک، جداسازی باندهای فرکانسی تصویر توسط فیلترکردن تصویر ورودی به کمک بانکی از فیلترهای میان‌گذر مطابق با شکل ۱-۲ انجام می‌شود. در نتیجه این عملیات فیلترینگ، تصویر ورودی بسته به تعداد فیلترهای موجود در فیلتربانک به چندین نسخه تصویر فیلترشده با محتوای فرکانسی متفاوت تجزیه می‌شود. به منظور تطبیق محتوای فرکانسی تصاویر فیلتر شده حاصل با خصوصیات سیستم بینایی انسان، فیلترها به صورت باندهای اکتاو^۲ (یعنی افزایش/کاهش عرض باندها به صورت توان صحیحی از ۲) طراحی و تنظیم شده‌اند. از آنجا که پهنای باند هر نسخه تصویر فیلتر شده کاهش یافته است، از نظر تئوری مطابق قضیه نایکوئیست^۳ می‌توان با نرخ کمتری از این تصاویر نمونه برداری کرد، و در نتیجه تعدادی زیرتصویر کاهش بعد یافته^۴ به دست آورد. در واحد کدگذار این زیرتصاویر پس از چندی‌سازی،

¹ Subband Coding

² Octave Bands

³ Nyquist

⁴ Reduced size subimage



شکل ۲-۱: بانکی از فیلترهای میان‌گذر [۱۶].

کدگذاری شده و ارسال می‌شوند. در واحد کدگشا، زیرتصویرهای دریافتی به ابعاد اولیه بازگردانده شده و از مجموعه‌ای از فیلترها به نام فیلتربانک ترکیب^۱ عبور داده می‌شوند. این فیلتربانک عمل درونیایی و ترکیب زیرتصویرها به منظور بازسازی تصویر اولیه را انجام می‌دهد.

۲-۱-۲ تبدیل موجک

تبدیل موجک حالت خاصی از کدگذاری زیرباند است که در زمینه کدگذاری تصویر و ویدیو بسیار فراگیر و متداول شده است. کدگذاری زیرباند تصاویر مبتنی بر تحلیل فرکانسی است حال آنکه تبدیل موجک بر تئوری تقریب^۲ استوار است. با این حال از آنجا که تصاویر طبیعی به طور محلی هموار بوده و می‌توان آن‌ها را با تقریب تکه‌ای-چندجمله‌ای^۳ مدل کرد، در صورت انتخاب مناسب تابع چندجمله‌ای می‌توان مشابه با ایده کدگذاری زیرباند به نوعی تحلیل فرکانسی دست یافت.

^۱ Synthesis filter bank

^۲ Approximation theory

^۳ Piecewise Polynomial

از نظر ریاضی، چنانچه $x(t)$ تابعی مربعی-انتگرالپذیر^۱ باشد، تبدیل موجک آن معادل با تجزیه این تابع برحسب مجموعه‌ای از توابع پایه به صورت زیر است:

$$X_w(a, b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}(t) dt \quad (۱-۲)$$

که در آن $\psi_{a,b}(t)$ به نام تابع پایه شناخته می‌شود. این تابع نسخه‌ای گسترش یافته و انتقال یافته^۲ از یک سیگنال میان‌گذر مانند $\psi(t)$ است که موجک مادر^۳ نامیده شده و به صورت زیر تعریف می‌شود:

$$\psi_{a,b} = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (۲-۲)$$

که در آن a و b به ترتیب پارامترهای گسترش^۴ و انتقال^۵ نامیده می‌شوند. تاثیر این پارامترها در شکل ۲-۲ نشان داده شده است. عرض تابع پایه با تغییر فاکتور گسترش (یا مقیاس) a به این شکل تغییر می‌کند که هرچه این فاکتور بزرگتر شود، عرض زمانی تابع پایه نیز بزرگتر می‌شود و بنابراین عرض (یا پهنای باند) فرکانسی آن کمتر می‌شود. بنابراین به کمک این پارامتر می‌توانیم درجه‌ی تفکیک زمانی و فرکانسی را (با مصالحه‌ای که بین یکدیگر دارند) در تبدیل موجک تغییر دهیم، و همین ویژگی جالب تبدیل موجک است که این تبدیل را برای تحلیل سیگنال‌هایی همچون تصاویر طبیعی که ویژگی‌هایی با اندازه‌های مختلف دارند، مناسب می‌سازد.

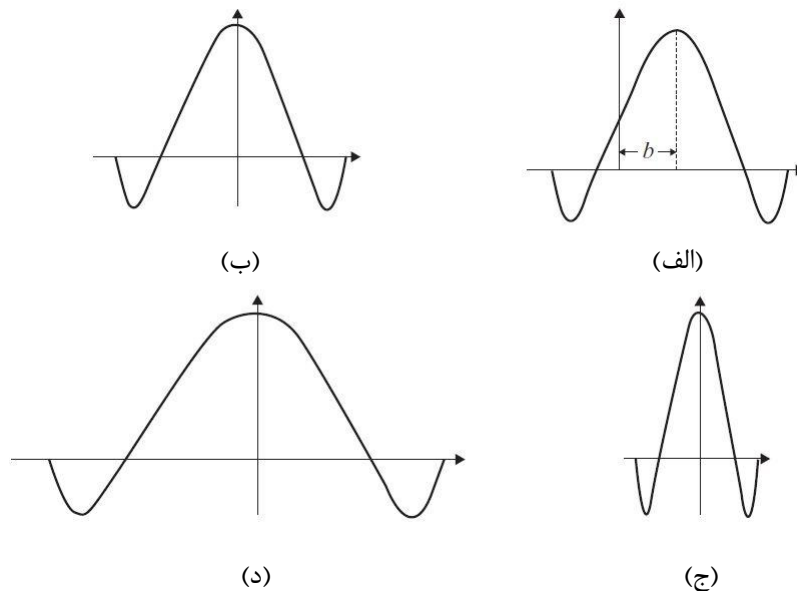
¹ Square-Integrable

² Dilation and Translation Version

³ Mother Wavelet

⁴ Dilation

⁵ Translation



شکل ۲-۲: تاثیر گسترش و انتقال زمانی روی موجک مادر. (الف) موجک مادر با $a = 1$ و $b = 0$. (ب) حالت

$a = 1$ و $b \neq 0$. (ج) حالت $a = 2$ و $b = 0$. (د) حالت $a = 1/2$ و $b = 0$. [۱۶]

متناظر با هر اندازه‌ی ویژگی^۱ یک تابع پایه $\psi_{a,b}(t)$ وجود دارد که آن ویژگی به بهترین وجه توسط این تابع پایه تحلیل می‌شود. به عنوان مثال، در تصویری از یک خانه که شخصی از پنجره بیرون را نگاه می‌کند، تابع پایه‌ی متناظر با مقدار a بزرگ، به راحتی کل خانه را تحلیل می‌کند. اما شخص کنار پنجره در مقیاسی کوچکتر بهتر تحلیل خواهد شد و همینطور چشم‌های شخص در مقیاسی باز هم کوچکتر. بنابراین، تبدیل موجک همانند تحلیل سیگنال با فیلترهای میان‌گذری است که فرکانس مرکزی آن‌ها متغیر و وابسته به پارامتر a بوده، اما دارای ضریب کیفیتی^۲ ثابت هستند. لازم به ذکر است که ضریب کیفیت یک فیلتر میان‌گذر، نسبت فرکانس مرکزی به پهنای باند آن فیلتر تعریف می‌شود.

¹ Feature Size

² Quality Factor

۳-۱-۲ تبدیل موجک گسسته

تبدیل موجک تعریف شده در رابطه‌ی (۱-۲) سیگنال یک بعدی $x(t)$ را به تابع دوبعدی $X_w(a, b)$ نگاشت می‌کند و بنابراین موجب تولید تزايد زیادی می‌شود. سیگنال اولیه را می‌توان از روی تبدیل موجک و به ازاء مقادیر گسسته‌ای از پارامترهای a و b به دست آورد [۲۰]. پارامتر a را می‌توان با انتخاب $a = a_0^m$ ، بطوریکه $a_0 > 1$ و m عددی صحیح باشد، گسسته کرد. با افزایش مقدار a ، پهنای باند (یا همان درجه تفکیک فرکانسی) تابع پایه کاهش می‌یابد و بنابراین برای پوشش یک ناحیه‌ی فرکانسی مشخص، نیاز به گام‌ها یا در اصطلاح سلول‌های درجه تفکیک^۱ بیشتری است. بطور مشابه، گسسته‌سازی پارامتر b نیز با نمونه‌برداری در محور زمان انجام می‌گیرد. فرکانس این نمونه‌برداری به پهنای باند سیگنالی که قرار است نمونه‌برداری شود وابسته بوده، و در نتیجه با پارامتر a تناسب معکوس دارد. پارامتر b را می‌توان با انتخاب $b = nb_0a_0^m$ گسسته کرد. به ازای $a_0 = 2$ و $b_0 = 1$ انتخاب‌هایی برای تابع $\psi(t)$ وجود دارند، بطوریکه توابع $\psi_{m,n}(t)$ حاصل از آن‌ها پایه‌های یکامتعامدی را در فضای توابع مربعی-انتگرال‌پذیر تشکیل می‌دهند. این امر بدین معناست که هر تابع مربعی-انتگرال‌پذیر $x(t)$ را می‌توان بصورت یک ترکیب خطی از توابع پایه همانند زیر نمایش داد:

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \alpha_{m,n} \psi_{m,n}(t) \quad (3-2)$$

که در آن ضرایب $\alpha_{m,n}$ ، به ضرایب تبدیل موجک $x(t)$ معروف هستند، و از طریق رابطه‌ی (۴-۲) بدست می‌آیند:

¹ Resolution Cell

$$\alpha_{m,n} = \int_{-\infty}^{\infty} x(t) \psi_{m,n}(t) dt \quad (4-2)$$

جالب است توجه شود که با هر بار افزایش m ، مقدار a دو برابر می‌شود، و این به معنی دو برابر شدن عرض زمانی و نصف شدن عرض یا پهنای باند فرکانسی است. این ویژگی معادل با تحلیل سیگنال به کمک تجزیه‌ی اکتاو باندهای فرکانسی و نیز تبدیل موجک توان-دو^۱ است. در تبدیل موجک توان-دو مشابه کدگذاری زیرباند، برای تحلیل سیگنال از یک تکنیک مبتنی بر فیلتربانک موجک که به الگوریتم هرمی نیز معروف است، استفاده شده است [۲۱]. همانطور که در بخش ۲-۱ ذکر شد، یک فیلتربانک از فیلترهایی تشکیل شده است که یک سیگنال را به باندهای فرکانسی مختلف تجزیه می‌کند [۲۲]. در الگوریتم هرمی، سیگنال اصلی توسط دو فیلتر موجک پایین‌گذر L و بالاگذر H به دو سیگنال فرکانس پایین و فرکانس بالا تجزیه شده و سپس زیرنمونه‌برداری^۲ می‌شوند. هریک از این سیگنال‌ها حاوی نیمی از محتوای فرکانسی سیگنال اصلی هستند. خروجی فیلترهای پایین‌گذر و بالاگذر به ترتیب موسوم به اطلاعات تقریب و جزئیات هستند. در تحلیل چند درجه تفکیک، سیگنال تقریب بدست آمده در سطح تجزیه^۳ اول، مجدداً با عبور از فیلترهای L و H زیرنمونه‌برداری به سیگنال‌های تقریب و جزئیات تجزیه شده و این روند برای سیگنال تقریب سطح تجزیه دوم به بعد تا حصول تعداد سطوح تجزیه مطلوب به همین شکل ادامه می‌یابد. تجزیه موجک چند سطحی سیگنال اصلی در نهایت منجر به تولید یک سیگنال تقریب مربوط به سطح آخر تجزیه، و دنباله‌ای از سیگنال‌های جزئیات مربوط به سطوح اول تا آخر تجزیه خواهد شد. با اینکه تبدیل

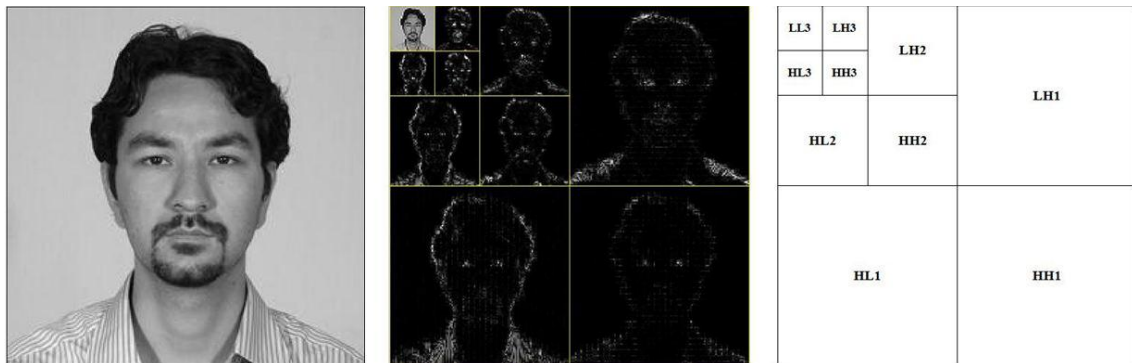
¹ Dyadic Wavelet Transform

² Downsampling

³ Decomposition Level

موجک توان-دو حالت خاصی از تبدیل موجک گسسته است، اما در اغلب اوقات منظور از تبدیل موجک گسسته همان تبدیل موجک توان-دو است. در این پایان نامه نیز از این پس منظور ما از تبدیل موجک گسسته، تبدیل موجک توان-دو خواهد بود.

در مورد سیگنال‌های دوبعدی مانند تصاویر، چنانچه سطرها و ستون‌ها به عنوان سیگنال‌هایی یک‌بعدی در نظر گرفته شوند، به روشی مشابه فوق می‌توان تبدیل موجک گسسته دوبعدی را بدست آورد [۲۳]. در این روش ابتدا در جهت افقی از تک تک سطرهای تصویر توسط دو فیلتر L و H تبدیل موجک گسسته یک‌بعدی گرفته می‌شود. در نتیجه این عمل تصویر اصلی به دو زیرتصویر مستطیلی هم‌اندازه با محتوای فرکانسی متفاوت تجزیه می‌شود. تعداد سطرهای این زیرتصاویر با تصویر اصلی برابر، و تعداد ستون‌های آن‌ها به دلیل زیرنمونه‌برداری نصف تصویر اصلی است. حال این بار در جهت عمودی مجدداً با اعمال فیلترهای L و H روی تک تک ستون‌های این زیرتصاویر، از آن‌ها تبدیل موجک گسسته یک‌بعدی گرفته می‌شود. با این کار در نهایت چهار زیرتصویر حاصل خواهد شد. زیرتصویری که در نتیجه دو بار اعمال متوالی فیلتر L در جهت‌های افقی و عمودی روی تصویر اصلی حاصل شده است، اصطلاحاً زیرتصویر تقریب نامیده شده و با LL نشان داده می‌شود. به همین ترتیب زیرتصاویر LH ، HL و HH به ترتیب زیرتصاویر جزئیات افقی، عمودی و قطری نامیده می‌شوند. اکنون یک تجزیه موجک یک‌سطحی از تصویر اصلی بدست آمده است. برای دستیابی به سطح دوم تجزیه موجک کافیهست همین روند را این بار روی زیر تصویر تقریب یعنی LL تکرار کرده و آن را به چهار زیر تصویر $LL2$ ، $LH2$ ، $HL2$ و $HH2$ تجزیه کنیم. این کار را می‌توان برای دستیابی به سطوح بالاتر تجزیه موجک همینطور ادامه داد. شکل ۲-۳ یک تصویر چهره پرسنلی و تبدیل موجک گسسته دوبعدی آن را با سه سطح تجزیه نشان می‌دهد. همانطور که در شکل ۲-۳ مشاهده می‌شود، زیرتصویر



شکل ۲-۳: تبدیل موجک گسسته دوبعدی یک تصویر چهره پرسنلی با سه سطح تجزیه.

LL3 یک نسخه تقریبی از تصویر چهره است که حاوی بیشترین انرژی تصویری اصلی در مقایسه با سایر زیرتصاویر است.

۲-۱-۴ تبدیل موجک مبتنی بر ناحیه

تبدیل موجک سیگنال داده شده را به زیرباندهای مختلف فرکانسی تجزیه می‌کند و امکان بهره‌مندی از قابلیت‌های تحلیل چندمقیاسی را در اختیارمان می‌گذارد. روش‌های فشرده‌سازی مبتنی بر موجک در نرخ‌های فشرده‌سازی بالاتر کیفیت تصویر بهتری را ارائه می‌دهند. علت این امر را می‌توان قابلیت بالای تبدیل موجک در متراکم‌سازی انرژی در زیرباندهای فرکانسی پایین دانست. در کنار مزایای فراوان تبدیل موجک، عدم امکان اعمال آن بر روی نواحی و یا اشیاء تصویر با شکل دلخواه یک نقص محسوب می‌شود. این نقص باعث دشواری کاربرد تبدیل موجک در روش‌های کدگذاری ROI شده است، چرا که اعمال و کاربرد آن را به نواحی مستطیلی محدود ساخته است. به منظور رفع این مشکل انواع گوناگونی از تبدیلات موجک مبتنی بر شکل دلخواه^۱ پیشنهاد شده‌اند

¹ Arbitrarily Shaped

[۱۴] و [۲۴-۲۶]. تبدیل موجک افقی شکل (SWAT)^۱، برای اعمال تبدیل موجک به هر ناحیه با شکل دلخواهی پیشنهاد شده است [۲۴] و [۲۵]. این الگوریتم به منظور تبدیل ناحیه با شکل دلخواه، تمامی پیکسل‌های سطرها و ستون‌های این ناحیه را در یک ردیف قرار می‌دهد. این تکنیک از نظر محاسباتی کارآمد است، اما ممکن است در تصویر بازسازی‌شده باعث اعوجاجات شدیدی میان پیکسل‌های درون ناحیه شود، چرا که شکاف‌ها^۲ یا حفره‌های درون ناحیه حذف خواهند شد. در روش دیگری به نام تبدیل موجک مبتنی بر شیء^۳، با تبدیل شیء با شکل دلخواه به یک ناحیه مستطیلی، تبدیل موجک معمولی بر روی آن اعمال می‌شود [۲۶]. این روش از نظر محاسباتی کارآمد نیست، چرا که فرآیند تجزیه در این روش نیازمند عملیات برونمایی^۴ و فیلترگذاری پایین‌گذر^۵ است. در [۱۴] روشی برای اعمال تبدیل موجک بر روی نواحی با شکل دلخواه پیشنهاد شده است، که علاوه بر پیچیدگی محاسباتی پایین، عملکرد بهتری نسبت به روش‌های پیشین از خود نشان داده است. این روش اصطلاحاً تبدیل موجک مبتنی بر ناحیه یا به اختصار RWT نامیده شده است.

تبدیل RWT قابل اعمال بر روی هر ناحیه و یا شیئی درون تصویر با شکل دلخواه است. در تبدیل RWT، عمل تجزیه به جای اعمال فیلترهای موجک به تصویر مستطیلی، با اعمال این فیلترها به ناحیه شیء توسعه‌یافته^۶ صورت می‌گیرد. در تبدیل RWT با حفظ شکل اصلی شیء در هر زیرباند

¹ Shape Adaptive Wavelet Transform

² Break-ups

³ Object-based Wavelet Transform

⁴ Extrapolation

⁵ Low pass filtering

⁶ Extended Object

امکان دستیابی به ماکزیمم ناهمبستگی^۱ نواحی تصویر فراهم است. از طرفی تبدیل معکوس RWT می‌تواند شکل اصلی ناحیه مورد نظر را بطور کامل و بی‌عیب بازسازی کند. از آنجا که تبدیل RWT تنها ناحیه شیء مورد نظر در تصویر اصلی را مورد پردازش قرار می‌دهد، لذا در مقایسه با تبدیل موجک متداول^۲، بطور قابل ملاحظه‌ای نیازمند محاسبات کمتری است.

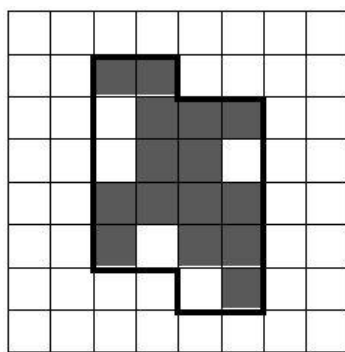
در شکل ۴-۲ الگوریتم تبدیل RWT نشان داده شده است. همانطور که مشاهده می‌شود در اینجا نیز همچون روش‌های تجزیه متداول، روشی تفکیک‌پذیر برای تجزیه در پیش گرفته شده است. در گام اول قبل از اعمال فیلترهای موجک، در جهت افقی از چپ به راست هر دو پیکسل مجاور هم درون تصویر مورد بررسی قرار می‌گیرند تا تعیین شود که آیا این پیکسل‌ها متعلق به شیء مورد نظر هستند یا خیر. چنانچه حداقل یکی از این دو پیکسل بخشی از شیء باشد، هر دوی آن‌ها درون ناحیه شیء توسعه‌یافته لحاظ می‌شوند. نواحی شیء اصلی و توسعه یافته به ترتیب در شکل ۴-۲(الف) و شکل ۴-۲(ب) نشان داده شده‌اند. پس از تعیین ناحیه شیء توسعه یافته، فیلتربانک موجک شامل فیلترهای پایین‌گذر و بالاگذر در جهت افقی به این ناحیه اعمال می‌شوند. نمونه‌های فیلتر شده ناحیه شیء توسعه‌یافته طوری درون یک زیرباند قرار می‌گیرند که شکل اصلی شیء در زیرباند حفظ شود. شکل ۴-۲(ج) نتیجه حاصل از تجزیه افقی را نشان می‌دهد. به روشی مشابه در جهت عمودی از بالا به پایین، تجزیه عمودی بر روی تصویر تجزیه‌شده افقی اجرا می‌شود. نتیجه حاصل از تجزیه با چهار زیرباند LL، LH، HL و HH در شکل ۴-۲(د) نشان داده شده است. باید توجه داشت که تبدیل RWT شکل اصلی شیء تصویر را در هر زیرباند حفظ می‌کند، که ویژگی مهمی در تحلیل

¹ Decorrelation

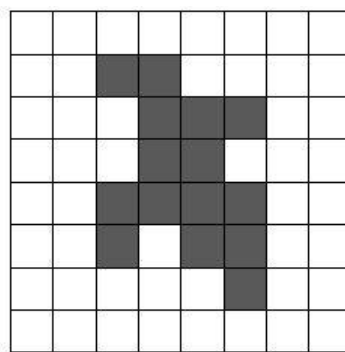
² Conventional Wavelet Transform

چندمقیاسی و کدگذاری کارآمد محسوب می‌شود، چرا که این روش مانع از ایجاد برخی اعوجاج‌ها میان پیکسل‌های تصویر شده و خودتشابهی شیء را در هر زیرباند حفظ می‌کند.

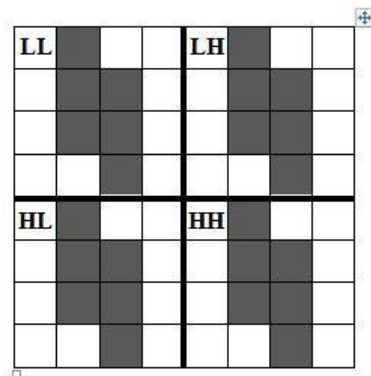
برای اعمال تبدیل RWT به تصویر حقیقی، تصویر قطعه‌بندی شده^۱ مورد نیاز است. شکل ۵-۲ (الف) تصویر "فروشنده" را نشان می‌دهد. این تصویر بصورت دستی به ناحیه شیء و پس‌زمینه قطعه‌بندی شده، که در شکل ۵-۲ (ب) نشان داده شده است. "شیء" شامل سر شخص و بخش بالای



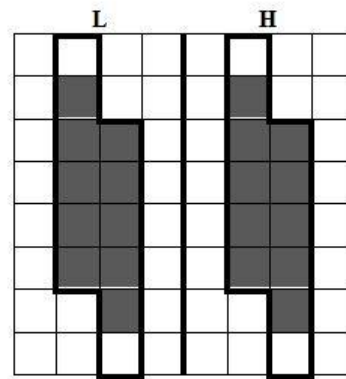
(ب)



(الف)



(د)



(ج)

شکل ۵-۲: گام‌های پردازشی تبدیل RWT. (الف) شیء اصلی. (ب) شیء توسعه‌یافته. (ج) نتیجه تجزیه افقی. (د)

نتیجه تجزیه عمودی اعمال شده روی نتیجه تجزیه افقی [۱۴].

¹ Segmented Image

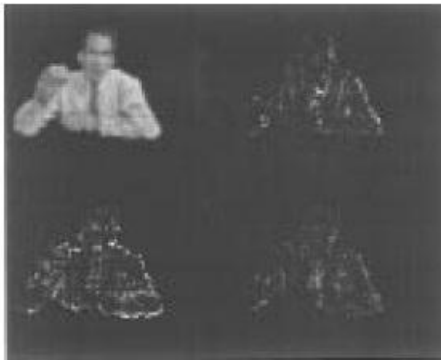
بدن اوست، و پس زمینه با مقدار ثابت صفر پر شده است تا لبه‌ها و مرز بین شیء و پس زمینه بطور کامل قابل تشخیص باشند. شکل ۵-۲ (ج) تصویر تجزیه شده حاصل از تبدیل RWT را نشان می دهد.



(ب)



(الف)



(د)



(ج)



(ه)

شکل ۵-۲: تصویر فروشنده. (الف) تصویر اصلی. (ب) تصویر قطعه بندی شده. (ج) تصویر تجزیه شده با استفاده از تبدیل RWT. (د) تصویر تجزیه شده با استفاده از SAWT. (ه) تصویر بازسازی شده با استفاده از معکوس RWT [۱۴].

به منظور مقایسه، تصویر تجزیه شده حاصل از تبدیل SAWT نیز در شکل ۲-۵(د) نشان داده شده است. مشاهده می شود که تبدیل RWT علاوه بر حفظ شکل اصلی شیء تصویر، مقدار اطلاعات فرکانس بالای کمتری نسبت به تبدیل SAWT تولید کرده است. تصویر بازسازی شده با استفاده از تبدیل معکوس RWT در شکل ۲-۵(ه) نشان داده شده است. میانگین مربعات خطا (MSE) میان شیء اصلی و شیء بازسازی شده برابر صفر است. بنابراین تبدیل معکوس RWT می تواند بطور کامل شیء اصلی را بازسازی کند.

۲-۴-۱ پیچیدگی محاسباتی تبدیل RWT

تبدیل RWT می تواند پیچیدگی محاسباتی تبدیل موجک را بطور قابل ملاحظه ای کاهش دهد. به منظور مقایسه پیچیدگی تبدیل RWT با تبدیل موجک متداول، تعداد عملیات مورد نیاز برای انجام این دو تبدیل ارزیابی می شوند. فرض کنید ابعاد تصویر $N \times N$ ، و طول فیلتر موجک برابر P باشد. اگر M و A ، به ترتیب بیانگر عملیات ضرب و جمع باشند، محاسبات مورد نیاز برای بدست آوردن خروجی فیلتر در هر پیکسل برابر $P.M + (P - 1).A$ خواهد بود. در تبدیل موجک متداول به منظور تجزیه تصویر در جهت افقی با استفاده از فیلتربانک شامل فیلترهای پایین گذر و بالا گذر، تعداد $[P.M + (P - 1).A].N^2$ عملیات مورد نیاز است. اعمال فیلتربانک در جهت عمودی به تصویر تجزیه شده افقی چهار زیرباند تولید می کند، که بطور مشابه نیازمند تعداد $[P.M + (P - 1).A].N^2$ عملیات است. بنابراین، تعداد کل عملیات مورد نیاز برای تجزیه سطح اول تصویر با استفاده از تبدیل موجک متداول، برابر $2.[P.M + (P - 1).A].N^2$ خواهد بود. تبدیل RWT می تواند بطور قابل ملاحظه ای تعداد عملیات ضرب و جمع را کاهش دهد، اگرچه عمل تجزیه توسط RWT نیازمند $\frac{3}{2}N^2$ عملیات اضافی مقایسه به منظور تعیین ناحیه شیء در تصویر است (N^2)

عملیات مقایسه برای تجزیه افقی و $\frac{1}{2}N^2$ عملیات مقایسه برای تجزیه عمودی). با فرض اینکه ناحیه شیء $\frac{1}{K}$ از تصویر را اشغال کرده باشد، تعداد عملیات ضرب و جمع مورد نیاز در RWT به $\frac{1}{K}$ تعداد این عملیات در تبدیل موجک متداول کاهش می‌یابد. بنابراین در تبدیل RWT تجزیه سطح اول تصویر می‌تواند با تعداد $2.P.N^2.\frac{1}{K}$ عملیات ضرب، $2.(P-1).N^2.\frac{1}{K}$ عملیات جمع، و $\frac{3}{2}N^2$ عملیات مقایسه انجام شود. جدول ۱-۲ بطور خلاصه مقایسه پیچیدگی محاسباتی تبدیل RWT و تبدیل موجک متداول را نشان می‌دهد.

جدول ۱-۲: مقایسه پیچیدگی محاسباتی [۱۴].

روش تبدیل	نوع عملیات		
	ضرب	جمع	مقایسه
تبدیل موجک متداول	$2.P.N^2$	$2.(P-1).N^2$	0
RWT	$2.P.N^2.\frac{1}{K}$	$2.(P-1).N^2.\frac{1}{K}$	$\frac{3}{2}N^2$

۲-۲ کدگذاری تصویر مبتنی بر موجک

ارسال سریع تصاویر از طریق دستگاه‌های قابل حمل یا متحرک، که عموماً در فضای حافظه، قدرت پردازش و عمر باتری محدودیت دارند، نیازمند الگوریتم‌هایی برای کدگذاری تصاویر است که بتوانند با پیچیدگی محاسباتی پایین تصاویر را به صورتی کارآمد فشرده کنند. همچنین جستجوی تصاویر در اینترنت از میان مجموعه داده‌های تصویری، نیازمند الگوریتم‌های کدگذاری و کدگشایی

سریعی است که از نظر نرخ-اعوجاج نیز کارایی بالایی داشته باشند. چنین الگوریتم‌هایی باید قابلیت ساخت و بازسازی پیشرو^۱ تصاویر کدشده مبتنی بر موجک، و دستیابی به تصاویری با کیفیت خوب در مراحل اولیه انتقال را داشته باشند. این ویژگی‌ها مخصوصاً هنگامی اهمیت بیشتری پیدا می‌کنند که جستجوی تصاویر وب از طریق خطوط بی‌سیم صورت گیرد، چرا که در این حالت محدودیت ظرفیت کانال، محدودیت حافظه و نیز محدودیت در پیچیدگی محاسباتی عواملی مسجل و قطعی خواهند بود.

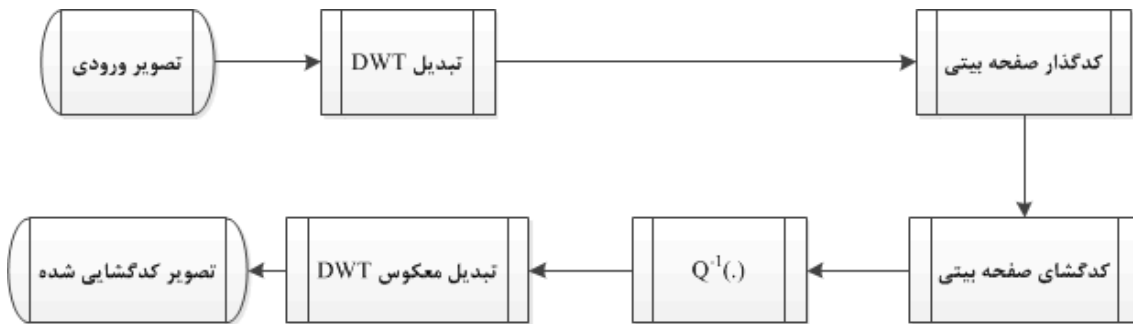
تبدیل موجک به نحو بسیار خوبی قادر به خوشه‌بندی انرژی در مکان و فرکانس است [۹]. بنابراین میزان موفقیت و کارایی یک الگوریتم کدگذاری تصویر مبتنی بر موجک، به میزان بهره‌برداری آن از این خاصیت خوشه‌بندی انرژی تبدیل موجک بستگی دارد. در طی سالیان گذشته، تعدادی الگوریتم بسیار موفق برای کدگذاری تصویر مبتنی بر موجک معرفی شده‌اند. این الگوریتم‌ها را می‌توان بطور کلی به دو گروه الگوریتم‌های مبتنی بر درخت صفر، و مبتنی بر بلوک صفر دسته‌بندی کرد. اغلب این الگوریتم‌ها کدگذارهای مبتنی بر صفحه بیتی^۲ هستند که در بخش ۲-۲-۱ به اختصار با این نوع کدگذاری بیشتر آشنا خواهیم شد. نکته مهم در مورد این روش‌ها این است که، هدف همگی آن‌ها کدگذاری و ارسال بیشترین ضرایب موجک مهم با کمترین هزینه نرخ‌بیت و زمان است. دستیابی به این هدف مستلزم ارائه روشی بهینه چه از نظر نرخ‌بیت و چه از نظر زمان، برای جستجوی ضرایب مهم خواهد بود. معیار ارزیابی اهمیت ضرایب در چنین روش‌هایی معمولاً اندازه و بزرگی ضرایب در نظر گرفته می‌شود، چرا که ثابت شده است ضرایب موجک بزرگتر، نقش و تاثیر بیشتری در کاهش اعوجاج تصویر فشرده‌شده دارند [۲۷].

¹ Progressive

² Bit Plane-Based

۲-۱-۲ کدگذاری مبتنی بر صفحه بیتی ضرایب موجک

شکل ۲-۶ نمودار بلوکی یک سیستم کدگذاری تصویر، مبتنی بر کدگذاری صفحه بیتی ضرایب موجک را نشان می‌دهد. ابتدا تصویر ورودی با اعمال تبدیل DWT ناهمبسته شده، و سپس ضرایب تبدیل موجک به کدگذار صفحه بیتی تحویل داده می‌شوند. کدگذاری صفحه بیتی در حقیقت یک روش تقریب ضرایب موجک مبتنی بر صفحه بیتی است. ایده اصلی در این نوع کدگذاری مشابه نمایش دودویی یک عدد حقیقی بصورت $c = \sum_{\{n\}} b_n 2^n$ است، که $n = n_{max}, n_{max} - 1, \dots$ و $b_n \in \{0, 1\}$ و n_{max} شماره پرارزش‌ترین صفحه بیتی (MSB)^۱ بزرگترین ضریب موجک هستند. هر بیت اضافی b_n ، دقت این نمایش دودویی را افزایش داده و بازه نامعینی یا عدم قطعیت^۲ را برای عدد c نصف می‌کند. بنابراین دقت نهایی نمایش یک عدد یا ضریب مانند c ، توسط تعداد کل بیت‌های دریافت شده آن کنترل می‌شود.

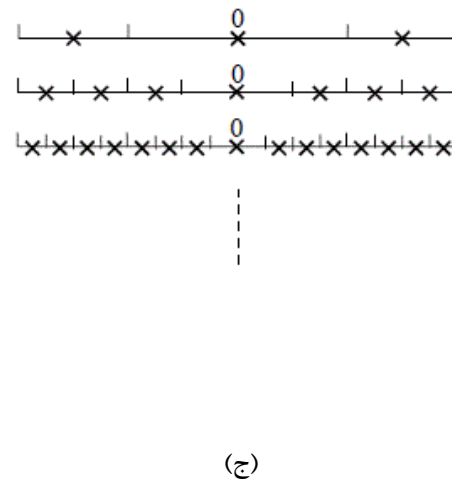
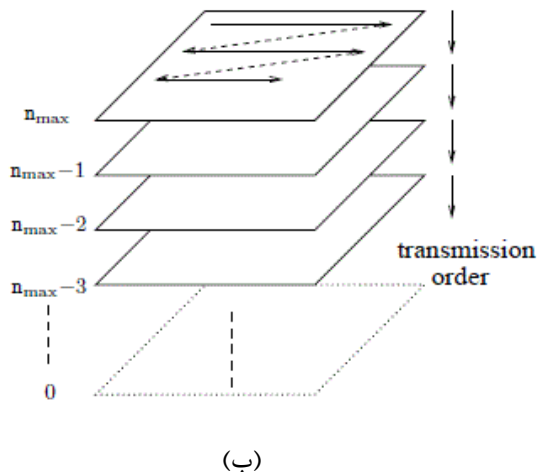
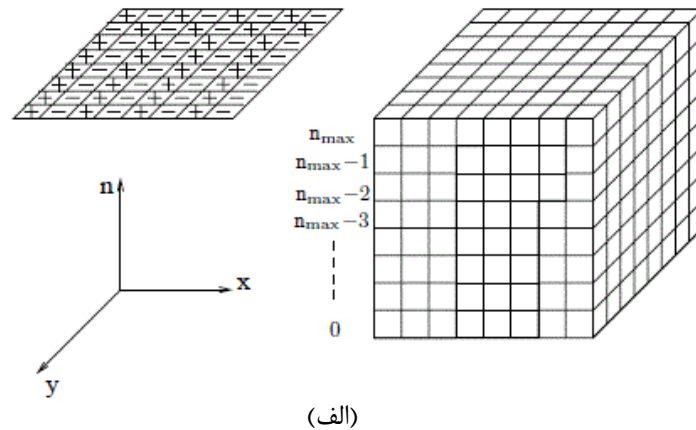


شکل ۲-۶: یک سیستم کدگذاری تصویر مبتنی بر کدگذاری صفحه بیتی ضرایب موجک [۲۸].

^۱ Most Significant Bit Plane

^۲ Uncertainty Interval

برای اعمال این ایده پایه به فشردگی پیشر و تصویر، ابتدا ضرایب موجک $w_{i,j}$ به شکل علامت-اندازه^۱ یعنی $w_{i,j} = \text{sgn}(w_{i,j})|w_{i,j}|$ (همانطور که در شکل ۷-۲ الف) نشان داده شده است) نمایش



شکل ۷-۲: توضیح نحوه نمایش، انتقال و چندی سازی ضرایب در کدگذاری صفحه بیتی موجک. (الف) نمایش دودویی ضرایب موجک در شکل علامت-اندازه. (ب) کدگذاری صفحه بیتی و ترتیب انتقال. (ج) چندی سازیهای موثر برای صفحات بیتی، مقادیر بازسازی شده توسط "x" علامت گذاری شده اند [۲۸].

¹ Sign-Magnitude Form

داده می‌شوند؛ که $sgn(.)$ علامت ضریب ورودی را برمی‌گرداند. فرآیند کدگذاری از پرارزش‌ترین صفحه بیتی (MSB) شروع شده و به سوی کم‌ارزش‌ترین صفحه بیتی (LSB)^۱ پیش می‌رود. در هر صفحه بیتی همانطور که در شکل ۲-۷ (ب) نشان داده شده است، تمامی داده‌های مربوط به اندازه ضرایب مطابق یک ترتیب پیمایش^۲ از پیش تعریف شده کدگذاری می‌شوند. در فرآیند کدگذاری بعد از مواجهه با بیت MSB (اولین بیت "۱") مربوط به اندازه یک ضریب برای اولین بار، بیت علامت آن نیز بلافاصله کدگذاری می‌شود. در طی فرآیند کدگذاری، یک ضریب معلوم ضربی مهم نامیده می‌شود، در صورتی که بیت MSB اندازه آن قبلاً کدگذاری شده باشد؛ و در غیر اینصورت ضربی غیرمهم^۳ نامیده می‌شود. در این روش، تمام ضرایب موجب به تدریج و بطور پیشرونده بیت به بیت تصحیح می‌شوند. فرآیند کدگذاری پس از اتمام بودجه بیتی اختصاص یافته و یا دستیابی به دقت مطلوب، می‌تواند در هر صفحه بیتی و هر موقعیت پیکسلی متوقف شود.

فرض کنید فرآیند کدگذاری دقیقاً در انتهای صفحه بیتی n ام متوقف شود. در این حالت از دید واحد کدگشا تمامی ضرایب واقع در بازه $(-2^n, 2^n)$ ، ضرایب غیرمهم شناخته می‌شوند. همچنین اندازه بازه‌های نامعینی برای تمامی ضرایب مهم برابر 2^n خواهد بود. در کدگذاری ترتیبی متداول^۴ نتایج یکسانی با استفاده از بازه‌های چندی‌سازی رابطه (۲-۵) بدست می‌آید.

¹ Least Significant Bit Plane

² Scanning Order

³ Insignificant Coefficient

⁴ Conventional Sequential Coding

$$\mathfrak{X}_q^n = \begin{cases} (q2^n - 2^n, q2^n] & \text{if } q < 0 \\ (-2^n, 2^n) & \text{if } q = 0 \\ [q2^n, q2^n + 2^n) & \text{if } q > 0 \end{cases} \quad (5-2)$$

که q مقدار خروجی چندی‌ساز^۱ برای یک نمونه ورودی $x \in \mathfrak{X}_q^n$ است. تابع چندی‌سازی مربوطه به سادگی متناظر است با چندی‌سازی آستانه‌ای یکنواخت (UTQ)^۲ با یک اندازه گام 2^n و یک ناحیه مرده^۳ دو برابر اندازه گام [۲۸]. بنابراین، اگرچه عمل چندی‌سازی بطور صریح و واضح انجام نمی‌گیرد، اما کدگذار صفحه بیتی ضرایب موجک بطوری کارآمد تمامی ضرایب را توسط یک خانواده از چندی‌سازهای یکنواخت جاسازی شده، بیت به بیت چندی‌سازی می‌کند. اندازه گام چندی‌سازی متوالی برای هر صفحه بیتی اضافی، با یک ضریب ۲ کاهش می‌یابد. شکل ۷-۲ (ج) چندی‌سازهای موثر مربوط به صفحات بیتی را نشان می‌دهد.

هنگامی که هیچ دانش قبلی راجع به آمار منبع در دسترس نیست، معمولاً نقطه میانی بازه‌های چندی‌سازی به عنوان مقادیر بازسازی شده ضرایب کدگشایی شده انتخاب می‌شوند. این نقاط توسط "x" در شکل ۷-۲ (ج) مشخص شده‌اند. تمامی ضرایب غیرمهم به مقدار صفر نگاشت خواهند شد.

۲-۲-۲ کدگذاری مبتنی بر درخت صفر

الگوریتم‌های مبتنی بر درخت صفر با گروه‌بندی ضرایب موجک متناظر با موقعیت و جهت مکانی مشابه، درخت‌هایی به نام درخت جهت‌دار مکانی یا SOT را تشکیل می‌دهند؛ این درخت‌ها از ضرایب

¹ Quantizer

² Uniform Threshold Quantization

³ Dead Zone

پایین ترین زیرباند فرکانسی به عنوان گره های ریشه^۱ سرچشمه گرفته اند. آنگاه درخت صفر تعریف می شود، یک SOT که نسبت به یک آستانه مشخص (2^n) دارای هیچ ضریب مهمی (بزرگتری) نباشد. کدگذارهای موجک پیشنهاد شده در [۱۲]، [۱۷]، [۲۹] و [۳۰] از نوع مبتنی بر درخت صفر هستند. در میان این کدگذارها، الگوریتم SPIHT پیشنهاد شده توسط Said و Pearlman [۱۲]، بیشترین شهرت و محبوبیت را دارد. الگوریتم SPIHT از کارایی بالایی در کدگذاری و پیچیدگی محاسباتی پایین برخوردار است، و در عین حال دنباله بیت های یکپارچه ای تولید می کند. کدگذاری SPIHT برای فشرده سازی و انتقال پیشرو بهینه^۲ طراحی شده است. فرآیند انتقال را می توان در هر زمان از فرآیند کدگذاری متوقف کرد و به بهترین کیفیت تصویر با نرخ بیت موجود تا آن زمان دست یافت. وابستگی پایین کارایی SPIHT به کدگذاری آنتروپی (معمولاً در حدود ۰.۵-۰.۲ dB)، به ما این اختیار را می دهد که از کدگذاری حسابی^۳ [۳۱] صرف نظر کنیم. در ادامه به توضیح مختصر نحوه عملکرد الگوریتم کدگذاری SPIHT می پردازیم.

در الگوریتم SPIHT به غیر از پایین ترین زیرباند فرکانسی، ضرایب موجود در سایر زیرباندها به آرایه هایی 2×2 تحت عنوان گره های فرزند^۴ یک ضریب از یکی از زیرباندهای سطح تجزیه بالاتر دسته بندی می شوند. همچنین ضرایب پایین ترین زیرباند فرکانسی به آرایه هایی 2×2 تقسیم می شوند، اما همانطور که در شکل ۲-۸ (الف) و (ب) مشاهده می شود، ضریب گوشه بالا-چپ این آرایه ها هیچ

¹ Root Nodes

² Optimal Progressive Transmission

³ Arithmetic Coding

⁴ Offspring

فرزندى ندارد. براساس نمادهای تعريف شده در [۱۲]، مجموعه تمام فرزندان گره (i, j) را با $O(i, j)$ ، مجموعه تمام نسل‌های^۱ گره (i, j) را با $D(i, j)$ (يك مجموعه از نوع "A") و مجموعه تمام نسل‌ها به جز فرزندان بی‌واسطه^۲ گره (i, j) را با $L(i, j)$ (يك مجموعه از نوع "B") نشان می‌دهیم. شكل ۲-۸ (ج) مجموعه‌های $O(i, j)$ ، $D(i, j)$ و $L(i, j)$ را در يك درخت SOT نشان می‌دهد. اطلاعات مهم در سه فهرست منظم ذخیره می‌شوند: فهرست پیکسل‌های غیرمهم (LIP)^۳، فهرست مجموعه‌های غیرمهم (LIS)^۴ و فهرست پیکسل‌های مهم (LSP)^۵. در نخستین گام مقداردهی اولیه این سه فهرست بدین صورت انجام می‌گیرد: ضرایب موجود در پایین‌ترین زیرباند فرکانسی (بالاترین سطح هرم) به LIP اضافه می‌شوند، و از این ضرایب تنها آنهایی که دارای نسل هستند به عنوان مجموعه نوع "A" به LIS اضافه می‌شوند. در ابتدای کار LSP يك فهرست خالی خواهد بود.

فرآیند کدگذاری از پرارزش‌ترین صفحه بيتی شروع شده و به سوی کم‌ارزش‌ترین صفحه بيتی پیش می‌رود. پیشروی کدگذاری در هر صفحه بيتی از طریق سه فهرست LIP، LIS و LSP صورت می‌گیرد. به بیان دیگر در هر صفحه بيتی، واحد کدگذار ابتدا کار را با LIP شروع کرده، در ادامه LIS را در دست گرفته و در انتها با LSP کار خواهد کرد. برای هر ضریب موجود در LIP، يك بیت برای توصیف "اهمیت" آن استفاده می‌شود. اگر ضریبی مهم نباشد در LIP باقی خواهد ماند و بیت بیشتری تولید نخواهد شد، در غیر اینصورت بیت علامت مربوطه تولید شده و آن ضریب به LSP جابجا

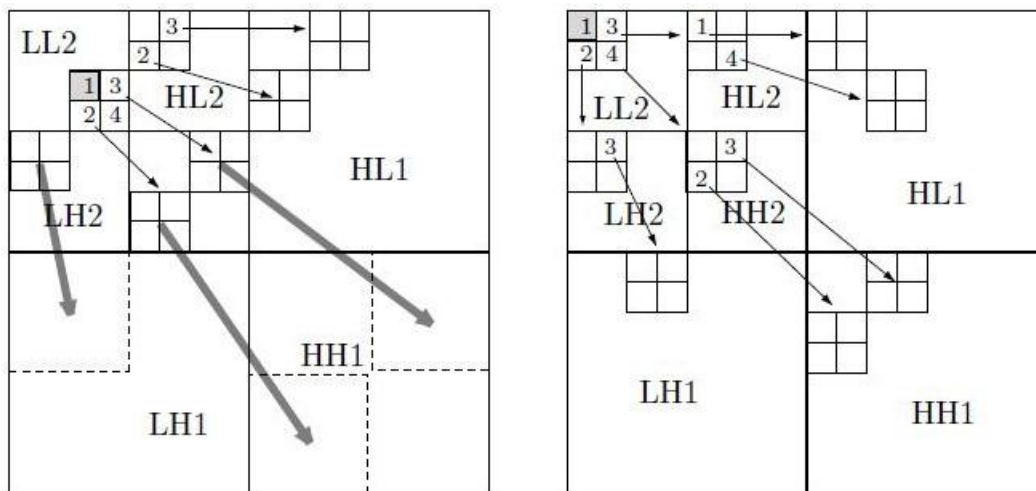
¹ Descendents

² Immediate Offsprings

³ List of Insignificant Pixels

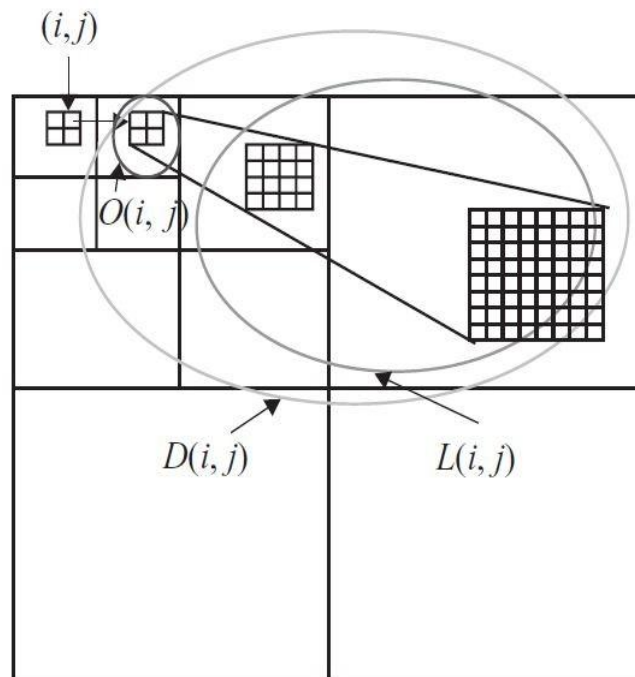
⁴ List of Insignificant Sets

⁵ List of Significant Pixels



(ب)

(الف)



(ج)

شکل ۸-۲: درخت‌های جهت‌دار مکانی و بخش‌بندی مجموعه در SPIHT [۱۶ و ۲۷].

می‌شود. بطور مشابه، هر مجموعه موجود در LIS نیازمند یک بیت برای اطلاعات "اهمیت" است. مجموعه‌های غیرمهم در LIS باقی می‌مانند، در حالیکه مجموعه‌های مهم به زیرمجموعه‌هایی بخش‌بندی می‌شوند. یک مجموعه مهم نوع "A" به یک مجموعه نوع "B" و چهار ضریب فرزند

بخش‌بندی خواهد شد؛ بطوریکه مجموعه نوع “B” به انتهای LIS افزوده شده، در حالیکه چهار ضریب فرزند بلافاصله برای تعیین اهمیت بررسی می‌شوند. یک مجموعه مهم نوع “B” به چهار مجموعه نوع “A” بخش‌بندی خواهد شد؛ بطوریکه تمام آن‌ها به انتهای LIS افزوده می‌شوند. از آنجایی که تمام مجموعه‌های غیرمهم اخیراً تولید شده به انتهای LIS افزوده می‌شوند، آن‌ها نیز در همان صفحه بیتی و به روشی یکسان پردازش خواهند شد تا زمانی که هر زیرمجموعه مهم دقیقاً شامل تنها یک ضریب باشد. در انتها، هر ضریب موجود در LSP، به غیر از آن‌هایی که در صفحه بیتی جاری اضافه شده‌اند، با یک بیت پالایش و تصحیح می‌شوند. سپس الگوریتم کدگذاری روند فوق را برای صفحه بیتی بعدی تکرار می‌کند.

۳-۲-۲ کدگذاری مبتنی بر بلوک صفر

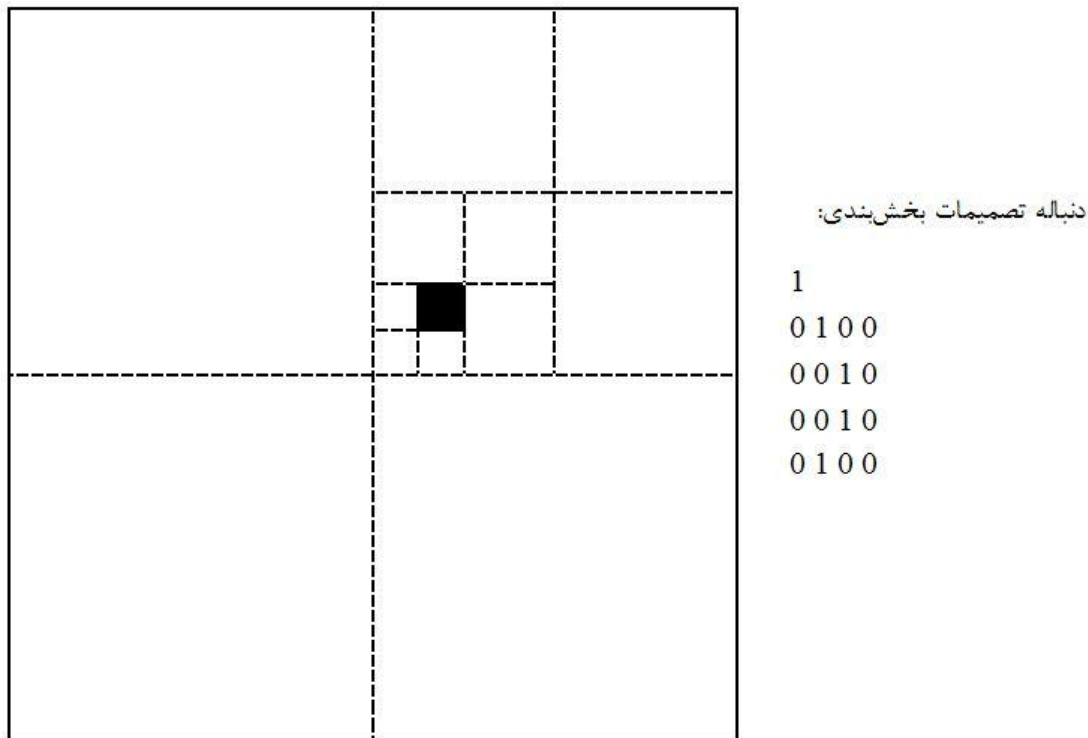
الگوریتم‌های مبتنی بر بلوک صفر تصویر تبدیل‌یافته را به بلوک‌هایی مجاور هم تقسیم کرده و آزمون اهمیت را بر روی این بلوک‌های منحصر بفرد اجرا می‌کنند. مشابه تعریف درخت صفر، در اینجا نیز یک بلوک صفر، بلوکی است که نسبت به یک آستانه مشخص (2^n) دارای هیچ ضریب مهمی (بزرگتری) نباشد. الگوریتم‌های مشهور در این دسته شامل کدگذار بلوک جاسازی‌شده بخش‌بندی مجموعه (SPECK)^۱ [۳۲]، بخش‌بندی سلسله‌مراتبی بلوک زیرباند (SBHP)^۲ [۳۳]، EBCOT [۹] و کدگذاری بلوک صفر جاسازی‌شده EZBC^۳ [۲۸ و ۳۴] می‌باشند.

¹ Set Partitioning Embedded Block

² Subband Block Hierarchical Partitioning

³ Embedded Zero Block Coding

در این الگوریتم‌ها اگر یک بلوک مهم باشد، به چهار زیربلوک با اندازه برابر بخش‌بندی می‌شود که این نوع بخش‌بندی را اصطلاحاً "بخش‌بندی درخت چهارگانه"^۱ نیز می‌نامند. سپس اهمیت هر زیربلوک بطور مجزا سنجیده شده و هر زیربلوک مهم مجدداً با استفاده از بخش‌بندی درخت چهارگانه تقسیم خواهد شد. بدین ترتیب یک بلوک مهم بطور بازگشتی به زیربلوک‌هایی بخش‌بندی می‌شود، و این عمل آنقدر ادامه می‌یابد تا زمانی که ضرایب مهم پیدا شوند. در شکل ۹-۲ می‌توان عملیات



شکل ۹-۲: نمایش بخش‌بندی درخت چهارگانه و کدگذاری در یک ناحیه بلوکی 16×16 . در اینجا آزمون اهمیت برای هر آرایه 2×2 یک ترتیب پیمایش Z را دنبال می‌کند. بیت "۱" در دنباله کدگذاری شده نشان‌دهنده ادامه بخش‌بندی یک گره معین است [۲۸].

¹ Quadtree Partitioning

بخش‌بندی درخت چهارگانه و کدگذاری را در یک ناحیه بلوکی 16×16 مشاهده کرد. مزیت کدگذاری بلوکی این است که بلوک‌های کوچک که نشان‌دهنده نواحی فرکانس بالا هستند، بطور مستقل از نواحی بزرگتر با محتوای فرکانسی-مکانی پایین کدگذاری می‌شوند.

۴-۲-۲ کدگذاری درخت بلوکی موجک

واضح است که الگوریتم‌های کدگذاری درخت صفر مانند SPIHT از طریق ساختار درختی تنها از همبستگی‌های بین‌بندی بهره می‌برند، در حالیکه الگوریتم‌های کدگذاری بلوک صفر مانند SPECK تنها از همبستگی‌های درون‌بندی بهره می‌برند. در [۱۳] الگوریتمی برای کدگذاری تصویر پیشنهاد شده است که ویژگی‌های الگوریتم‌های مبتنی بر درخت صفر و بلوک صفر را با یکدیگر ترکیب می‌کند. این امر باعث بهره‌گیری همزمان این روش از همبستگی درون‌بندی و بین‌بندی ضرایب موجک می‌شود. الگوریتم پیشنهادی تصویر تبدیل‌یافته را به بلوک‌های ضرایبی تقسیم کرده و سپس درخت‌های بلوکی^۱ را با ریشه‌هایی واقع در پایین‌ترین زیربند فرکانسی، به سبک الگوریتم‌های درخت صفر تشکیل می‌دهد. در یک درخت بلوکی، بلوک‌های مهم با استفاده از مفهوم "بخش‌بندی درخت"^۲ مورد استفاده در SPIHT پیدا می‌شوند، در حالیکه ضرایب مهم درون هر بلوک با استفاده از مفهوم "بخش‌بندی درخت چهارگانه" مورد استفاده در SPECK یافت می‌شوند. یک درخت بلوکی مهم بصورت بازگشتی آنقدر بخش‌بندی و تفکیک می‌شود (ترکیبی از بخش‌بندی درختی و بلوکی)، تا زمانی که ضرایب مهم پیدا شوند. این الگوریتم بنا به دلایل آشکار ذکر شده، در اصطلاح "کدگذاری

^۱ Blocktree

^۲ Tree Partitioning

درخت بلوکی موجک^۱ یا به اختصار WBTC نامیده شده است. اگرچه استفاده از کدگذاری حسابی کارایی WBTC را بهبود می بخشد، اما بطور کلی استفاده از آن ضروری نیست. الگوریتم WBTC با بهره گیری همزمان از مزایای الگوریتم های کدگذاری مبتنی بر بلوک و مبتنی بر درخت، به کارایی بالاتر در نرخ بیت های خیلی پایین، سرعت بالاتر در فرآیند کدگذاری و نیز کاهش حافظه مورد نیاز برای ذخیره سازی فهرست های کمکی دست یافته است.

می توان گفت کدگذاری WBTC به نوعی نسخه توسعه یافته SPIHT است. در WBTC ضرایب موجک بصورت بلوک هایی با ابعاد $m \times n$ گروه بندی شده و سپس درخت های بلوکی با ریشه هایی واقع در پایین ترین زیرباند فرکانسی (زیرباند LL بالاترین سطح تجزیه) شکل می گیرند. یک درخت بلوکی، درختی از تمام بلوک های نسل^۱ یک بلوک ریشه^۲ است. این روش سه مزیت نسبت به SPIHT دارد. اولاً این روش تعداد زیادی از درخت های صفر خوشه بندی شده الگوریتم SPIHT را ترکیب می کند، که معمولاً در گذرهای اولیه^۳ رخ می دهد؛ بنابراین درخت های صفری با تعداد عناصر بیشتر ایجاد خواهد شد. دوماً به دلیل ماهیت مبتنی بر بلوک این روش، در مقایسه با تکنیک های مبتنی بر پیکسل^۴، حافظه مورد نیاز برای ذخیره سازی فهرست ها و نیز زمان کدگذاری بطور قابل ملاحظه ای کاهش می یابد.

¹ Descendent Blocks

² Root Block

³ Early Passes

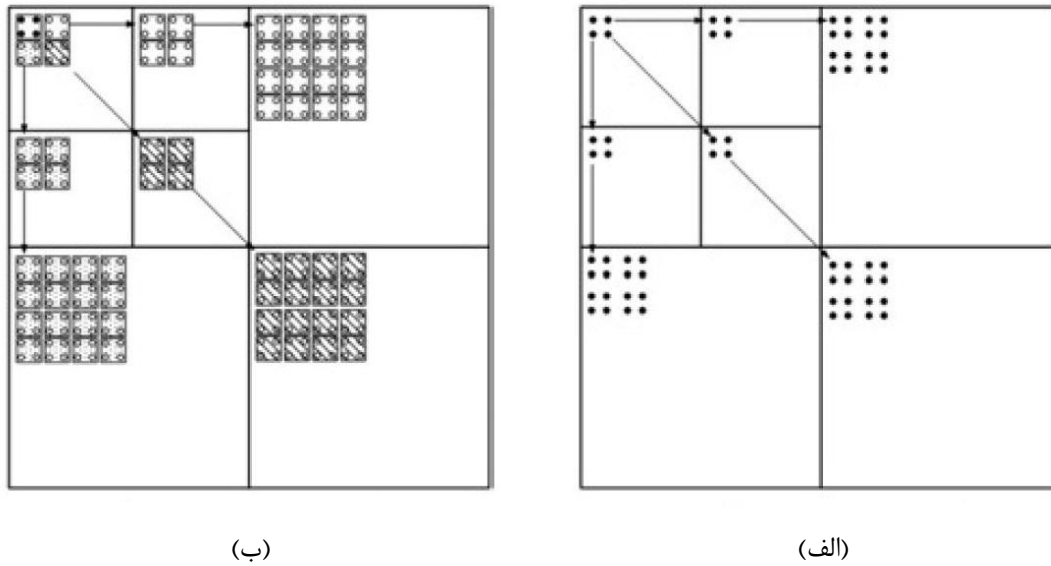
⁴ Pixel-Based

در الگوریتم WBTC به غیر از زیرباند LL بالاترین سطح تجزیه موجک و نیز زیرباندهای LH، HL و HH پایین ترین سطح تجزیه، هر بلوک دارای چهار بلوک فرزند در سطح تجزیه پایین تر (یا بطور معادل در زیرباندهای فرکانسی بالاتر) است، بطوریکه این بلوک های فرزند از نظر مکانی در موقعیت و جهتی مشابه بلوک مربوطه قرار گرفته اند. در زیرباند LL بالاترین سطح تجزیه، در هر گروه 2×2 از بلوک ها، بلوک واقع در گوشه بالا-چپ دارای هیچ نسلی نخواهد بود، و هر یک از سه بلوک دیگر چهار بلوک فرزند در همان سطح تجزیه خواهند داشت؛ که این بلوک های فرزند بسته به جهت مکانی بلوک مربوطه ممکن است در زیرباند LH، HL و یا HH قرار گرفته باشند. هر بلوکی که دارای فرزند باشد اصطلاحاً "بلوک والد"^۱ نامیده می شود. با ایجاد یک درخت بلوکی، تعداد زیادی از درخت های SOT مورد استفاده در SPIHT با یکدیگر ترکیب شده و تنها در یک درخت SOT الگوریتم WBTC قرار می گیرند. شکل ۲-۱۰ ساختار درخت های SOT را در الگوریتم های SPIHT و WBTC نشان می دهد. بطور خاص با اندازه بلوک 2×2 همانطور که در شکل ۲-۱۰ (ب) نشان داده شده است، هر چهار درخت SOT از SPIHT با یکدیگر ترکیب شده و تنها در یک SOT از WBTC قرار می گیرند.

نمادها و اصطلاحات استفاده شده در الگوریتم WBTC عبارتند از:

$B_{k,l}^{m,n}$: یک بلوک دلخواه شامل مجموعه ای از ضرایب موجک بصورت $\{C_{i,j} | k \leq i < k + m, l \leq j < l + n\}$ بیانگر ابعاد بلوک و (k, l) بیانگر مختصات گوشه بالا-چپ بلوک است (مطابق شکل ۲-۱۱ الف)). لازم به ذکر است که هر بلوک توسط مختصات گوشه بالا-چپ خود آدرس دهی می شود.

¹ Parent Block



شکل ۲-۱۰: درخت‌های جهت‌دار مکانی. (الف) استفاده شده در SPIHT. (ب) استفاده شده در WBTC [۱۳].

$H_{k,l}^{m,n}$: یک مجموعه از تمامی بلوک‌های ریشه SOT.

$O_{k,l}^{m,n}$: یک مجموعه از تمامی بلوک‌های فرزند بلوک ریشه $B_{k,l}^{m,n}$ به جز بالاترین و پایین‌ترین سطوح

هرم، $O_{k,l}^{m,n}$ بصورت زیر تعریف می‌شود:

$$O_{k,l}^{m,n} = \{B_{2k,2l}^{m,n}, B_{2k+m,2l}^{m,n}, B_{2k,2l+n}^{m,n}, B_{2k+m,2l+n}^{m,n}\}$$

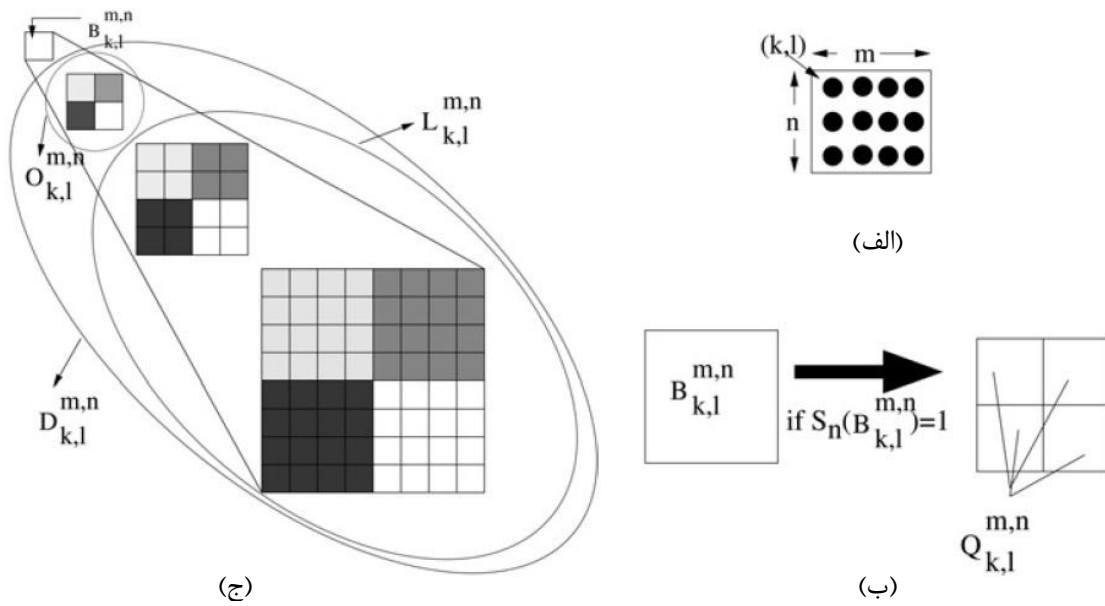
$Q_{k,l}^{m,n}$: یک مجموعه از تمامی بلوک‌های چهارگانه^۱ یک بلوک والد $B_{k,l}^{m,n}$ ، که بصورت زیر تعریف

می‌شود:

$$Q_{k,l}^{m,n} = \{B_{k,l}^{m/2,n/2}, B_{k+m/2,l}^{m/2,n/2}, B_{k,l+n/2}^{m/2,n/2}, B_{k+m/2,l+n/2}^{m/2,n/2}\}$$

$D_{k,l}^{m,n}$: یک مجموعه از تمامی بلوک‌های نسل بلوک ریشه $B_{k,l}^{m,n}$ (یک مجموعه از نوع "A").

^۱ Quad-Blocks



شکل ۲-۱۱: بخش‌بندی درخت چهارگانه و درخت بلوکی در WBTC. (الف) توصیف یک بلوک $B_{k,l}^{m,n}$. (ب) بخش‌بندی درخت چهارگانه یک بلوک $B_{k,l}^{m,n}$. (ج) بخش‌بندی یک SOT بلوکی با بلوک ریشه $B_{k,l}^{m,n}$. [۱۳].

$L_{k,l}^{m,n}$: یک مجموعه از تمامی بلوک‌های نسل بلوک ریشه $B_{k,l}^{m,n}$ ، به جز بلوک‌های فرزند بی‌واسطه (یک مجموعه از نوع "B").

$S_n(T)$: تابع اهمیت در صفحه بیتی n ام که به مجموعه T اعمال می‌شود و اهمیت آن را مورد ارزیابی قرار می‌دهد. این تابع بصورت زیر تعریف می‌شود:

$$S_n(T) = \begin{cases} 1, & \max(\{|C_{i,j}|\}) \geq 2^n, \quad C_{i,j} \in T \\ 0, & \text{otherwise} \end{cases} \quad (۶-۲)$$

در رابطه (۶-۲)، مجموعه T ممکن است یک SOT بلوکی، یک بلوک $m \times n$ از ضرایب موجک و یا یک ضریب موجک تنها باشد.

در WBTC مشابه SPIHT از سه فهرست منظم استفاده می‌شود، فهرست بلوک‌های بی‌اهمیت (LIB)، فهرست مجموعه بلوک‌های بی‌اهمیت (LIBS) و فهرست پیکسل‌های مهم (LSP). این فهرست‌ها حاوی مختصات ضرایب و بلوک‌های غیر ریشه و یا بلوک‌های ریشه درخت‌های SOT

می‌باشند. در گام مقداردهی اولیه، مختصات تمامی بلوک‌های موجود در زیرباند LL بالاترین سطح تجزیه به فهرست LIB اضافه می‌شوند، و تنها آن دسته از بلوک‌های این زیرباند که دارای نسل هستند، به عنوان مجموعه‌های نوع "A" به فهرست LIBS اضافه می‌شوند. فهرست LSP در ابتدای کار خالی خواهد بود.

الگوریتم WBTC نیز مشابه SPIHT و SPECK یک روش کدگذاری مبتنی بر صفحه بیتی است که در هر صفحه بیتی شامل دو فاز دسته‌بندی^۱ و پالایش^۲ است. این روش کدگذاری فرآیندی تکراری بوده، بطوریکه با شروع از پرارزش‌ترین صفحه بیتی و کدگذاری آن، در صفحات بیتی پایین‌تر نیز فرآیند کدگذاری مشابهی تکرار می‌شود. تعداد تکرارها حداکثر برابر $n_{max} + 1$ خواهد بود، که n_{max} شماره پرارزش‌ترین صفحه بیتی بوده و از رابطه زیر بدست می‌آید:

$$n_{max} = \log_2(\max(|C_{i,j}|)) \quad (۷-۲)$$

در رابطه (۷-۲)، i و j سطر i ام و ستون j ام ماتریس ضرایب C بدست آمده از تبدیل موجک است. فرآیند کدگذاری تا زمانی که نرخ بیت مورد نیاز تصویر فشرده شده تحقق نیافته ادامه خواهد یافت، و به محض رسیدن به این نرخ بیت متوقف می‌شود.

در فاز دسته‌بندی، واحد کدگذار ابتدا اهمیت بلوک‌های ضرایب موجود در LIB را با استفاده از رابطه (۶-۲) نسبت به آستانه‌ی جاری ارزیابی می‌کند. رابطه (۶-۲) را آزمون اهمیت نیز می‌گویند. آستانه‌های مورد استفاده در آزمون اهمیت بسته به شماره صفحه بیتی یا همان n تغییر می‌کنند.

¹ Sorting

² Refinement

مقادیری که n می‌تواند اختیار کند از ۰ تا n_{max} خواهد بود. برای هر بلوک موجود در LIB یک بیت برای توصیف اهمیت آن استفاده می‌شود. چنانچه بلوکی مهم نباشد، این بلوک یک بلوک صفر بوده و لذا یک بیت با مقدار "۰" به واحد کدگشا ارسال می‌شود؛ از طرفی بلوک مربوطه در LIB باقی مانده و هیچ بیت اضافه‌ای تولید نخواهد شد. در اینجا اطلاعات غیرمهم $m \times n$ ضریب منحصر بفرد، تنها توسط یک تک بیت "۰" حمل می‌شوند، در حالیکه در SPIHT این بلوک صفر تعداد $m \times n$ بیت "۰" تولید خواهد کرد. این توضیح تا حدودی نحوه بهره‌گیری الگوریتم WBTC از همبستگی درون‌بندی ضرایب را آشکار می‌سازد. حال چنانچه بلوکی مهم باشد، این بلوک یک بلوک غیرصفر بوده و لذا یک بیت با مقدار "۱" به واحد کدگشا ارسال می‌شود. یک بلوک مهم همانطور که در شکل ۲-۱۱ (ب) نشان داده شده است، با استفاده از بخش‌بندی درخت چهارگانه به چهار بلوک مجاور هم بخش‌بندی می‌شود. عملیات تقسیم‌بندی بطور بازگشتی آنقدر تکرار می‌شود تا زمانی که دیگر تقسیم بیشتری نیاز نباشد و یا کوچکترین اندازه ممکن بلوک یعنی یک ضریب منحصر بفرد بدست آمده باشد. در این مرحله اهمیت‌های چهار ضریب بدست آمده بطور مجزا سنجیده می‌شوند. اگر ضریبی غیرمهم باشد، یک بیت با مقدار "۰" فرستاده شده و ضریب مربوطه به عنوان یک بلوک تک ضریب به LIB منتقل می‌شود. در غیراینصورت، چنانچه ضریبی مهم باشد، یک بیت با مقدار "۱" فرستاده شده و بیت علامت آن نیز کدگذاری می‌شود و همچنین ضریب مربوطه به LSP منتقل می‌شود. پس از ارزیابی تمام چهار ضریب منحصر بفرد درون بلوک، بلوک جاری از LIB حذف می‌شود.

سپس واحد کدگذار به بررسی LIBS پرداخته و یک آزمون اهمیت را روی هر مجموعه اجرا می‌کند. مجموعه‌های غیرمهم درون LIBS باقی می‌مانند، در حالیکه مجموعه‌های مهم به زیرمجموعه‌هایی بخش‌بندی می‌شوند. یک مجموعه مهم نوع "A" با بلوک ریشه $B_{k,l}^{m,n}$ مطابق شکل ۲-۱۱ (ج) به یک مجموعه نوع "B" ($L_{k,l}^{m,n}$) و چهار بلوک فرزند ($O_{k,l}^{m,n}$) بخش‌بندی خواهد شد.

مجموعه نوع "B" به انتهای LIBS افزوده شده، در حالیکه چهار بلوک فرزند (هر یک هم‌اندازه با بلوک ریشه $B_{k,l}^{m,n}$) بلافاصله اهمیتشان به همان روشی که برای بلوک‌های درون LIB توضیح داده شد، بررسی می‌شوند. در اینجا نیز هر بلوک صفری منجر به صرفه‌جویی بیتی خواهد شد. یک مجموعه مهم نوع "B" به چهار مجموعه نوع "A" بخش‌بندی خواهد شد؛ و تمام آن‌ها به انتهای LIBS افزوده می‌شوند. از آنجایی که تمام مجموعه‌های غیرمهم اخیراً تولید شده به انتهای LIBS افزوده می‌شوند، آن‌ها نیز در همان سطح آستانه و به روشی یکسان پردازش خواهند شد تا زمانی که هریک از آن‌ها بررسی شده باشد. پس از فاز دسته‌بندی طی یک فاز پالایش ضرایب موجود در LSP، به غیر از آن‌هایی که در صفحه بیتی جاری اضافه شده‌اند، با یک بیت پالایش و تصحیح می‌شوند. سپس الگوریتم کدگذاری روند فوق را با کاهش سطح آستانه جاری با یک ضریب ۲، تکرار می‌کند و این فرآیند تکراری تا زمانی که نرخ بیت مطلوب حاصل شود ادامه می‌یابد. باید توجه داشت که با یک اندازه بلوک 1×1 (تک ضریب)، الگوریتم WBTC به الگوریتم SPIHT تبدیل می‌شود. بنابراین، SPIHT یک حالت خاص از الگوریتم WBTC است.

۳-۲ کدگذاری ROI در حوزه موجک

تکنیک کدگذاری مبتنی بر ناحیه مطلوب (ROI) تصویر، به معنی فشردسازی بی‌اتلاف و یا کم‌اتلاف نواحی مطلوب درون یک تصویر، و فشردسازی نواحی غیرمطلوب با اتلاف بیشتر است. براساس این ایده می‌توان علاوه بر دستیابی به یک نرخ فشردسازی بالا، اطلاعات مهم تصویر را نیز بدون اتلاف و یا با اتلاف کم حفظ کرد [۳۵].

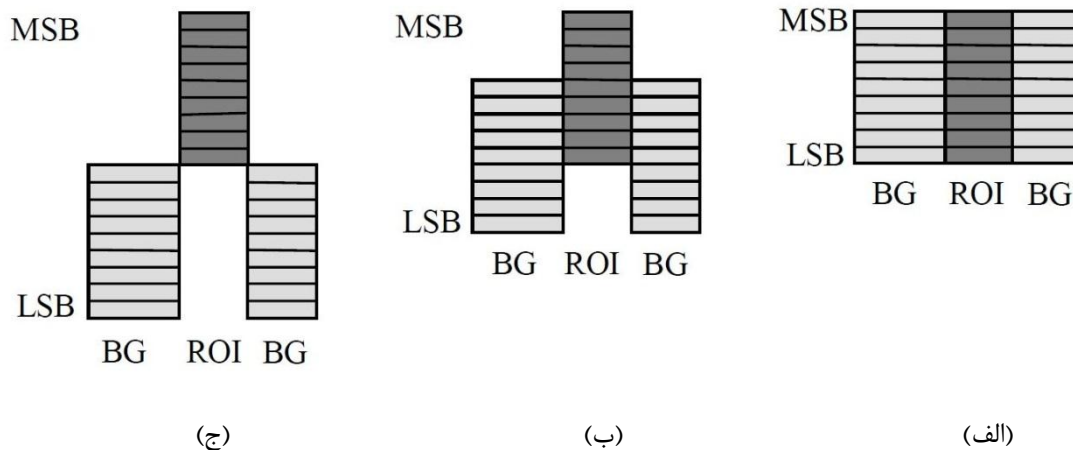
دو روش کدگذاری ROI در استاندارد JPEG2000 معرفی شده است: روش انتقال ماکزیمم (Maxshift) و روش مقیاس گذاری کلی [۳۶]. هر دوی این روش‌ها مبتنی بر انتقال رو به پایین^۱ بیت‌های ضرایب پس‌زمینه (BG) از پرارزش‌ترین صفحه بیتی (MSB) به سمت کم ارزش‌ترین صفحه بیتی (LSB)، و در نتیجه قرار گرفتن بیت‌های ضرایب مربوط به ROI در صفحه‌های بیتی بالاتر نسبت به BG هستند. بنابراین ضرایب ROI در کدگذاری صفحه بیتی جاسازی شده می‌توانند با اولویت بالاتری کدگذاری شوند. در واحد کدگشا، صفحه‌های بیتی بازسازی می‌شوند و قبل از اعمال تبدیل موجک معکوس، ضرایب غیر ROI با انتقال رو به بالا^۲ به صفحه‌های بیتی اصلی و اولیه خود برگردانده می‌شوند. اگر قبل از اینکه تصویر بطور کامل کدگذاری/کدگشایی شود، دنباله بیت کدگذاری شده برش خورد یا فرآیند کدگذاری/کدگشایی متوقف شود، نواحی ROI کیفیت بالاتری نسبت به BG خواهند داشت. شکل ۲-۱۲ نحوه جابجایی صفحه‌های بیتی را در روش‌های کدگذاری ROI مورد استفاده در استاندارد JPEG2000 نشان می‌دهد.

۲-۳-۱ مقیاس گذاری کلی

توصیفی از روش مقیاس گذاری کلی در شکل ۲-۱۲ (ب) به تصویر کشیده شده است. در این شکل مقدار مقیاس گذاری صفحه‌های بیتی با S نشان داده شده است. در روش مقیاس گذاری کلی، باید مقدار مقیاس گذاری و اطلاعات شکل نواحی ROI نیز به دنباله کدگذاری شده پیوست شوند. روش مقیاس گذاری کلی سه اشکال و نقص بزرگ دارد. اول اینکه در این روش نمی‌توان به راحتی با

¹ Down-Shifting

² Up-Shifting



شکل ۲-۱۲: روش‌های کدگذاری ROI در استاندارد JPEG2000. (الف) بدون مقیاس‌گذاری. (ب) روش

مقیاس‌گذاری کلی، $S = 4$. (ج) روش Maxshift، $S = 9$. [۳۵]

زیربافتهای مختلف موجب به طرق مختلف برخورد کرد؛ که گاهی اوقات این مسئله یکی از خواسته‌های کاربران است. دوماً، این روش نیازمند کدگذاری و ارسال اطلاعات شکل نواحی ROI به واحد کدگشا است. این امر بطور قابل ملاحظه‌ای پیچیدگی پیاده‌سازی کدگذار/کدگشا را افزایش می‌دهد. سوم اینکه اگر ROI با شکل‌های دلخواه مطلوب باشد، آنگاه کدگذاری شکل ROI تعداد بیت‌های زیادی را مصرف خواهد کرد، و در نتیجه کارایی کلی کدگذاری بنحو چشمگیری کاهش خواهد یافت. برای حل این مشکل بخش ۲^۱ استاندارد JPEG2000 [۳۷]، شکل‌های ROI را به مستطیل و بیضی محدود کرد، چرا که این شکل‌ها را می‌توان با تعداد بیت‌های کمی کدگذاری کرد و واحد کدگشا نیز به طریقی ساده قادر به بازسازی ماسک ROI خواهد بود. هرچند این ترفند دامنه کاربرد کدگذاری ROI را محدود می‌کند.

¹ Part 2

۲-۳-۲ انتقال ماکزیمم

یک راه حل بسیار موثر برای جبران نقایص مقیاس گذاری کلی، استفاده از روش انتقال ماکزیمم است، که در بخش ۱ استاندارد JPEG2000 معرفی شد و مورد استفاده قرار گرفت [۳۸]. روش Maxshift را می توان حالت خاصی از روش مقیاس گذاری کلی در نظر گرفت، که در آن مقدار مقیاس گذاری آنقدر بزرگ است که هیچ همپوشانی میان صفحه های بیتی ROI و BG وجود ندارد. بنابراین در روش Maxshift مقدار مقیاس گذاری S باید شرط زیر را ارضا کند:

$$S \geq \max(M_b) \quad (۸-۲)$$

در رابطه (۸-۲) $\max(M_b)$ بیشترین تعداد صفحات بیتی اندازه ضرایب پس زمینه است. پس از مقیاس گذاری، تمامی بیت های با ارزش مربوط به ROI در صفحات بیتی بالاتری نسبت به تمام بیت های با ارزش مربوط به BG قرار خواهند داشت. شکل ۲-۱۲ (ج) انتقال صفحه بیتی را در روش Maxshift نشان می دهد. بنابراین در روش Maxshift شکل ROI برای واحد کدگشا مفهوم بوده، و لذا کدگذاری ROI با شکل دلخواه قابل پشتیبانی است. همچنین این روش می تواند به راحتی با زیرباندهای موجک مختلف به طرق مختلف رفتار کند. محدودیت بزرگ روش Maxshift این است که نمی تواند به راحتی اهمیت نسبی میان نواحی ROI و BG را با تنظیم مقدار مقیاس گذاری کنترل کند. این بدان معنی است که در تمامی زیرباندهایی که میان ROI و BG تمایز ایجاد شده است، هیچ اطلاعاتی راجع به ضرایب غیر ROI نمی تواند دریافت شود، تا زمانیکه همه جزئیات ضرایب ROI بطور کامل کدگشایی شده باشند، حتی اگر این جزئیات نویز تصادفی نامحسوس و غیر قابل مشاهده باشند.

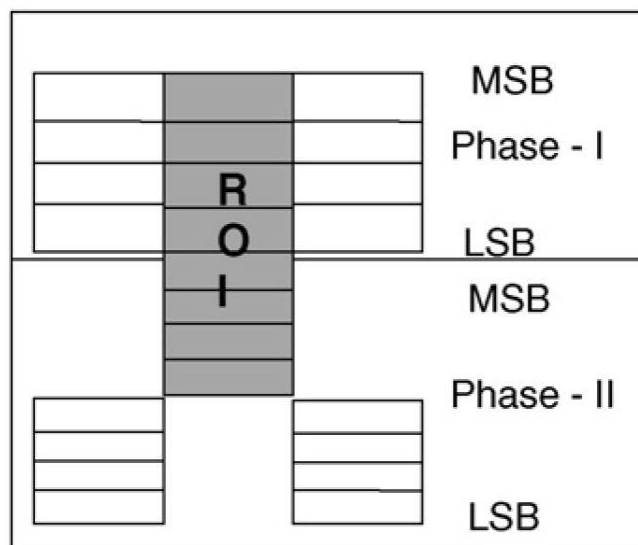


روش های پیشنهادی

۳-۱ روش پیشنهادی اول

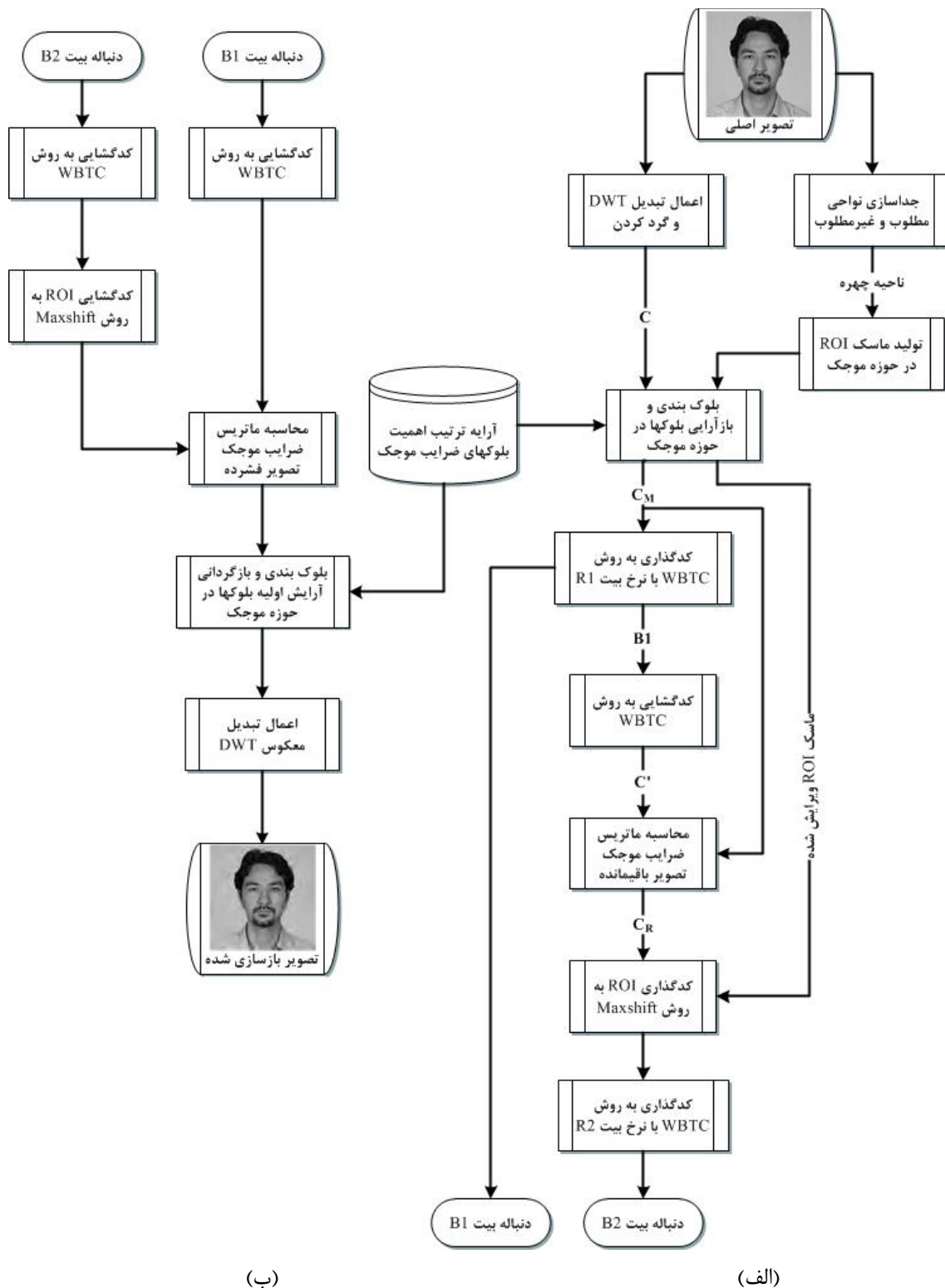
در روش پیشنهادی اول از تکنیک کدگذاری ROI ارائه شده در [۱۱] برای بهبود کیفیت ناحیه چهره در نرخ‌بیت‌های پایین استفاده شده است. در این تکنیک با بهره‌گیری از روش Maxshift و دو مرحله کدگذاری تصویر، مشکل کنترل اهمیت نسبی میان نواحی مطلوب و غیرمطلوب در روش Maxshift تا حدود زیادی برطرف شده است. در این تکنیک ابتدا با کدگذاری تصویر چهره در یک نرخ‌بیت مشخص و سپس بازسازی آن، یک نسخه کلی و تقریبی از تصویر چهره تولید می‌شود. در مرحله دوم با کدگذاری تصویر باقیمانده چهره در یک نرخ‌بیت دیگر و سپس بازسازی و افزودن آن به تصویر تقریبی چهره، کیفیت ناحیه ویژگی‌های چهره در تصویر بازسازی‌شده بهبود می‌یابد. تصویر باقیمانده از تفاضل تصویر اصلی و تصویر تقریبی بدست آمده، و قبل از کدگذاری این تصویر، با استفاده از روش Maxshift اولویت کدگذاری ناحیه ویژگی‌های چهره ارتقا می‌یابد. در حقیقت در این تکنیک، در مرحله اول یک نرخ‌بیت یکسان و در نتیجه اهمیتی برابر به سرتاسر تصویر شامل نواحی مطلوب و غیرمطلوب اختصاص داده می‌شود. به بیان دیگر در مرحله اول ناحیه غیرمطلوب نیز به موازات ناحیه مطلوب با نرخ‌بیتی برابر کدگذاری می‌شود. سپس در مرحله دوم با استفاده از روش Maxshift ابتدا ناحیه مطلوب کدگذاری شده و پس از پایان آن در صورتی که نرخ‌بیت بیشتری موجود باشد، عملیات کدگذاری با ناحیه غیرمطلوب ادامه خواهد یافت. شکل ۳-۱ تکنیک کدگذاری ROI ارائه شده در [۱۱] را نشان می‌دهد.

بنابراین با استفاده از این تکنیک می‌توان اهمیت و کیفیت نسبی میان نواحی مطلوب و غیرمطلوب را کنترل کرد؛ بدین ترتیب که ابتدا تصویر اصلی بسته به کیفیت مورد نیاز ناحیه غیرمطلوب در نرخ‌بیتی مناسب کدگذاری شده، و سپس ناحیه مطلوب در تصویر باقیمانده با استفاده از روش Maxshift با توجه به کیفیت مورد نیاز در نرخ‌بیت مناسب کدگذاری می‌شود. تلفیق این



شکل ۳-۱: تکنیک کدگذاری ROI ارائه شده در [۱۱].

تکنیک با الگوریتم کدگذاری WBTC و تکنیک بازآرایی بلوک‌های ضرایب موجک در روش پیشنهادی اول باعث شده است که کیفیت ناحیه چهره در نرخ‌بیت‌های پایین نسبت به روش [۱۱] بهبود یابد، و همچنین کارایی نرخ-عوجاج آن در نرخ‌بیت‌های پایین از روش‌هایی همچون استاندارد JPEG2000 و SPIHT بالاتر باشد. علت این امر کارایی بالای الگوریتم WBTC در نرخ‌بیت‌های پایین و نیز بهبود عملکرد این الگوریتم با استفاده از تکنیک بازآرایی بلوک‌های ضرایب موجک است. از طرفی به دلیل سرعت بالای کدگذاری و کدگذاری الگوریتم WBTC، در روش پیشنهادی اول بازسازی تصاویر در نرخ‌بیت‌های پایین با سرعت بالایی انجام گرفته و سرعت فشرده‌سازی نیز در حد مناسبی است. بنابراین این روش مناسب کاربردهایی است که نیازمند فشرده‌سازی و بازسازی سریع و نیز ارسال و دریافت سریع تصاویر پرسنلی در نرخ‌بیت‌های پایین هستند، که از آن جمله می‌توان سیستم‌های تشخیص هویت بلادرنگ را مثال زد. نمودار بلوکی فرآیند فشرده‌سازی و بازسازی روش پیشنهادی اول در شکل ۳-۲ نشان داده شده است.



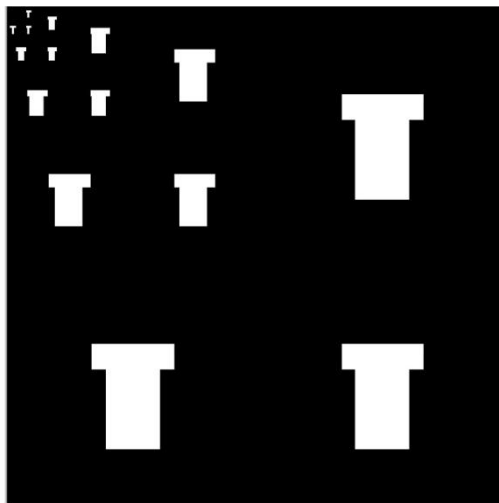
شکل ۳-۲: نمودار بلوکی روش پیشنهادی اول. (الف) فرآیند فشرده‌سازی. (ب) فرآیند بازسازی.

۳-۱-۱ فرآیند فشرده‌سازی

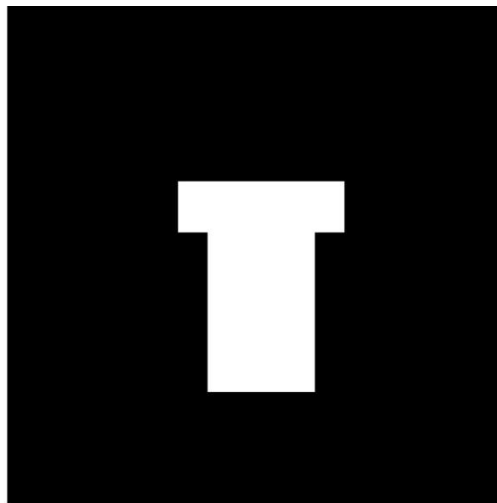
در فرآیند فشرده‌سازی همانطور که در شکل ۳-۲(الف) مشاهده می‌شود، در مرحله اول از تصویر پرسنلی تبدیل موجک گسسته گرفته شده و سپس ضرایب موجک حاصل به نزدیکترین مقدار صحیح گرد می‌شوند. این عمل گرد کردن ضرایب معادل یک چندی‌سازی یکنواخت^۱ با گام "۰.۵" است، که کمترین خطا را برای دستیابی به یک نمایش صحیح از ضرایب نتیجه می‌دهد. ماتریس ضرایب موجک حاصل از اعمال تبدیل DWT و گرد کردن را C می‌نامیم. به موازات این مرحله و یا پس از آن باید نواحی مطلوب و غیرمطلوب تصویر پرسنلی جداسازی شده و ماسک ROI در حوزه موجک تولید شود. در روش پیشنهادی اول تنها ناحیه چهره شامل چشم‌ها، ابروها، بینی، دهان و حتی چانه به عنوان ناحیه مطلوب در نظر گرفته شده است، چرا که این ناحیه بیشترین اهمیت و تاثیر را در شناسایی صحیح هویت فرد دارد. البته لازم به ذکر است که در نظر گرفتن سایر نواحی به عنوان نواحی غیرمطلوب به معنای بی‌اهمیت بودن این نواحی و لذا عدم توجه و اختصاص نرخ‌بیت به آن‌ها نیست؛ بلکه به معنای اهمیت کمتر این نواحی در مقایسه با ناحیه چهره و در نتیجه اختصاص نرخ‌بیتی کمتر به آن‌هاست. برای جداسازی نواحی مطلوب و غیرمطلوب از یک تکنیک پیشنهادی سریع مبتنی بر عملیات لبه‌گیری و شکل‌شناسی استفاده شده است که در بخش ۳-۴ به تشریح آن خواهیم پرداخت.

منظور از ماسک ROI در حوزه مکان یک تصویر دودویی هم‌اندازه با تصویر اصلی است که در آن پیکسل‌های متعلق به ناحیه مطلوب با مقدار "۱" مشخص شده‌اند، و پیکسل‌های سایر نواحی مقدار "۰" دارند. شکل ۳-۳(الف) یک ماسک ROI را در حوزه مکان نشان می‌دهد. بطور مشابه ماسک ROI

¹ Uniform Quantization



(ب)



(الف)

شکل ۳-۳: ماسک ROI. (الف) در حوزه مکان. (ب) در حوزه موجک با ۵ سطح تجزیه.

در حوزه موجک نیز تصویری دودویی هم‌اندازه با تصویر اصلی است که در آن پیکسل‌های متعلق به ناحیه مطلوب در زیرباندهای مختلف تجزیه موجک با مقدار "۱" مشخص شده‌اند. در حقیقت ماسک ROI در حوزه موجک، مجموعه‌ای از نسخه‌های کاهش بعد یافته از ماسک ROI در حوزه مکان است، که همانند چیدمان زیرتصاویر متناظر با زیرباندهای تصویر تجزیه‌شده موجک^۱ در کنار یکدیگر قرار گرفته‌اند. با یک زیرنمونه‌برداری از مرتبه ۲ در جهت‌های افقی و عمودی روی ماسک ROI در حوزه مکان، می‌توان یک نسخه کاهش بعد یافته با ضریب ۲ از این ماسک تولید کرد. با تکرار این روند روی نسخه‌های کاهش بعد یافته حاصل می‌توان به نسخه‌هایی با ابعاد کوچکتر دست یافت. شکل ۳-۳(ب) ماسک ROI متناظر با شکل ۳-۳(الف) را در حوزه موجک و با ۵ سطح تجزیه نشان می‌دهد.

¹ Wavelet Decomposed Image

در مرحله بعد ماتریس C و نیز ماسک ROI در حوزه موجک، به بلوک‌هایی مجاور هم و مجزا با ابعاد $n \times n$ تقسیم شده، و سپس آرایش این بلوک‌ها مطابق آرایه ترتیب اهمیت بدست آمده در فاز شبه-آموزش و ترتیب جستجوی ضرایب مهم در الگوریتم WBTC تغییر می‌کند. در حقیقت هدف از این مرحله تمرکز بیشتر انرژی تصویر و بطور معادل تمرکز ضرایب مهم‌تر در زیرباندهای فرکانسی پایین‌تر است. این امر باعث افزایش کارایی کدگذاری WBTC و بطور کلی الگوریتم‌های کدگذاری ضرایب موجک می‌گردد، چرا که در این حالت زمان و نرخ بیت کمتری صرف یافتن ضرایب مهم یا همان ضرایب بزرگتر خواهد شد. در بخش ۳-۶ توضیح بیشتری راجع به تاثیر بازآرایی بلوک‌های ضرایب موجک در کارایی الگوریتم WBTC، و نیز نحوه بدست آوردن آرایه ترتیب اهمیت بلوک‌های ضرایب در فاز شبه-آموزش داده خواهد شد. پس از بازآرایی بلوک‌های ضرایب موجک ماتریس C ، ماتریس حاصل به نام C_M توسط روش WBTC با نرخ بیت $R1$ bpp کدگذاری می‌شود. دنباله بیت حاصل به نام $B1$ ذخیره و یا به واحد کدگشا فرستاده می‌شود. به موازات، دنباله بیت $B1$ کدگشایی می‌شود تا ماتریس دیگری به نام C' تولید شود. از آنجا که در نرخ بیت‌های پایین C' حاوی ضرایب زیرباندهای فرکانسی پایین است، چنانچه آرایش بلوک‌های ضرایب آن به حالت اولیه برگردانده شده و سپس از آن تبدیل موجک معکوس گرفته شود یک تصویر کلی و تقریبی از چهره بدست خواهد آمد. در مرحله بعد همانطور که در شکل ۳-۲(الف) نشان داده شده است، یک ماتریس باقیمانده،

$C_R = C_M - C'$ محاسبه می‌شود. سپس با استفاده از روش Maxshift و ماسک ROI ویرایش شده حاصل از مرحله بازآرایی بلوک‌ها، صفحات بیتی ضرایب موجک مربوط به ناحیه چهره در ماتریس C_R ، به بالا منتقل شده، و توسط الگوریتم WBTC با نرخ بیت $R2$ bpp کدگذاری می‌شوند. دنباله بیت حاصل از این مرحله به عنوان دنباله بیت $B2$ ذخیره و یا به واحد کدگشا فرستاده می‌شود. لازم به ذکر است که مقادیر نرخ بیت‌های $R1$ و $R2$ باید به گونه‌ای انتخاب شوند که مجموع آن‌ها برابر کل

نرخ‌بیت مورد نیاز یا همان R باشد، یعنی $R = R1 + R2$. از طرفی به منظور کسب کیفیت بهتر در سرتاسر تصویر، از کل نرخ‌بیت مورد نیاز R برای تصویر فشرده‌شده، سهم بیشتری به $R1$ برای کدگذاری تصویر تقریبی چهره نسبت به $R2$ برای کدگذاری تصویر باقیمانده چهره اختصاص داده می‌شود، چرا که تصویر تقریبی کیفیت سرتاسر تصویر را تحت‌الشعاع قرار می‌دهد در حالیکه تصویر باقیمانده تنها روی کیفیت ناحیه چهره تأثیر می‌گذارد.

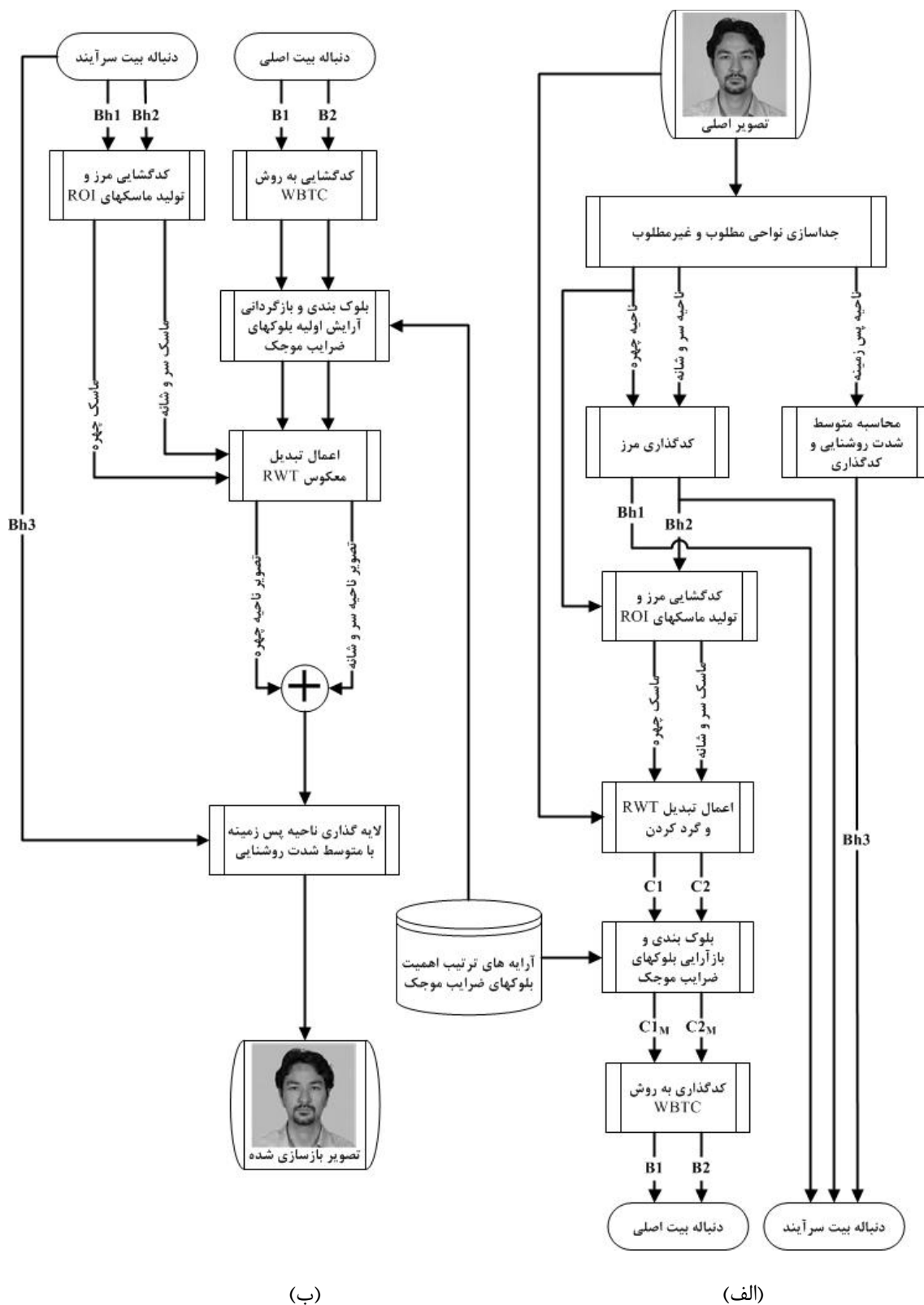
۲-۱-۳ فرآیند بازسازی

در فرآیند بازسازی دنباله‌های بیتی $B1$ و $B2$ بصورت جداگانه با استفاده از الگوریتم WBTC کدگذاری شده و ماتریس‌های ضرایب حاصل با یکدیگر جمع می‌شوند. البته قبل از جمع این دو ماتریس، بایستی ابتدا ناحیه ROI در ماتریس ضرایب حاصل از کدگذاری $B2$ ، به روش Maxshift کدگذاری شود. این کار با انتقال رو به پایین صفحات بیتی ضرایب موجک مربوط به ناحیه چهره محقق می‌شود. لازم به ذکر است که در اینجا نیازی به ماسک ROI در حوزه موجک نیست، چرا که ضرایب موجک ناحیه ROI به راحتی توسط اندازه‌شان قابل تشخیص هستند. در مرحله بعد آرایش بلوک‌های ضرایب موجک با استفاده از آرایه ترتیب اهمیت بدست آمده در فاز شبه-آموزش به حالت اولیه برگردانده می‌شود. در انتها با اعمال تبدیل موجک معکوس به ماتریس ضرایب، تصویر بازسازی‌شده بدست می‌آید. شکل ۲-۳ (ب) مراحل مختلف فرآیند بازسازی تصویر را نشان می‌دهد.

۲-۳ روش پیشنهادی دوم

در روش پیشنهادی دوم از تکنیک جدیدی برای فشرده‌سازی مبتنی بر ROI تصاویر چهره پرسنلی بهره گرفته شده است. در این روش با استفاده از تبدیل موجک مبتنی بر ناحیه (RWT)،

فشرده‌سازی نواحی مطلوب و غیرمطلوب بصورت کاملاً مجزا انجام شده است. بنابراین در این روش می‌توان هر ناحیه را دقیقاً با نرخ بیت دلخواه و مورد نیاز فشرده و بازسازی کرد؛ ویژگی‌ای که در روش پیشنهادی اول و روش [۱۱] به سختی امکان‌پذیر است. یک عیب این روش نیاز تبدیل RWT و معکوس آن به ماسک نواحی مطلوب است، که در نتیجه کدگذاری مرز نواحی مطلوب و ارسال آن به واحد کدگشا را ضروری می‌کند. در این روش برای کدگذاری مرز نواحی چهره و سر و شانه به ترتیب از تکنیک‌هایی مبتنی بر اختلاف فاصله رئوس چندضلعی و چندجمله‌ای درونیاب بهره گرفته‌ایم که در بخش ۳-۵ به اختصار توضیح داده خواهند شد. از طرفی سرعت فشرده‌سازی و بازسازی تصاویر در این روش نسبت به روش پیشنهادی اول کمتر است، چرا که پیچیدگی محاسباتی در این روش افزایش یافته است، و علت آن را می‌توان فشرده‌سازی نواحی بصورت مجزا، اضافه شدن تعداد مراحل بازآرایی بلوک‌های ضرایب بسته به تعداد نواحی مطلوب و نیز اضافه شدن مراحل کدگذاری و کدگشایی مرز نواحی مطلوب دانست. البته در این روش امکان فشرده‌سازی و بازسازی نواحی مختلف تصویر بصورت موازی وجود دارد، که در صورت استفاده از این قابلیت سرعت فشرده‌سازی و بازسازی بسته به تعداد نواحی چند برابر خواهد شد. نتایج شبیه‌سازی بیانگر کارایی بالاتر این روش در ناحیه چهره و در نرخ بیت‌های پایین در مقایسه با روش پیشنهادی اول، روش [۱۱]، استاندارد JPEG2000 و نیز SPIHT است. شکل ۳-۴ نمودار بلوکی مراحل مختلف فرآیند فشرده‌سازی و بازسازی تصاویر را در روش پیشنهادی دوم نشان می‌دهد.



شکل ۳-۴: نمودار بلوکی روش پیشنهادی دوم. (الف) فرآیند فشرده‌سازی. (ب) فرآیند بازسازی.

۳-۲-۱ فرآیند فشرده‌سازی

در روش پیشنهادی دوم تصویر چهره پرسنلی به دو ناحیه مطلوب چهره و سر و شانه و یک ناحیه غیرمطلوب پس‌زمینه تقسیم می‌شود. در این میان ناحیه چهره نسبت به ناحیه سر و شانه اهمیت بالاتری دارد، چرا که این ناحیه شامل ویژگی‌های مهم چهره از قبیل چشم‌ها، بینی و دهان است که نقش مهمی را در تشخیص صحیح هویت فرد بعهدہ دارند. در اینجا نیز مشابه روش پیشنهادی اول، برای جداسازی نواحی مطلوب و غیرمطلوب از تکنیک پیشنهادی ارائه شده در بخش ۳-۴ استفاده شده است. در فرآیند فشرده‌سازی مطابق شکل ۳-۴(الف) پس از جداسازی نواحی مطلوب و غیرمطلوب، مرز نواحی چهره و سر و شانه بصورت مجزا با استفاده از تکنیک‌های ارائه شده در بخش ۳-۵، با نرخ‌بیتی مشخص کدگذاری شده و به ترتیب دنباله بیت‌های $Bh1$ و $Bh2$ را تولید می‌کنند. این دنباله بیت‌ها به عنوان دنباله بیت سرآیند در حافظه ذخیره و یا به واحد کدگشا ارسال می‌شوند. علت این امر نیاز واحد کدگشا به دانستن مرز نواحی مطلوب و بازسازی ماسک‌های ROI متناظر با این نواحی، به منظور استفاده در مرحله اعمال تبدیل معکوس RWT است. نرخ‌بیت کدگذاری مرز نواحی مطلوب در واحد کدگذار، تعیین کننده دقت بازسازی این مرزها و ماسک‌های ROI متناظر با آن‌ها در واحد کدگشا خواهد بود. از طرفی برای عملکرد صحیح تبدیل RWT و معکوس آن، باید ماسک‌های ROI مورد استفاده در واحدهای کدگذار و کدگشا دقیقاً یکسان باشند. این عوامل سبب می‌شوند که در واحد کدگذار نیز مشابه واحد کدگشا ماسک‌های ROI از طریق کدگشایی و بازسازی مرز نواحی مطلوب تولید شوند تا بدین ترتیب یکسان بودن ماسک‌ها در هر دو واحد تضمین شود.

در مرحله بعدی فرآیند فشرده‌سازی، تبدیل RWT با استفاده از ماسک‌های ROI بدست آمده در مرحله قبل بر روی هر یک از نواحی مطلوب بصورت مجزا اعمال گشته و ضرایب موجک حاصل به نزدیکترین مقادیر صحیح گرد می‌شوند. ماتریس‌های ضرایب موجک بدست آمده برای دو ناحیه چهره

و سر و شانه را به ترتیب $C1$ و $C2$ می‌نامیم. در مرحله بعدی مشابه روش پیشنهادی اول ماتریس‌های $C1$ و $C2$ به بلوک‌هایی $n \times n$ تقسیم شده و این بلوک‌ها مطابق آرایه‌های ترتیب اهمیت بدست آمده در فاز شبه-آموزش و الگوی جستجوی ضرایب مهم در الگوریتم WBTC بازآرایی می‌شوند. لازم به ذکر است که بازآرایی بلوک‌های ضرایب هر یک از ماتریس‌های $C1$ و $C2$ بصورت مستقل و به ترتیب با استفاده از آرایه ترتیب اهمیت مخصوص ناحیه چهره و مخصوص ناحیه سر و شانه صورت می‌گیرد. ماتریس‌های ضرایب موجک بازآرایی شده مربوط به نواحی چهره و سر و شانه به ترتیب $C1_M$ و $C2_M$ نامیده می‌شوند. در انتها ماتریس‌های $C1_M$ و $C2_M$ هر کدام با نرخ‌بیتی مشخص و متمایز با استفاده از روش WBTC کدگذاری شده و به ترتیب دنباله بیت‌های $B1$ و $B2$ تولید می‌شوند. این دنباله بیت‌ها به عنوان دنباله بیت اصلی در حافظه ذخیره و یا به واحد کدگشا ارسال می‌شوند. مراحل که تاکنون شرح داده شد، مربوط به کدگذاری و فشرده‌سازی نواحی مطلوب چهره و سر و شانه بود. در مورد ناحیه غیرمطلوب پس‌زمینه، کدگذاری به روش بسیار ساده‌ای انجام می‌شود. در این روش متوسط شدت روشنایی ناحیه پس‌زمینه محاسبه شده و پس از گرد کردن آن به نزدیکترین مقدار صحیح، نمایش دودویی آن با نام $Bh3$ به دنباله بیت سرآیند افزوده می‌شود. علت انتخاب این روش کدگذاری برای ناحیه پس‌زمینه، اطلاعات بسیار اندک نهفته در این ناحیه و نبود هیچ اطلاعات مهمی برای تعیین هویت فرد است.

۲-۲-۳ فرآیند بازسازی

در فرآیند بازسازی ابتدا دنباله‌های بیتی $B1$ و $B2$ بصورت جداگانه با استفاده از الگوریتم WBTC کدگشایی شده، و سپس آرایش بلوک‌های هریک از ماتریس‌های ضرایب موجک حاصل بصورت مستقل با استفاده از آرایه ترتیب اهمیت مخصوص خود به حالت اولیه برگردانده می‌شود. مرحله بعد

اعمال تبدیل معکوس RWT بر روی ماتریس‌های ضرایب موجک حاصل از مرحله قبل است. این مرحله نیازمند ماسک‌های ROI مربوط به نواحی چهره و سر و شانه است. بنابراین قبل از انجام این مرحله بایستی دنباله بیت‌های سرآیند $Bh1$ و $Bh2$ کدگذاری شده و به ترتیب ماسک‌های چهره و سر و شانه تولید شوند. پس از تولید ماسک‌های ROI، تبدیل معکوس RWT با استفاده از این ماسک‌ها بصورت مجزا بر روی ماتریس‌های ضرایب موجک اعمال شده و تصاویر نواحی مطلوب بازسازی شده بدست می‌آیند. در انتها با جمع دو تصویر ناحیه چهره و ناحیه سر و شانه بدست آمده از مرحله قبل، و سپس لایه‌گذاری ناحیه پس‌زمینه با متوسط شدت روشنایی حاصل از کدگذاری $Bh3$ ، تصویر بازسازی شده بدست می‌آید. شکل ۳-۴ (ب) مراحل مختلف فرآیند بازسازی تصویر را در روش پیشنهادی دوم نشان می‌دهد.

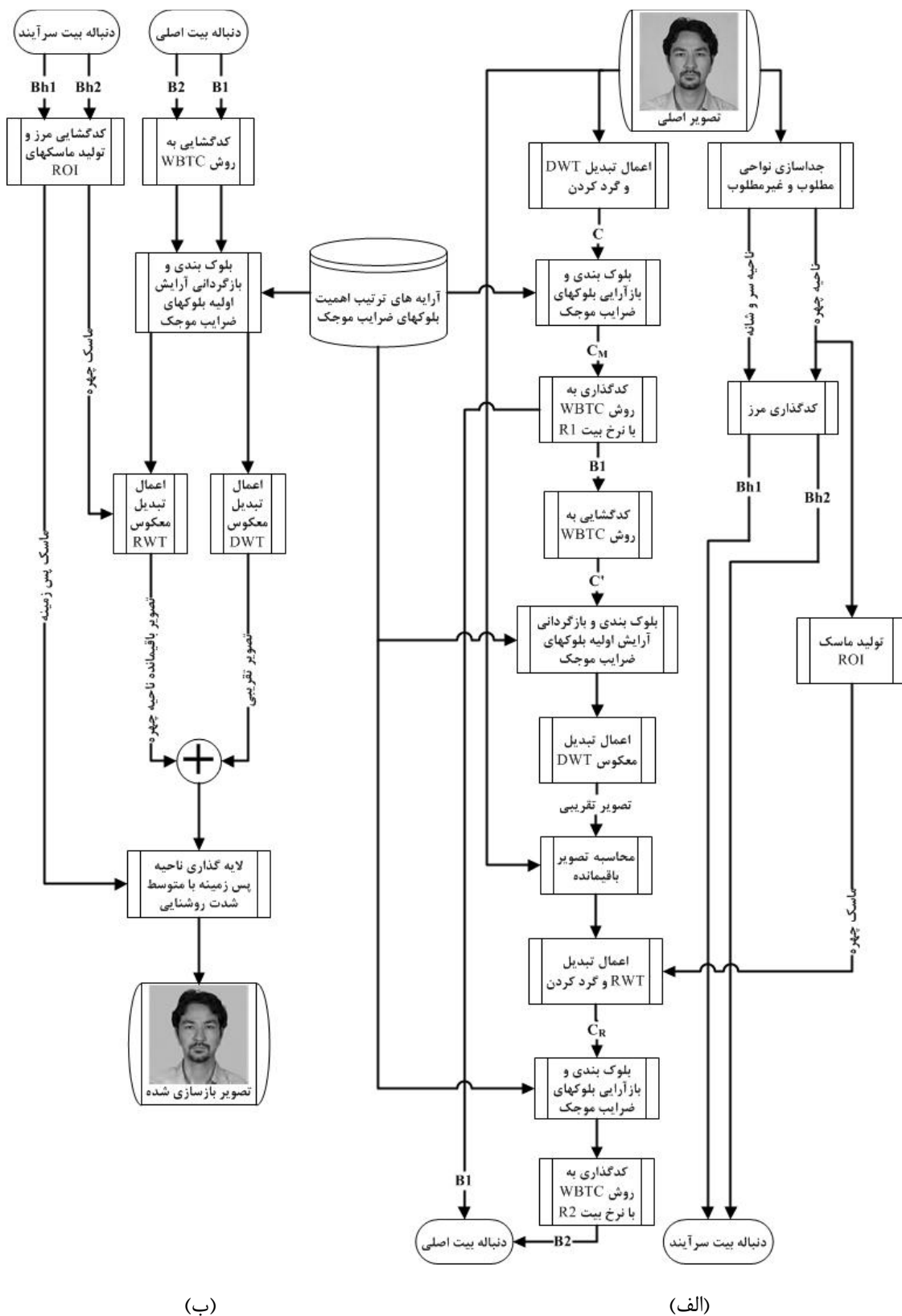
۳-۳ روش پیشنهادی سوم

روش پیشنهادی سوم در حقیقت تلفیقی از روش‌های پیشنهادی اول و دوم است. در این روش با بهره‌گیری توأم از تبدیل RWT و ایده کدگذاری ROI ارائه شده در [۱۱]، توانسته‌ایم علاوه بر دستیابی به کارایی بالاتر نرخ-اعوجاج در نرخ‌بیت‌های پایین نسبت به روش‌های پیشنهادی اول و دوم، سرعت فشرده‌سازی و بازسازی تصاویر را نیز نسبت به روش پیشنهادی دوم افزایش دهیم. در این روش نیز مشابه روش پیشنهادی اول، عمل فشرده‌سازی/بازسازی سرتاسر تصویر تا یک نرخ‌بیت مشخص بطور همزمان صورت می‌گیرد. سپس برای افزایش کیفیت نواحی مطلوب، با اعمال تبدیل RWT روی این نواحی در تصویر باقیمانده و کدگذاری آن‌ها با نرخ‌بیتی دیگر، اهمیت و نرخ‌بیت بیشتری به این نواحی اختصاص داده می‌شود. این روش در مقایسه با روش‌های پیشنهادی اول و دوم و به تبع آن روش [۱۱]، JPEG2000 و SPIHT کارایی بالاتری را در ناحیه چهره و بخصوص در

نرخ‌بیت‌های پایین، در شبیه‌سازی‌های انجام شده نشان داده است. نمودار بلوکی فرآیند فشرده‌سازی و بازسازی روش پیشنهادی سوم در شکل ۳-۵ نشان داده شده است.

۳-۳-۱ فرآیند فشرده‌سازی

در فرآیند فشرده‌سازی ابتدا از تصویر پرسنلی تبدیل موجک گرفته شده و سپس ضرایب موجک حاصل به نزدیکترین مقادیر صحیح گرد می‌شوند. ماتریس ضرایب موجک حاصل از اعمال تبدیل DWT و گرد کردن را C می‌نامیم. در مرحله بعد ماتریس C به بلوک‌هایی مجاور هم با ابعاد $n \times n$ تقسیم شده، و سپس آرایش این بلوک‌ها مطابق آرایه ترتیب اهمیت بدست آمده در فاز شبه-آموزش و الگوی جستجوی ضرایب مهم در الگوریتم WBTC تغییر می‌کند. پس از بازآرایی بلوک‌های ضرایب موجک ماتریس C ، ماتریس حاصل به نام C_M توسط روش WBTC با نرخ‌بیت $R1$ bps کدگذاری می‌شود. دنباله بیت حاصل به نام $B1$ به عنوان دنباله بیت اصلی ذخیره و یا به واحد کدگشا فرستاده می‌شود. به موازات، دنباله بیت $B1$ کدگشایی می‌شود تا ماتریس دیگری به نام C' تولید شود. با بازگردانی آرایش اولیه بلوک‌های ضرایب ماتریس C' و سپس اعمال تبدیل موجک معکوس به آن، یک تصویر تقریبی از چهره بدست خواهد آمد. در مرحله بعد همانطور که در شکل ۳-۵ (الف) نشان داده شده است، با تفاضل تصویر تقریبی از تصویر اصلی یک تصویر باقیمانده محاسبه می‌شود. مرحله بعد اعمال تبدیل RWT روی ناحیه مطلوب در تصویر باقیمانده است. از آنجا که تبدیل RWT برای عملکرد خود نیازمند ماسک ROI است، بایستی قبل از این مرحله ماسک ROI تولید شده باشد. در روش پیشنهادی سوم نیز همانند روش پیشنهادی اول تنها ناحیه چهره به عنوان ناحیه مطلوب در نظر گرفته شده است، چرا که این ناحیه بیشترین اهمیت و تاثیر را در شناسایی صحیح هویت فرد دارد. در اینجا نیز همانند روش‌های پیشنهادی اول و دوم، برای جداسازی نواحی مطلوب و غیرمطلوب



شکل ۳-۵: نمودار بلوکی روش پیشنهادی سوم. (الف) فرآیند فشرده‌سازی. (ب) فرآیند بازسازی.

از تکنیک پیشنهادی ارائه شده در بخش ۳-۴ استفاده شده است. در روش پیشنهادی سوم برای تضعیف اثرات مزاحم بصری ناشی از نشت ضرایب موجک، که به علت اختلاف زیاد شدت روشنایی بین مرز پس‌زمینه و پیش‌زمینه (بخصوص ناحیه سر) در این ناحیه مرزی مشهودتر هستند، در فرآیند بازسازی تصویر مشابه روش پیشنهادی دوم از یک مرحله لایه‌گذاری پس‌زمینه استفاده شده است. این مرحله اضافی در فرآیند بازسازی مستلزم یک مرحله اضافی کدگذاری مرز پیش‌زمینه در فرآیند فشرده‌سازی است، چرا که برای لایه‌گذاری پس‌زمینه به ماسک پس‌زمینه یا پیش‌زمینه نیاز است. پس از جداسازی نواحی مطلوب و غیرمطلوب، ماسک ROI مربوط به ناحیه چهره تولید شده، و همچنین مرز نواحی سر و شانه (پیش‌زمینه) و چهره با استفاده از تکنیک‌های ارائه شده در بخش ۳-۵ کدگذاری شده و به ترتیب دنباله بیت‌های $Bh1$ و $Bh2$ تولید می‌شوند. این دنباله بیت‌ها به عنوان دنباله بیت سرآیند ذخیره شده و یا به واحد کدگشا ارسال می‌شوند. اکنون می‌توان با استفاده از ماسک ROI تولید شده، تبدیل RWT را بر روی ناحیه چهره تصویر باقیمانده محاسبه شده در مرحله قبل اعمال کرد. ماتریس ضرایب موجک حاصل را پس از گرد کردن ضرایب موجک به نزدیکترین مقادیر صحیح، ماتریس باقیمانده C_R می‌نامیم. در مرحله بعد ماتریس باقیمانده C_R بلوک‌بندی شده و سپس بلوک‌های ضرایب آن براساس آرایه ترتیب اهمیت مخصوص ناحیه چهره بازآرایی می‌شوند. در انتها ماتریس ضرایب بازآرایی شده حاصل از مرحله قبل توسط الگوریتم WBTC با نرخ بیت $R2$ کدگذاری شده و دنباله بیت حاصل به نام $B2$ به دنباله بیت اصلی افزوده می‌شود.

۳-۲-۳ فرآیند بازسازی

شکل ۳-۵ (ب) مراحل مختلف فرآیند بازسازی تصویر را در روش پیشنهادی سوم نشان می‌دهد. در فرآیند بازسازی ابتدا دنباله‌های بیتی $B1$ و $B2$ بصورت جداگانه با استفاده از الگوریتم WBTC

کدگشایی شده، و سپس آرایش بلوک‌های هریک از ماتریس‌های ضرایب موجک حاصل بصورت مستقل با استفاده از آرایه ترتیب اهمیت مخصوص خود به حالت اولیه برگردانده می‌شود. در مرحله بعد تبدیل معکوس DWT بر روی ماتریس ضرایب حاصل از کدگشایی دنباله بیت $B1$ اعمال شده و یک تصویر تقریبی بازسازی می‌شود. به موازات این عمل، تبدیل معکوس RWT نیز بر روی ماتریس ضرایب حاصل از کدگشایی دنباله بیت $B2$ اعمال شده و یک تصویر باقیمانده از ناحیه چهره بازسازی می‌شود. البته باید توجه داشت که برای اعمال تبدیل معکوس RWT نیاز به ماسک ROI است و بنابراین ابتدا بایستی دنباله بیت سرآیند $Bh2$ کدگشایی شده و ماسک ناحیه چهره تولید شود. در انتها با جمع تصویر تقریبی و تصویر باقیمانده ناحیه چهره بدست آمده از مرحله قبل، و سپس لایه‌گذاری ناحیه پس‌زمینه با متوسط شدت روشنایی این ناحیه، تصویر بازسازی شده بدست می‌آید. برای لایه‌گذاری ناحیه پس‌زمینه و نیز محاسبه متوسط شدت روشنایی این ناحیه، نیاز به ماسک پس‌زمینه داریم. این ماسک را می‌توان با کدگشایی دنباله بیت $Bh1$ ، بازسازی مرز پیش‌زمینه و سپس پر کردن ناحیه خارج مرز بدست آورد.

۳-۴ جداسازی نواحی مطلوب و غیر مطلوب

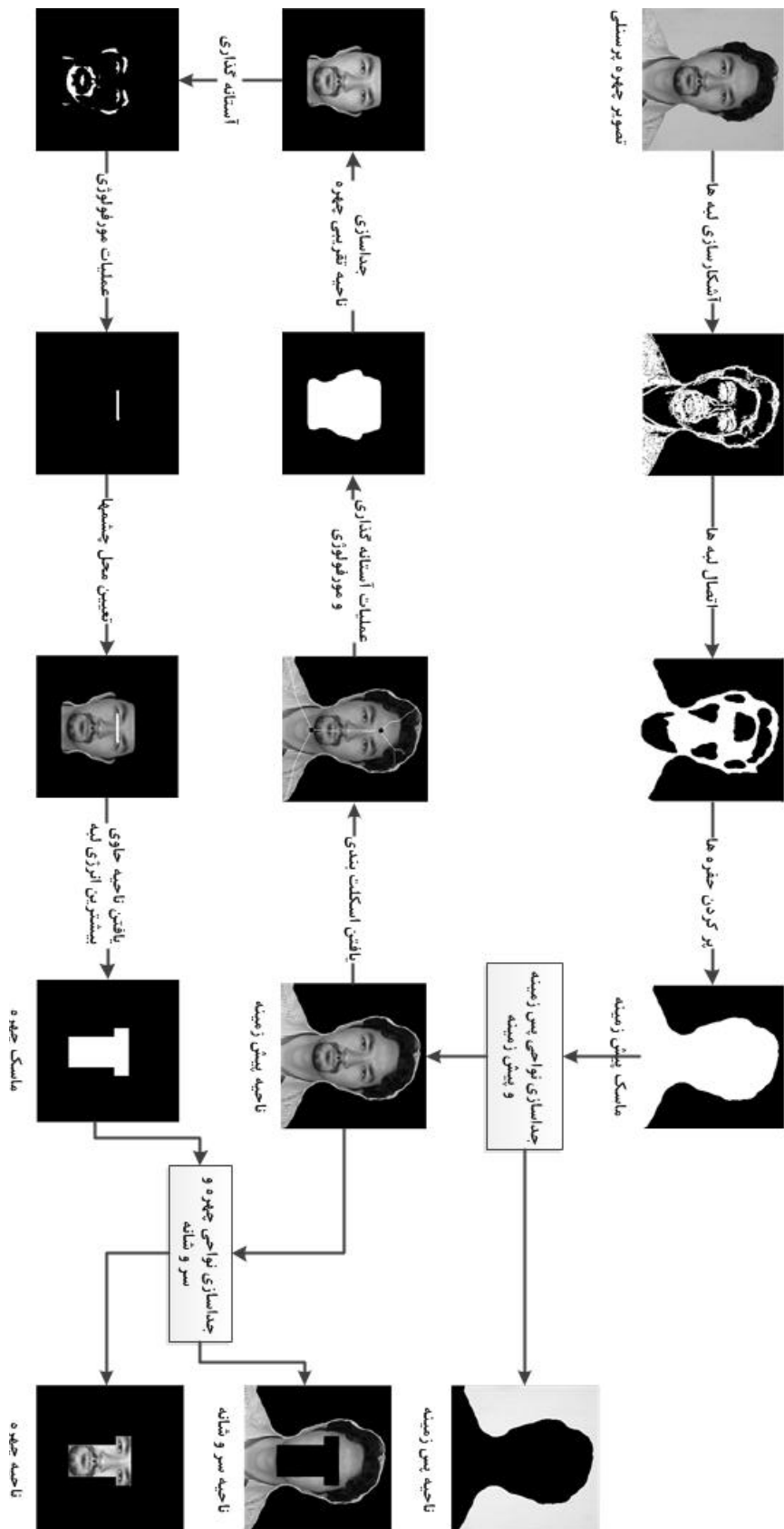
روش‌های متعددی برای تقطیع تصویر^۱ و جداسازی نواحی مطلوب از جمله ناحیه چهره و ویژگی‌های چهره معرفی شده است. اما اغلب این روش‌ها دارای پیچیدگی زیاد و سرعت کم هستند. از آنجا که در فرآیند فشرده‌سازی و به خصوص در انتقال مخابراتی پارامتر سرعت اهمیت فراوانی دارد،

¹ Image Segmentation

در این پایان‌نامه سعی کرده‌ایم تکنیکی سریع برای جداسازی نواحی مطلوب و غیرمطلوب در تصاویر چهره پرسنلی ارائه دهیم. این تکنیک مبتنی بر عملیات لبه‌گیری و شکل‌شناسی است. شکل ۳-۶ مراحل مختلف این تکنیک را به تصویر کشیده است.

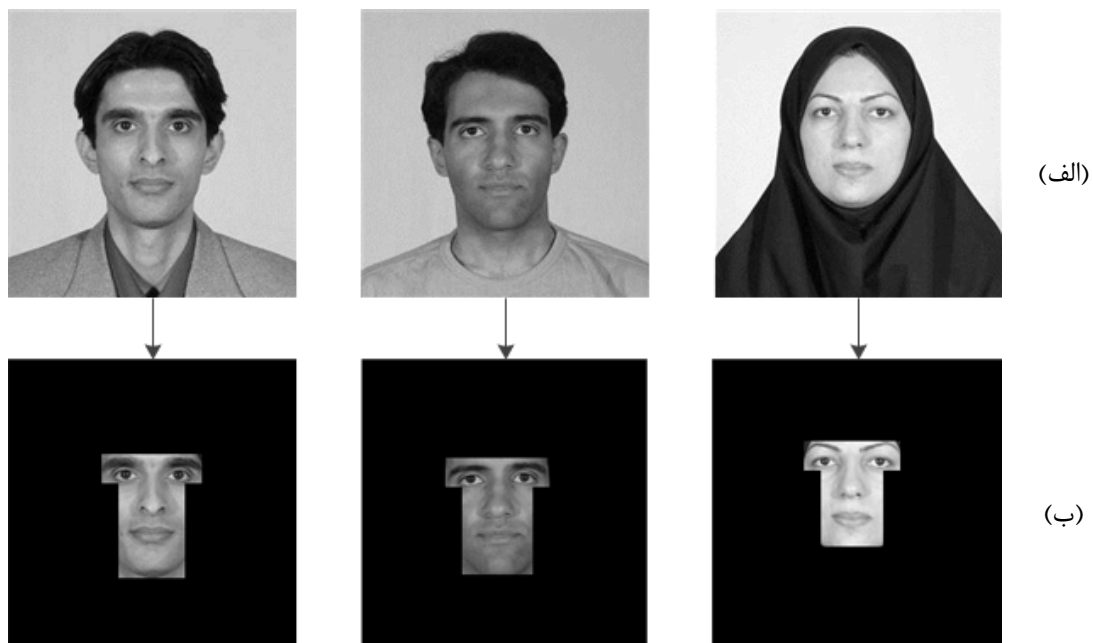
در این روش ابتدا ناحیه پیش‌زمینه از پس‌زمینه جدا می‌شود. این کار به راحتی با لبه‌گیری از تصویر، اتصال لبه‌های موجود در تصویر به منظور رفع شکستگی لبه‌ها در مرز بین پس‌زمینه و پیش‌زمینه و در نهایت پرکردن ناحیه داخلی مرز صورت می‌گیرد. حال ناحیه تقریبی چهره را از ناحیه پیش‌زمینه جدا می‌کنیم. بدین منظور باید اسکلت ناحیه پیش‌زمینه را بدست آوریم. در این اسکلت دو نقطه انشعاب در امتداد و یا حول محور تقارن چهره وجود دارد که دارای ویژگی‌های منحصر بفردی هستند. یکی از این نقاط انشعاب همواره در ناحیه بالای چشم‌ها تا ابتدای پیشانی قرار گرفته است. نقطه دیگر اغلب اوقات در ناحیه زیر دهان تا زیر چانه (گاهی اوقات تا وسط گردن) قرار گرفته است. حال با توجه به اینکه شدت روشنایی پوست نسبت به مو خیلی بیشتر است به راحتی می‌توان با آستانه‌گیری از ناحیه بین این دو نقطه و عملیات مورفولوژی یا همان شکل‌شناسی، ناحیه تقریبی چهره را بدست آورد. مرحله بعدی بدست آوردن ناحیه چشم‌ها است. با توجه به اینکه سطح روشنایی عنبیه نسبت به ناحیه اطرافش بسیار پایین‌تر است لذا انتظار داریم که با آستانه‌گیری از ناحیه تقریبی چهره، عنبیه چشم‌ها بصورت دو شیء دایره مانند نسبتاً کوچک در تصویر دودویی ظاهر شوند. با دانستن اندازه تقریبی عنبیه می‌توان اشیاء کوچکتر و بزرگتر را حذف کرد. همچنین از آنجا که چشم‌ها در یک راستای افقی قرار گرفته‌اند می‌توان با استفاده از عملیات مورفولوژی "بستن"^۱ و المان

¹ Closing



شکل ۳-۶: مراحل مختلف تکنیک پیشنهادی برای جداسازی نواحی مطلوب و غیرمطلوب تصاویر چهره پرسنلی.

ساختاری^۱ "خط با زاویه صفر درجه نسبت به افق" نقاط مربوط به عنبیه دو چشم را به هم وصل کرد. بدین ترتیب ناحیه بین دو عنبیه بصورت یک مستطیل در تصویر دودویی، کاملاً از سایر نقاط و اشیاء متمایز شده و قابل تشخیص است. تشخیص محل بینی و دهان به سادگی یافتن چشم‌ها نیست. لذا در این روش ما به یافتن ناحیه‌ای مستطیلی حول محور تقارن چهره و زیر چشم‌ها که دارای بیشترین انرژی لبه است، به عنوان ناحیه بینی و دهان و حتی چانه اکتفا می‌کنیم. در نهایت با تفاضل ناحیه چهره، شامل نواحی مستطیلی بدست آمده برای بینی و دهان و چشم‌ها و ابروها، از ناحیه پیش‌زمینه ناحیه سر و شانه نیز بدست خواهد آمد. شکل ۳-۷ چند نمونه تصویر چهره پرسنلی و نتایج حاصل از این تکنیک را برای جداسازی ناحیه چهره نشان می‌دهد.



شکل ۳-۷: نتایج حاصل از تکنیک پیشنهادی برای جداسازی ناحیه چهره. (الف) تصاویر چهره پرسنلی. (ب) نواحی چهره جداسازی شده

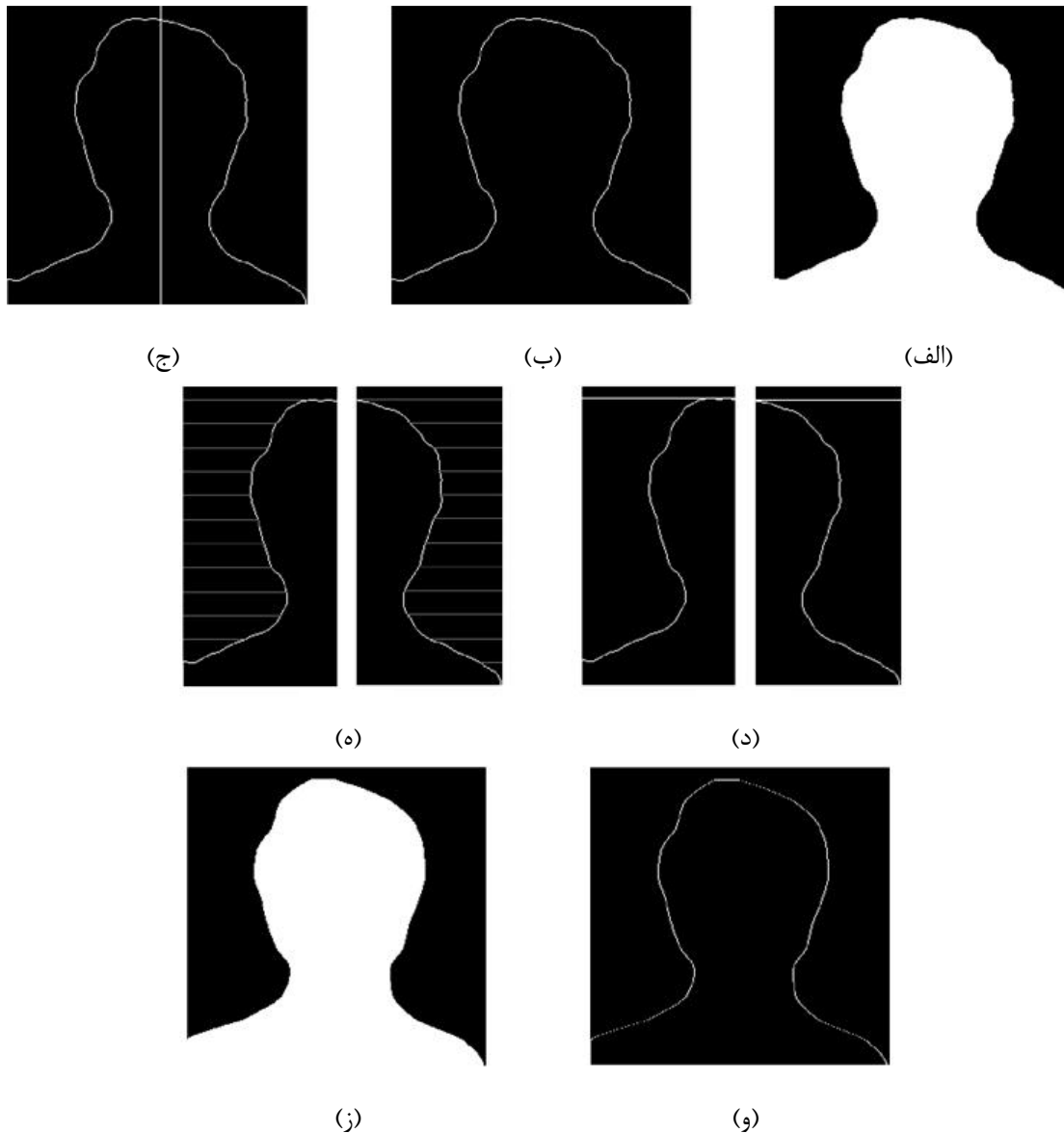
¹ Structure Element

۳-۵ کدگذاری مرز نواحی مطلوب

یکی از رایج‌ترین روش‌های کدگذاری مرز نواحی تصویر، کد زنجیره‌ای فریمن^۱ است. به عنوان مثال در [۸] از کد زنجیره‌ای ۸ جهتی فریمن و کدگذاری حسابی برای کدگذاری مرز ناحیه پیش‌زمینه یک تصویر چهره با ابعاد 256×256 پیکسل استفاده شده است. دنباله بیت حاصل از این کدگذاری نیازمند ۱۶۶ بیت حافظه است، که با توجه به ابعاد تصویر چهره نرخ‌بیتی معادل ۰.۰۲ bpp از تصویر فشرده‌شده را به خود اختصاص خواهد داد. این میزان نرخ‌بیت ممکن است زیاد قابل توجه نباشد، البته تا زمانیکه فشرده‌سازی در نرخ‌بیت‌های نسبتاً بالا صورت بگیرد. در این پایان‌نامه با توجه به اینکه با فشرده‌سازی تصاویر چهره پرسنلی در نرخ‌بیت‌های پایین سر و کار داریم، و از طرفی کدگذاری مرز دو ناحیه پیش‌زمینه و چهره مطرح است، بنابراین روش کد زنجیره‌ای فریمن گزینه مناسبی نبوده و کارایی لازم را نخواهد داشت.

در این پایان‌نامه از یک تکنیک پیشنهادی مبتنی بر چندجمله‌ای درونیاب برای کدگذاری مرز ناحیه پیش‌زمینه استفاده شده است. شکل ۳-۸ برخی از مراحل این تکنیک را به تصویر کشیده است. در این روش پس از استخراج مرز ناحیه پیش‌زمینه توسط عملیات مورفولوژی (شکل ۳-۸(ب))، تصویر مرز حاصل بصورت عمودی به دو بخش مساوی تقسیم می‌شود (شکل ۳-۸(ج)). با فرض اینکه تصویر چهره پرسنلی دارای ابعاد $M \times N$ پیکسل باشد، ابعاد هر یک از این دو بخش $M \times (\frac{N}{2})$ پیکسل خواهد بود. حال مرز موجود در هر یک از این دو بخش را می‌توان همانند یک منحنی یا تابع چند مقدار در نظر گرفت. چنانچه x و y به ترتیب بیانگر شماره سطر و ستون در تصویر باشند، آنگاه

¹ Freeman Chain Code



شکل ۳-۸: تکنیک پیشنهادی برای کدگذاری مرز ناحیه پیش‌زمینه. (الف) ماسک پیش‌زمینه. (ب) مرز پیش‌زمینه. (ج) تقسیم تصویر مرز بصورت عمودی به دو بخش مساوی. (د) سطر شروع مرز در دو بخش چپ و راست. (ه) موقعیت مجموعه نقاط X روی مرز دو بخش چپ و راست با فرض $n = 13$. (و) مرز بازسازی‌شده با فرض $n = 13$. (ز) ماسک بازسازی‌شده با فرض $n = 13$.

هر پیکسل از مرز واقع در سطر x_i و ستون y_i تصویر، همانند یک نقطه از منحنی پیوسته $y = f(x)$ واقع در مختصات (x_i, y_i) از صفحه xy خواهد بود. هدف ما این است که با داشتن مختصات تعداد

مشخصی از نقاط مرز، بتوانیم مختصات سایر نقاط مرز را تخمین بزنیم. بدین منظور به سراغ مبحث درونیابی توابع خواهیم رفت. فرض کنیم f_L و f_R به ترتیب بیانگر منحنی پیوسته مرز بخش چپ و راست باشند. سطر شروع و انتهای مرز را به ترتیب x_{first} و x_{last} می‌نامیم. مقادیر x_{last} و x_{first} در هر دو بخش برابر بوده و همواره $x_{last} = M$ خواهد بود. شکل ۳-۸(د) سطر شروع مرز را در هر دو بخش چپ و راست مرز نشان می‌دهد. می‌خواهیم توابع $f_L(x)$ و $f_R(x)$ را در بازه $[x_{first}, x_{last}]$ درونیابی کنیم. همانطور که گفته شد این توابع، توابعی چند مقدارده بوده و لذا قدم اول و ضروری برای درونیابی، تبدیل آنها به توابعی تک مقدارده است. بدین منظور در هر سطر از هر دو بخش، تنها بیرونی‌ترین نقطه مرزی را نگاه داشته و مابقی را حذف می‌کنیم. حال از بازه $[x_{first}, x_{last}]$ ، n نقطه با مقادیر صحیح شامل نقاط ابتدا و انتهای بازه انتخاب کرده بطوریکه داشته باشیم:

$$x_{first} = x_1 < x_2 < \dots < x_n = x_{last} \quad (۱-۳)$$

مجموعه این n نقطه را $X = \{x_i | i = 1, \dots, n\}$ می‌نامیم. با فرض $n = 13$ و فواصل مساوی بین نقاط، در شکل ۳-۸(ه) موقعیت این n نقطه روی مرز در هر دو بخش چپ و راست توسط خطوطی افقی مشخص شده‌اند. حال فرض کنیم مجموعه $Y_L = \{f_L(x_i) | i = 1, \dots, n\}$ و

$Y_R = \{f_R(x_i) | i = 1, \dots, n\}$ به ترتیب شماره ستون‌های بیرونی‌ترین نقاط مرزی بخش چپ و راست مرز باشند. در روش درونیابی مکعبی تکه‌ای^۱ ثابت شده است، چنانچه f_L و f_R توابعی تکه‌ای

^۱ Piecewise Cubic Interpolation

هموار^۱ و یکنوا^۲ باشند، آنگاه می‌توان توابع درونیاب مکعبی تکه‌ای مانند $p_L(x)$ و $p_R(x)$ را در بازه $[x_{first}, x_{last}]$ بدست آورد، بطوریکه:

$$p_L(x_i) = f_L(x_i) \quad , \quad p_R(x_i) = f_R(x_i) \quad (2-3)$$

$$i = 1, \dots, n$$

و $p_L(x)$ و $p_R(x)$ توابعی تکه‌ای یکنوا باشند. در حقیقت $p_L(x)$ و $p_R(x)$ در هر زیربازه $[x_i, x_{i+1}]$ یک چندجمله‌ای درونیاب مکعبی هرمیتی^۳ هستند. برای اطلاعات بیشتر راجع به روش درونیابی مکعبی تکه‌ای به [۳۹] رجوع شود.

با توجه به مطالب گفته شده، واحد کدگشا تنها با دریافت سه مجموعه X ، Y_L و Y_R قادر خواهد بود مرز بخش چپ و راست را با دقتی مشخص درونیابی و بازسازی کند. دقت مرز بازسازی شده وابسته به تعداد نقاط نمونه‌برداری شده از مرز یا همان n است. البته انتخاب نقاط نیز به نوبه خود بر روی دقت درونیابی و بازسازی تاثیر دارد، بطوریکه اگر نقاط به نحوی انتخاب شوند که در هر زیربازه f_L و f_R توابعی یکنوا باشند دقت بالاتری حاصل می‌شود. در روش پیشنهادی به منظور کاهش نرخ بیت کدگذاری مرز پیش‌زمینه تا حد ممکن، سه راه حل در نظر گرفته شده است.

(۱) از مجموعه X تنها نقاط x_{first} و x_{last} به واحد کدگشا فرستاده می‌شوند و سایر نقاط با فواصل مساوی بین این دو نقطه و با مقادیر صحیح ایجاد می‌شوند.

¹ Smooth

² Monotone

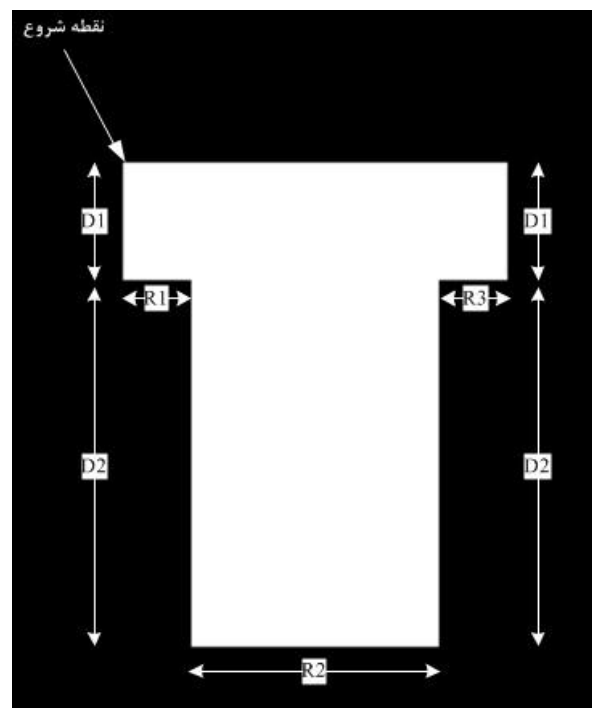
³ Hermite Cubic Interpolating Polynomial

(۲) از آنجا که x_{last} در نیمه پایینی تصویر و مجموعه نقاط Y_R در نیمه راست تصویر واقع شده‌اند، ابتدا به ترتیب مقادیر $\frac{M}{2}$ و $\frac{N}{2}$ از آنها کسر شده و سپس به واحد کدگشا ارسال می‌شوند.

(۳) مقدار n تا حد ممکن کوچک در نظر گرفته شده است، بطوریکه حداقل دقت مورد انتظار از بازسازی مرز تحقق یابد.

شکل ۸-۳ (و) و شکل ۸-۳ (ز) به ترتیب مرز و ماسک بازسازی شده را برای ناحیه پیش‌زمینه با فرض $n = 13$ نشان می‌دهند. لازم به ذکر است که دنباله بیت حاصل از کدگذاری مرز پیش‌زمینه با فرض $n = 13$ تنها ۱۹۶ بیت از حافظه را اشغال می‌کند. با توجه به ابعاد 256×256 پیکسلی تصاویر پرسنلی مورد استفاده در این پایان‌نامه، این تعداد بیت معادل یک نرخ بیت حدود 0.003 bpp خواهد بود که حتی در نرخ‌بیت‌های خیلی پایین نیز قابل چشم‌پوشی است.

کدگذاری مرز ناحیه چهره نسبت به پیش‌زمینه به روش بسیار ساده‌تری انجام می‌گیرد. دلیل این امر شکل مستطیل مانند این مرز می‌باشد. در حقیقت همانطور که در شکل ۹-۳ نشان داده شده است، تنها با ارسال مختصات یک رأس مرز مستطیلی ناحیه چهره به عنوان نقطه شروع، و نیز جابجایی‌های انجام شده نسبت به این نقطه در راستای افقی و عمودی، واحد کدگشا قادر خواهد بود این مرز را بطور کامل بازسازی کند. در این روش تنها حدود ۵۰ بیت برای کدگذاری مرز ناحیه چهره مصرف می‌شود.



شکل ۳-۹: نمایش نقطه شروع و میزان جابجایی‌ها نسبت به این نقطه برای کدگذاری مرز ناحیه چهره.

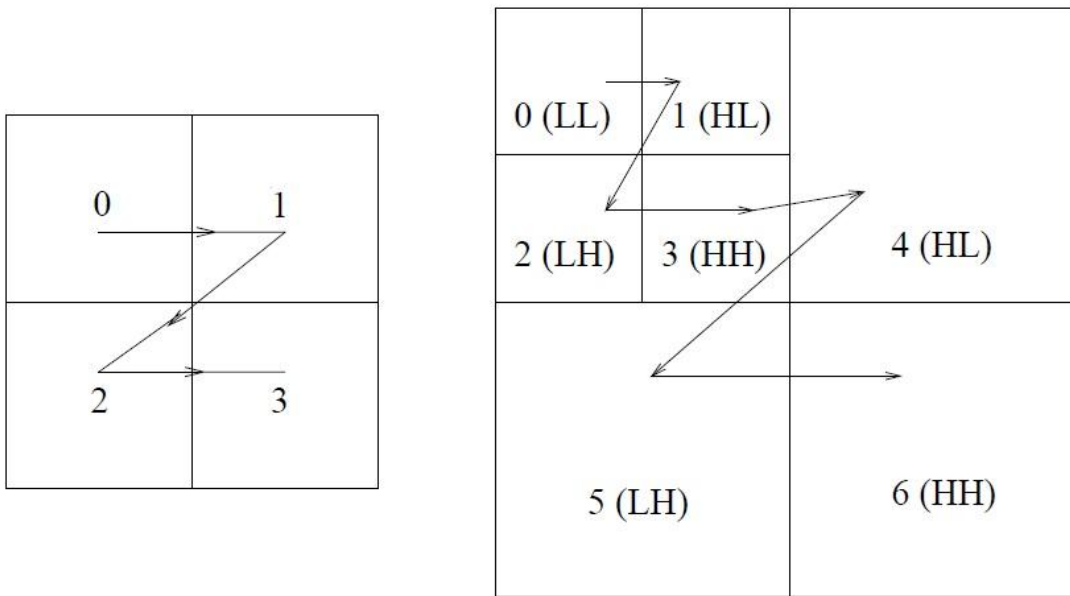
۳-۶ بهبود کارایی کدگذاری WBTC

می‌توان گفت کدگذاری WBTC به نوعی یک الگوریتم جستجوی منظم برای یافتن ضرایب موجک مهم و کدگذاری آن‌ها است. طبق این تعریف چنانچه بتوان آرایش و چیدمان ضرایب موجک را به نحوی تغییر داد که ضرایب مهم با جستجوی کمتری یافت شوند، کارایی کدگذاری WBTC افزایش خواهد یافت، چرا که در این حالت زمان و نرخ بیت کمتری صرف یافتن ضرایب مهم یا همان ضرایب بزرگتر خواهد شد. الگوریتم جستجو در کدگذاری WBTC و نیز اغلب روش‌های کدگذاری مبتنی بر موجک به گونه‌ای تعریف شده است که ابتدا زیرباندهای فرکانسی پایین‌تر مورد جستجو قرار می‌گیرند و سپس زیرباندهای فرکانسی بالاتر. دلیل این امر بهره‌گیری از خاصیت تبدیل موجک در متمرکز کردن و فشرده‌سازی انرژی تصویر یا بطور معادل تمرکز ضرایب بزرگتر در زیرباندهای

فرکانسی پایین‌تر است. بنابراین باید ضرایب موجک را براساس اندازه به ترتیب از بزرگترین تا کوچکترین ضریب، در زیرباندهای فرکانسی پایین تا زیرباندهای فرکانسی بالا قرار داد. ترتیب قرارگیری ضرایب موجک در زیرباندهای فرکانسی مطابق با الگوی جستجوی مورد استفاده در WBTC است.

در این پایان‌نامه سعی شده است که با تطبیق الگوی جستجوی ضرایب مهم در الگوریتم WBTC با یک ترتیب پیمایش Z ، آرایش و ترتیبی مشخص برای قرارگیری ضرایب موجک بر حسب اهمیت در زیرباندهای فرکانسی تعریف شود. در ترتیب پیمایش Z ، جستجوی ضرایب مهم درون هر زیرباند و نیز بین زیرباندها از الگویی به شکل حرف Z تبعیت می‌کند، که این الگوی Z بصورت تکرار شونده و در مقیاس‌های مختلف ظاهر می‌شود. شکل ۳-۱۰ ترتیب پیمایش بین زیرباندها و نیز ترتیب پیمایش درون هر بلوک را در الگوریتم WBTC تطبیق یافته با الگوی پیمایش Z نشان می‌دهد. براساس الگوری پیمایش Z می‌توان ترتیب قرارگیری ضرایب موجک را بر حسب اهمیت، در یک ماتریس ضرایب 16×16 با دو سطح تجزیه مطابق شکل ۳-۱۱ در نظر گرفت. در شکل ۳-۱۱ فرض شده است که واحد پایه درخت‌های بلوکی مورد استفاده در الگوریتم WBTC، بلوک‌هایی با ابعاد 2×2 هستند و با این فرض درخت‌های بلوکی بصورت متمایز نشان داده شده‌اند. این نحوه چیدمان ضرایب قابل تعمیم به ماتریس‌های ضرایب بزرگتر با تعداد سطوح تجزیه بیشتر است.

ارسال ترتیب کلیه ضرایب برای هر تصویر به عنوان اطلاعات جانبی باعث کاهش شدید کارایی فشرده‌سازی خواهد شد. برای رفع این مشکل می‌توان ماتریس ضرایب را به بلوک‌های $n \times n$ تقسیم کرده و سپس این بلوک‌ها را براساس متوسط انرژی ضرایب موجود در آن‌ها بصورت نزولی مرتب کرد. حال این بلوک‌های ضرایب مطابق شکل ۳-۱۱ به ترتیب اهمیت در موقعیت‌های جدید قرار داده می‌شوند. هرچه اندازه بلوک‌ها یعنی n افزایش یابد، تعداد بلوک‌ها و در نتیجه اطلاعات جانبی کاهش



(ب)

(الف)

شکل ۳-۱۰: الگوی پیمایش Z. (الف) ترتیب پیمایش بین زیربافتها. (ب) ترتیب پیمایش درون بلوک [۲۸].

۱	۲	۵	۶	۱۷	۱۸	۲۱	۲۲	۶۵	۶۶	۶۹	۷۰	۸۱	۸۲	۸۵	۸۶
۳	۴	۷	۸	۱۹	۲۰	۲۳	۲۴	۶۷	۶۸	۷۱	۷۲	۸۳	۸۴	۸۷	۸۸
۹	۱۰	۱۳	۱۴	۲۵	۲۶	۲۹	۳۰	۷۳	۷۴	۷۷	۷۸	۸۹	۹۰	۹۳	۹۴
۱۱	۱۲	۱۵	۱۶	۲۷	۲۸	۳۱	۳۲	۷۵	۷۶	۷۹	۸۰	۹۱	۹۲	۹۵	۹۶
۳۳	۳۴	۳۷	۳۸	۴۹	۵۰	۵۳	۵۴	۹۷	۹۸	۱۰۱	۱۰۲	۱۱۳	۱۱۴	۱۱۷	۱۱۸
۳۵	۳۶	۳۹	۴۰	۵۱	۵۲	۵۵	۵۶	۹۹	۱۰۰	۱۰۳	۱۰۴	۱۱۵	۱۱۶	۱۱۹	۱۲۰
۴۱	۴۲	۴۵	۴۶	۵۷	۵۸	۶۱	۶۲	۱۰۵	۱۰۶	۱۰۹	۱۱۰	۱۲۱	۱۲۲	۱۲۵	۱۲۶
۴۳	۴۴	۴۷	۴۸	۵۹	۶۰	۶۳	۶۴	۱۰۷	۱۰۸	۱۱۱	۱۱۲	۱۲۳	۱۲۴	۱۲۷	۱۲۸
۱۲۹	۱۳۰	۱۳۳	۱۳۴	۱۴۵	۱۴۶	۱۴۹	۱۵۰	۱۹۳	۱۹۴	۱۹۷	۱۹۸	۲۰۹	۲۱۰	۲۱۳	۲۱۴
۱۳۱	۱۳۲	۱۳۵	۱۳۶	۱۴۷	۱۴۸	۱۵۱	۱۵۲	۱۹۵	۱۹۶	۱۹۹	۲۰۰	۲۱۱	۲۱۲	۲۱۵	۲۱۶
۱۳۷	۱۳۸	۱۴۱	۱۴۲	۱۵۳	۱۵۴	۱۵۷	۱۵۸	۲۰۱	۲۰۲	۲۰۵	۲۰۶	۲۱۷	۲۱۸	۲۲۱	۲۲۲
۱۳۹	۱۴۰	۱۴۳	۱۴۴	۱۵۵	۱۵۶	۱۵۹	۱۶۰	۲۰۳	۲۰۴	۲۰۷	۲۰۸	۲۱۹	۲۲۰	۲۲۳	۲۲۴
۱۶۱	۱۶۲	۱۶۵	۱۶۶	۱۷۷	۱۷۸	۱۸۱	۱۸۲	۲۲۵	۲۲۶	۲۲۹	۲۳۰	۲۴۱	۲۴۲	۲۴۵	۲۴۶
۱۶۳	۱۶۴	۱۶۷	۱۶۸	۱۷۹	۱۸۰	۱۸۳	۱۸۴	۲۲۷	۲۲۸	۲۳۱	۲۳۲	۲۴۳	۲۴۴	۲۴۷	۲۴۸
۱۶۹	۱۷۰	۱۷۳	۱۷۴	۱۸۵	۱۸۶	۱۸۹	۱۹۰	۲۳۳	۲۳۴	۲۳۷	۲۳۸	۲۴۹	۲۵۰	۲۵۳	۲۵۴
۱۷۱	۱۷۲	۱۷۵	۱۷۶	۱۸۷	۱۸۸	۱۹۱	۱۹۲	۲۳۵	۲۳۶	۲۳۹	۲۴۰	۲۵۱	۲۵۲	۲۵۵	۲۵۶

شکل ۳-۱۱: ترتیب قرارگیری ضرایب موجک مطابق الگوی Z در یک ماتریس ضرایب ۱۶×۱۶ با دو سطح تجزیه.

پیدا می‌کند، اما در عین حال این امر منجر به کاهش تاثیر بازآرایی ضرایب خواهد شد و بایستی بین این دو مسئله مصالحه‌ای صورت گیرد.

۳-۶-۱ فاز شبه-آموزش

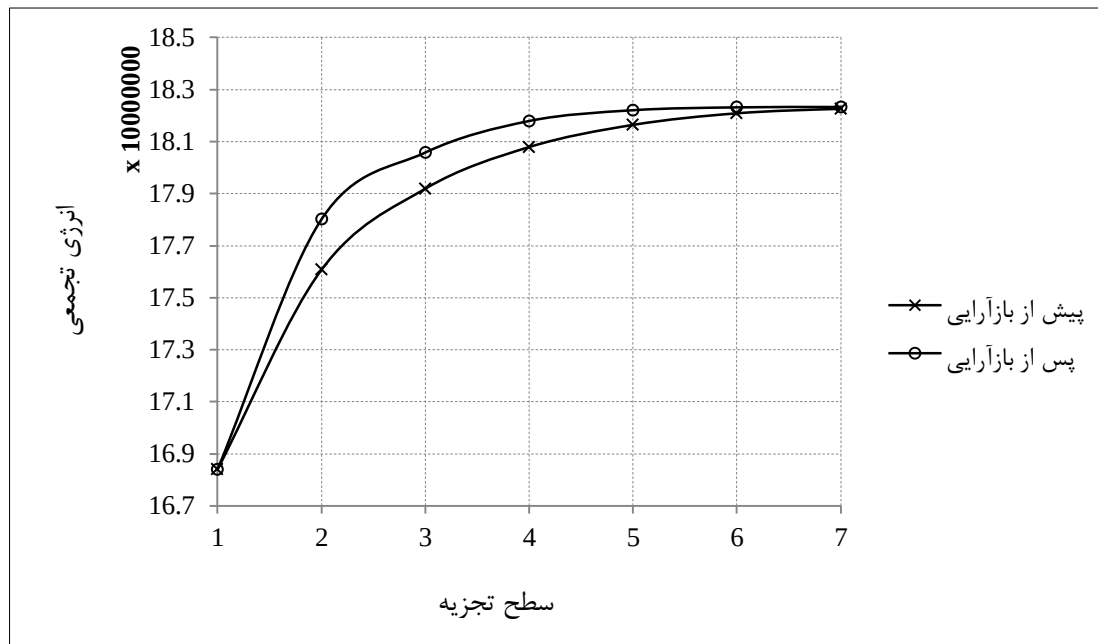
راه حل دوم برای کاهش اطلاعات جانبی و یا در حقیقت حذف اطلاعات جانبی، استفاده از یک آرایه ترتیب اهمیت یکسان، برای ضرایب موجک کلیه تصاویر چهره پرسنلی موجود در پایگاه داده است. این آرایه ترتیب اهمیت را می‌توان در مرحله‌ای به نام فاز شبه-آموزش بدست آورده و سپس در هر دو واحد کدگذار و کدگشا نگهداری کرد. دلیل نامگذاری این مرحله به فاز شبه-آموزش، شباهت تقریبی آن به فاز آموزش در روش‌های فشرده‌سازی مبتنی بر فرهنگ داده است. بدین ترتیب هزینه اطلاعات جانبی جای خود را به هزینه حافظه جانبی برای ذخیره‌سازی آرایه ترتیب اهمیت در واحدهای کدگذار و کدگشا خواهد داد، که البته این حجم حافظه جانبی در قیاس با حجم حافظه مصرف شده برای پایگاه‌های بزرگ داده قابل چشم‌پوشی و صرف‌نظر است.

یک راه ساده برای بدست آوردن آرایه ترتیب اهمیت ضرایب در فاز شبه-آموزش، میانگین‌گیری از ماتریس ضرایب موجک روی کل تصاویر پایگاه داده و سپس مرتب کردن ضرایب ماتریس میانگین حاصل بر حسب اندازه بصورت نزولی است. از آنجا که طرح کلی تصاویر چهره پرسنلی تقریباً یکسان است و در حوزه موجک این طرح کلی حفظ خواهد شد، می‌توان ادعا کرد که این روش ساده بطور متوسط باعث تمرکز بیشتر انرژی ضرایب تصاویر پایگاه داده در زیرباندهای فرکانسی پایین‌تر، و در نتیجه افزایش کارایی کدگذاری WBTC خواهد شد.

با توجه به اینکه پوشش زنان و مردان در پایگاه داده مورد استفاده در این پایان‌نامه متفاوت است، این مسئله ممکن است باعث ایجاد اختلاف‌هایی در طرح کلی تصاویر مربوط به هر جنسیت شده و در

نتیجه در برخی موارد تاثیر بازآرایی ضرایب را کاهش دهد. در روش‌های فشرده‌سازی پیشنهادی برای اجتناب از این مسئله، در فاز شبه-آموزش آرایه ترتیب اهمیت برای هر جنسیت بصورت جداگانه بدست می‌آید. علاوه بر این با توجه به اینکه در روش پیشنهادی دوم، دو ناحیه چهره و سر و شانه به عنوان نواحی مطلوب در نظر گرفته شده‌اند، در فاز شبه-آموزش آرایه ترتیب اهمیت متناظر با هر یک از این نواحی نیز بصورت جداگانه محاسبه می‌شود. در روش پیشنهادی سوم نیز بطور مشابه آرایه‌های ترتیب اهمیت مجزایی برای سرتاسر تصویر و ناحیه چهره بدست خواهد آمد. بنابراین در روش‌های پیشنهادی دوم و سوم در کل باید چهار آرایه ترتیب اهمیت برای هر روش در فاز شبه-آموزش بدست آوریم. با توجه به اینکه ابعاد تصاویر پرسنلی مورد استفاده در این پایان‌نامه 256×256 پیکسل است، فضای حافظه‌ای حدود ۱۲۸ کیلوبایت برای هر آرایه ترتیب اهمیت و لذا ۵۱۲ کیلوبایت برای چهار آرایه مورد نیاز است. به منظور کاهش حافظه جانبی مورد نیاز برای ذخیره‌سازی این تعداد آرایه، در فاز شبه-آموزش با تقسیم ماتریس میانگین ضرایب به بلوک‌هایی با ابعاد $n \times n$ ، به جای چهار آرایه ترتیب اهمیت ضرایب، چهار آرایه ترتیب اهمیت بلوک‌های ضرایب بدست آمده و در نتیجه فضای حافظه مورد نیاز برای هر آرایه کاهش می‌یابد. به عنوان مثال با فرض اندازه بلوک 2×2 فضای حافظه مورد نیاز برای هر آرایه به ۲۸ کیلوبایت کاهش پیدا می‌کند. شکل ۳-۱۲ منحنی انرژی تجمعی میانگین ضرایب موجک ناحیه چهره جنسیت مذکر را پیش و پس از بازآرایی بلوک‌های ضرایب بصورت تابعی از سطح تجزیه موجک نشان می‌دهد. این منحنی‌ها با فرض ۷ سطح تجزیه موجک، اندازه بلوک 2×2 و با میانگین‌گیری روی ۳۴ تصویر مرد پایگاه داده بدست آمده‌اند. در شکل ۳-۱۲ سطح تجزیه ۱ در محور افقی متناظر با پایین‌ترین باند فرکانسی تصاویر بوده، و انرژی تجمعی متناظر با سطح تجزیه L برابر مجموع انرژی ضرایب میانگین موجود در سطوح تجزیه ۱ تا L است. دو نکته مهم از شکل ۳-۱۲ قابل درک است. نکته اول قابلیت فوق‌العاده تبدیل موجک در تراکم‌سازی انرژی

در زیرباندهای فرکانسی پایین می‌باشد. نکته دوم کارایی روش ساده استفاده شده برای بازآرایی بلوک‌های ضرایب، در تراکم بیشتر انرژی در زیرباندهای فرکانسی پایین نسبت به حالت پیش از بازآرایی است.



شکل ۳-۱۲: منحنی‌های سطح تجزیه-انرژی تجمعی میانگین ضرایب موجک ناحیه چهره پیش و پس از بازآرایی.



نتایج شبیه سازی

۴-۱ معرفی پایگاه داده

در این پایان نامه از پایگاه تصاویر چهره دانشگاه صنعتی امیرکبیر استفاده شده است. این پایگاه تصاویر شامل ۴۸۰ تصویر از ۴۰ شخص مختلف شامل ۳۴ مرد و ۶ زن ایرانی می باشد. از هر شخص در ۱۲ حالت متفاوت از قبیل خنثی، شاد، عصبی، متعجب، با عینک، با چشمان بسته، چرخش سر در جهت های مختلف و نیز زوایای نورپردازی مختلف تصویر گرفته شده است. این تصاویر رنگی بوده و دارای ابعاد 2048×1536 پیکسل می باشند. از آنجا که تصاویر مورد نظر ما، تصاویر چهره پرسنلی هستند، از این پایگاه تصاویر به ازای هر فرد تنها تصویر خنثی انتخاب شد. سپس از هر تصویر انتخابی یک ناحیه مربعی شامل سر، بخشی از شانه ها و پس زمینه برش داده شده و به ابعاد 256×256 پیکسل تغییر یافتند. در نهایت تمام ۴۰ تصویر حاصل به تصاویر سطح خاکستری با نرخ بیت ۸ bpp تبدیل شدند. این پایگاه جدید تصاویر چهره به عنوان تصاویر چهره پرسنلی جهت انجام آزمایش ها مورد استفاده قرار گرفته است. در شکل ۴-۱ چند نمونه از این تصاویر چهره پرسنلی نشان داده شده اند.



شکل ۴-۱: چند نمونه تصویر چهره پرسنلی از پایگاه داده.

۴-۲ معیارهای ارزیابی

دو معیار مهم برای ارزیابی کارایی روش‌های فشرده‌سازی تصاویر عبارتند از:

(۱) میزان فشرده‌سازی

(۲) میزان اعوجاج یا میزان کیفیت

میزان فشرده‌سازی با کمیتی به نام نرخ فشرده‌سازی سنجیده می‌شود. البته کمیت دیگری به نام نرخ‌بیت نیز وجود دارد که خیلی زیاد مورد استفاده قرار می‌گیرد. منظور از نرخ‌بیت تعداد بیت‌هایی است که برای ذخیره‌سازی هر پیکسل مصرف می‌شود. چنانچه فرض کنیم f تصویر اصلی^۱ با ابعاد $M \times N$ پیکسل و نرخ‌بیت b ، و f^c تصویر فشرده‌شده^۲ با ابعاد $M \times N$ پیکسل و نرخ‌بیت $B.R$ باشد، نرخ فشرده‌سازی بصورت زیر تعریف خواهد شد:

$$C.R = \frac{S^f}{S^{f^c}} \quad (۱-۴)$$

که در آن S^f تعداد کل بیت‌های لازم برای ذخیره‌سازی تصویر اصلی f و S^{f^c} تعداد کل بیت‌های لازم برای ذخیره‌سازی تصویر فشرده‌شده f^c است. با معلوم بودن ابعاد تصویر f و نرخ‌بیت آن داریم:

$$S^f = N M b \quad (۲-۴)$$

بنابراین نرخ‌بیت تصویر فشرده‌شده را می‌توان از رابطه زیر بدست آورد:

$$B.R = \frac{b}{C.R} \quad (۳-۴)$$

^۱ Original Image

^۲ Compressed Image

هرچه مقدار $C.R$ بیشتر و یا بطور معادل هرچه مقدار $B.R$ کمتر باشد، میزان فشرده سازی تصویر بالاتر خواهد بود.

میزان اعوجاج یا میزان کیفیت معیار مهم دیگری در ارزیابی کارایی فشرده سازی است. این معیار را اغلب با کمیتی به نام مقدار بیشینه سیگنال به نویز^۱ مورد سنجش قرار می دهند. در اینجا نیز چنانچه فرض کنیم f تصویر اصلی با ابعاد $M \times N$ پیکسل و نرخ بیت b ، و f^{dec} تصویر بازسازی شده^۲ با ابعاد $M \times N$ پیکسل باشد، مقدار بیشینه سیگنال به نویز بصورت زیر تعریف می شود:

$$PSNR(f, f^{dec}) = 10 \log_{10} \left(\frac{MN(2^b - 1)^2}{\|f^{dec} - f\|^2} \right) [dB] \quad (4-4)$$

در این رابطه عبارت $MN(2^b - 1)^2$ نشان دهنده مقدار بیشینه انرژی تصویر اصلی و عبارت $\|f^{dec} - f\|^2$ نشان دهنده انرژی نویز حاصل از فرآیند فشرده سازی است. هرچه مقدار $PSNR$ بیشتر باشد میزان کیفیت تصویر بالاتر یا بطور معادل میزان اعوجاج تصویر پایین تر خواهد بود.

در یک فرآیند فشرده سازی هدف این است که میزان فشرده سازی تا حد ممکن بالا و میزان اعوجاج تا حد ممکن پایین باشد، که البته این دو معیار همواره عکس یکدیگر رفتار کرده و لذا بایستی مصالحه ای میان آنها صورت گیرد.

¹ Pick Signal to Noise Ratio

² Reconstructed Image

۳-۴ نتایج شبیه سازی

برای مقایسه کارایی روش های پیشنهادی با سایر روش ها، نتایج بدست آمده با نتایج حاصل از روش ارائه شده در [۱۱]، نسخه اصلی کدگذاری SPIHT بدون کدگذاری حسابی [۱۲] و استاندارد JPEG2000 [۳۸] مقایسه شده اند. روش های پیشنهادی، روش کدگذاری SPIHT و نیز روش [۱۱]، با استفاده از نرم افزار Matlab Version 7.1 و اجرا روی یک رایانه قابل حمل^۱ با پردازنده Corei5 2.53 GHz و حافظه 4-GByte RAM شبیه سازی شده اند. همچنین برای کسب نتایج JPEG2000 از نرم افزار JPEG2000 Compressor استفاده شده است.

در هر سه روش پیشنهادی برای دستیابی به بیشترین کارایی تکنیک بازآرایی ضرایب و در عین حال اختصاص حدی مناسب از حافظه جانبی به آرایه های ترتیب اهمیت، در مرحله بازآرایی بلوک های ضرایب و نیز فاز شبه-آموزش اندازه بلوک ها 2×2 در نظر گرفته شده اند. از طرفی در روش پیشنهادی اول برای کاهش بیشتر حافظه جانبی، تنها بلوک های چهار سطح انتهایی تجزیه موجک بازآرایی می شوند. اگرچه این عمل باعث کاهش کارایی تکنیک بازآرایی ضرایب می شود، اما همچنان کارایی بالاتر روش پیشنهادی اول را مخصوصاً در نرخ بیت های پایین و در ناحیه چهره حفظ خواهد کرد. علت این امر باز هم به قابلیت بسیار خوب تبدیل موجک در متراکم سازی انرژی برمی گردد؛ بطوریکه بخش قابل توجهی از انرژی تصویر در سطوح انتهایی تجزیه موجک متمرکز شده و در نرخ بیت های خیلی پایین و پایین عموماً ابتدا ضرایب موجک موجود در این سطوح تجزیه کدگذاری می شوند. در روش های پیشنهادی دوم و سوم برای ذخیره سازی چهار آرایه ترتیب اهمیت به ۱۱۲ کیلوبایت حافظه

¹ Laptop

جانبی نیاز است، در حالیکه این میزان حافظه جانبی در روش پیشنهادی اول به ۵۱۲ بایت برای ذخیره‌سازی دو آرایه ترتیب اهمیت کاهش یافته است. البته در پایگاه‌های داده بزرگ که دارای صدها و یا هزاران تصویر هستند، حتی ۱۱۲ کیلوبایت حافظه جانبی نیز در مقایسه با حجم حافظه مورد نیاز برای ذخیره‌سازی این تصاویر قابل چشم‌پوشی است. در شبیه‌سازی‌های انجام شده، زمان اجرای^۱ فاز شبه-آموزش روی ۴۰ تصویر پایگاه داده در روش‌های پیشنهادی اول، دوم و سوم به ترتیب ۰.۶ ثانیه، ۳۲ ثانیه و ۲۰.۷ ثانیه بوده است. البته این زمان‌های اجرا با برنامه‌نویسی غیر بهینه و نیز اجرا در یک رایانه قابل حمل بدست آمده‌اند، و لذا مسلماً با برنامه‌نویسی بهینه و اجرا در یک رایانه شخصی^۲ کاهش قابل توجهی خواهند داشت.

با افزایش تعداد سطوح تجزیه موجک، میزان تراکم انرژی در سطوح انتهایی تجزیه نیز افزایش می‌یابد. این مسئله از یک طرف باعث افزایش کارایی کدگذاری و فشرده‌سازی شده و از طرف دیگر باعث افزایش پیچیدگی محاسباتی خواهد شد. در شبیه‌سازی‌های انجام شده، این افزایش پیچیدگی محاسباتی در مورد تبدیل RWT نسبت به تبدیل DWT بیشتر به چشم می‌آید. البته این به معنای نقض ادعای مطرح شده در [۱۴] و بخش ۲-۱-۴ در مورد پایین‌تر بودن پیچیدگی محاسباتی تبدیل RWT نسبت به DWT نیست. در حقیقت علت بالاتر بودن پیچیدگی محاسباتی تبدیل RWT نسبت به DWT نقص در نحوه پیاده‌سازی و برنامه‌نویسی این تبدیل است؛ چرا که تبدیل DWT را می‌توان بصورت ماتریسی پیاده‌سازی کرد، در حالیکه در تبدیل RWT این قابلیت وجود ندارد.

¹ Runtime

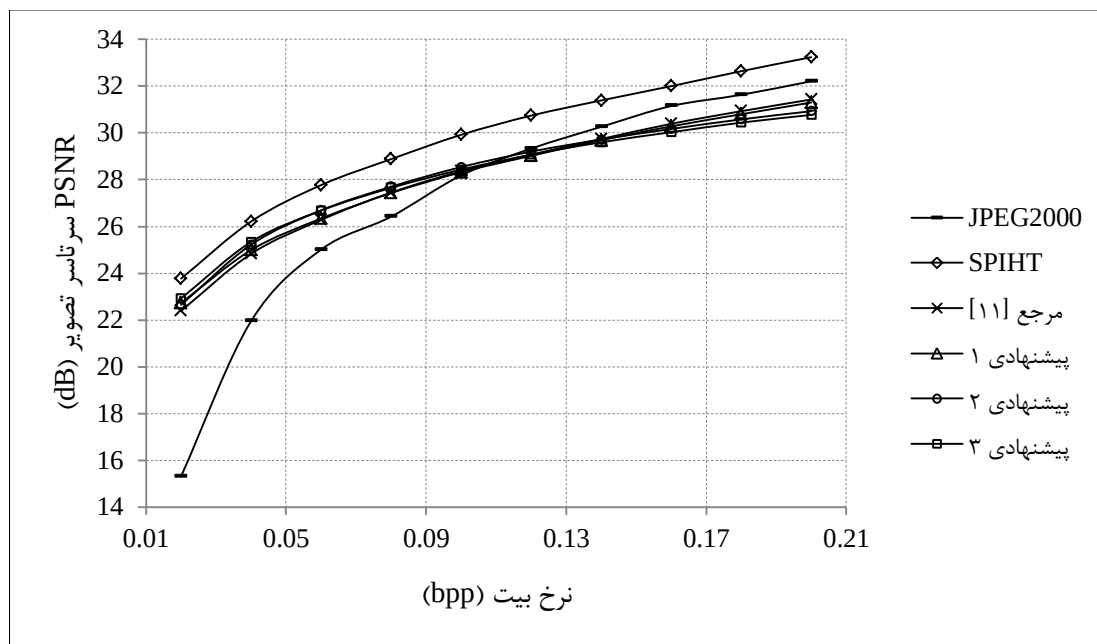
² Personal Computer

مسئله دیگری که با افزایش تعداد سطوح تجزیه موجک باید به آن توجه شود، محدودتر شدن انتخاب‌ها برای اندازه بلوک مورد استفاده در الگوریتم WBTC است. به عنوان مثال در یک تصویر با ابعاد 256×256 پیکسل چنانچه از ۷ سطح تجزیه موجک استفاده شود؛ با توجه به اینکه زیرباند LL بالاترین سطح تجزیه تنها شامل 2×2 ضریب موجک است، اندازه بلوک در الگوریتم WBTC تنها می‌تواند 2×2 و یا 1×1 اختیار شود. بنابراین چنانچه بخواهیم از ماکزیمم سطوح تجزیه موجک ممکن یعنی ۸ استفاده کنیم، الگوریتم WBTC با اختیار اندازه بلوک 1×1 به الگوریتم SPIHT تبدیل خواهد شد. در روش‌های پیشنهادی برای اینکه بتوانیم از قابلیت‌های الگوریتم WBTC همزمان با تعداد سطوح تجزیه موجک بالا بهره ببریم، تعداد سطوح تجزیه موجک را بصورت غیر واقعی و موهومی به الگوریتم WBTC تحویل می‌دهیم. به عنوان مثال فرض کنیم از یک تصویر با ۸ سطح تجزیه تبدیل موجک گرفته شود. حال چنانچه در الگوریتم WBTC به جای ۸ سطح تجزیه، بصورت موهومی ۶ سطح تجزیه تعریف شود، امکان بهره‌گیری از اندازه بلوک‌های 4×4 ، 2×2 و 1×1 فراهم می‌شود. بنابراین با استفاده از این تکنیک نه تنها می‌توان از ویژگی تمرکز بیشتر انرژی در تعداد سطوح تجزیه بالا بهره برد، بلکه با انتخاب اندازه بلوک بزرگتر از 1×1 در الگوریتم WBTC می‌توان از همبستگی‌های درون‌بندی ضرایب نیز به منظور افزایش کارایی کدگذاری و فشرده‌سازی بهره‌برداری کرد.

در شبیه‌سازی‌های انجام شده تعداد سطوح تجزیه واقعی در روش‌های پیشنهادی اول، دوم و سوم به ترتیب ۷، ۸ و ۸ در نظر گرفته شده است. همچنین در هر سه روش پیشنهادی تعداد سطوح تجزیه موهومی و اندازه بلوک‌های مورد استفاده در الگوریتم WBTC به ترتیب ۶ و 4×4 در نظر گرفته شده‌اند. در مورد روش‌های [۱۱] و SPIHT نیز تعداد سطوح تجزیه موجک، ماکزیمم تعداد ممکن یعنی ۸ اختیار شده است.

در [۱۱] به منظور کسب کیفیت بهتر در سرتاسر تصویر، از کل نرخ بیت مورد نیاز برای تصویر فشرده شده، سهم بیشتری به کدگذاری تصویر تقریبی چهره نسبت به تصویر باقیمانده چهره اختصاص داده شده است، چرا که تصویر تقریبی کیفیت سرتاسر تصویر را تحت الشعاع قرار می دهد در حالیکه تصویر باقیمانده تنها روی کیفیت ناحیه چهره تأثیر می گذارد. در روش های پیشنهادی اول و سوم نیز با توجه به بهره گیری از تکنیک کدگذاری ROI ارائه شده در [۱۱]، همین قاعده رعایت شده است. همچنین در روش پیشنهادی دوم با توجه به اینکه ناحیه سر و شانه نسبت به ناحیه چهره بخش بزرگتری از تصویر را تشکیل می دهد، در اینجا نیز به منظور کسب کیفیت بهتر در سرتاسر تصویر، بطور مشابه از کل نرخ بیت مورد نیاز سهم بیشتری به کدگذاری ناحیه سر و شانه نسبت به ناحیه چهره اختصاص داده شده است. در شبیه سازی های انجام شده، ۶۰ درصد از نرخ بیت کل به کدگذاری ناحیه سر و شانه در روش پیشنهادی دوم و تصویر تقریبی چهره در روش پیشنهادی اول و روش [۱۱]، و ۴۰ درصد باقیمانده به کدگذاری ناحیه چهره در روش پیشنهادی دوم و تصویر باقیمانده چهره در روش پیشنهادی اول و روش [۱۱] اختصاص داده شده است. در مورد روش پیشنهادی سوم، ۶۵ درصد از نرخ بیت کل به کدگذاری تصویر تقریبی چهره و ۳۵ درصد باقیمانده به کدگذاری تصویر باقیمانده چهره اختصاص یافته است.

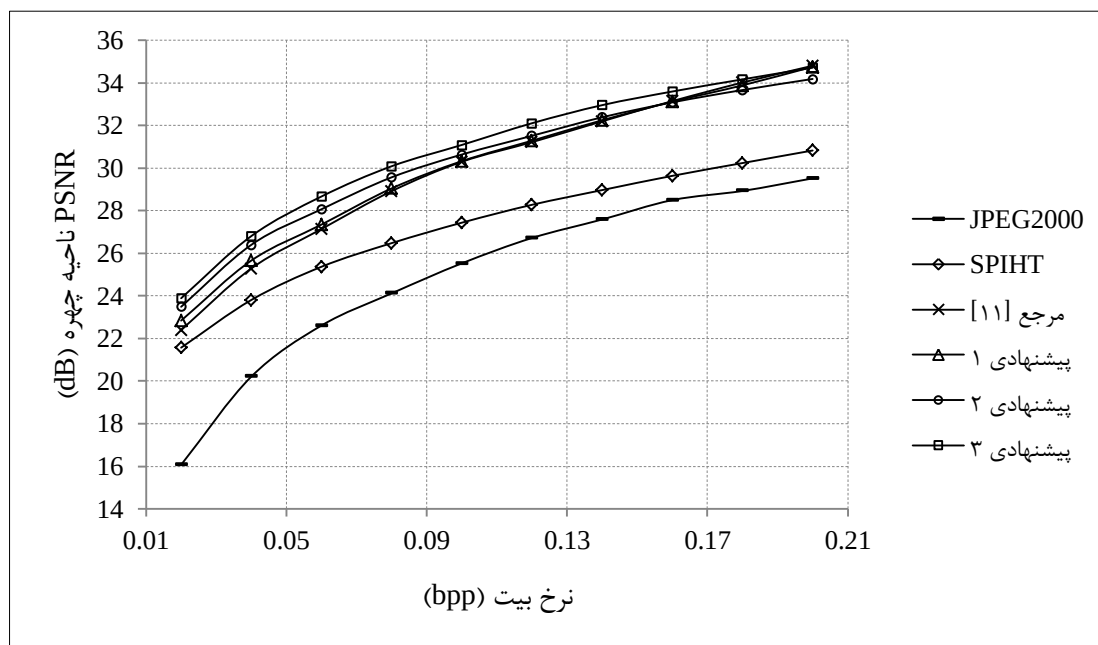
شکل ۲-۴ و شکل ۳-۴ منحنی های نرخ-اعوجاج حاصل از روش های مختلف را به ترتیب برای سرتاسر تصویر و ناحیه چهره نشان می دهد. این منحنی ها با میانگین گیری روی ۴۰ تصویر پایگاه داده بدست آمده اند. از شکل ۲-۴ ملاحظه می کنیم که روش SPIHT نسبت به سایر روش های مورد مقایسه همواره کارایی بالاتری را در سرتاسر تصویر ارائه می دهد. همچنین هر سه روش پیشنهادی در مقایسه با استاندارد JPEG2000 کارایی بالاتری را در نرخ بیت های کمتر از ۰.۱ bpp ارائه می دهند.



شکل ۴-۲: منحنی‌های نرخ-اعوجاج سرتاسر تصویر برای روش‌های مختلف.

این کارایی بالا در نرخ بیت‌های پایین تر خود را بهتر نشان می‌دهد، بطوریکه در نرخ بیت ۰.۰۲ bpp روش‌های پیشنهادی اول تا سوم به ترتیب به بهبودی حدود ۷.۴۲ dB، ۷.۳۴ dB و ۷.۶۱ dB در کارایی نرخ-اعوجاج نسبت به استاندارد JPEG2000 دست یافته‌اند که بسیار قابل توجه است. در مقایسه با روش [۱۱]، در روش‌های پیشنهادی سعی شده است که در نرخ بیت‌های پایین کارایی نرخ-اعوجاج در سرتاسر تصویر حداقل با روش [۱۱] برابر بوده و حتی امکان بهبود یابد؛ چرا که در این پایان‌نامه هدف اصلی بهبود کارایی و کیفیت ناحیه چهره در نرخ بیت‌های پایین، و در درجه دوم حفظ حداقل کیفیت مورد نیاز برای سرتاسر تصویر است. روش‌های پیشنهادی در نرخ بیت‌های کمتر از ۰.۱۲ bpp کارایی قابل مقایسه و حتی بالاتری را نسبت به روش [۱۱] ارائه می‌دهند. همانطور که انتظار می‌رود با کاهش نرخ بیت کارایی روش‌های پیشنهادی در مقایسه با روش [۱۱] افزایش می‌یابد، بطوریکه در نرخ بیت ۰.۰۲ bpp روش‌های اول تا سوم به ترتیب به بهبودی حدود ۰.۳۴ dB، ۰.۲۶ dB و ۰.۵۳ dB را در مقایسه با روش [۱۱] نشان می‌دهند.

شکل ۳-۴ کارایی بالاتر روش‌های پیشنهادی را در ناحیه چهره در مقایسه با سایر روش‌ها به اثبات می‌رساند. همانطور که مشاهده می‌شود در تمامی نرخ‌بیت‌های مورد آزمایش، روش‌های پیشنهادی همواره بهتر از استاندارد JPEG2000 و روش SPIHT عمل می‌کنند. در نرخ‌بیت‌های کمتر از ۵.۲۴ dB، روش‌های پیشنهادی اول تا سوم در مقایسه با JPEG2000 به ترتیب بهبودی فراتر از ۵.۲۴ dB و ۴.۶۷ dB و ۵.۲ dB را نشان می‌دهند و این بهبود با کاهش نرخ‌بیت افزایش می‌یابد. همچنین در مقایسه با SPIHT این روش‌ها توانسته‌اند به ترتیب به حداقل بهبودی حدود ۱.۲۷ dB، ۱.۹۲ dB و ۲.۳۲ dB دست یابند، با این تفاوت که در اینجا با افزایش نرخ‌بیت به بهبود بیشتری دست خواهیم یافت. روش [۱۱] نیز در مقایسه با استاندارد JPEG2000 و روش SPIHT از نظر کارایی در ناحیه چهره، رفتار تقریباً مشابهی با روش‌های پیشنهادی را از خود نشان می‌دهد. البته روش‌های پیشنهادی در مقایسه با روش مبتنی بر ناحیه مطلوب [۱۱] نیز توانسته‌اند به کارایی بالاتری در ناحیه چهره، مخصوصاً در نرخ‌بیت‌های پایین دست یابند که بدین شرح است:



شکل ۳-۴: منحنی‌های نرخ-عوجاج ناحیه چهره برای روش‌های مختلف.

۱) در محدوده نرخ‌بیت‌های ۰.۱۴ bps تا ۰.۰۲ bps، دستیابی به بهبود کارایی از ۰.۰۵ dB تا ۰.۴۴ dB در روش پیشنهادی اول.

۲) در محدوده نرخ بیت‌های ۰.۱۴ bps تا ۰.۰۲ bps دستیابی به بهبود کارایی از ۰.۱۹ dB تا ۰.۰۹ dB در روش پیشنهادی دوم.

۳) در محدوده نرخ‌بیت‌های ۰.۱۸ bps تا ۰.۰۲ bps دستیابی به بهبود کارایی از ۰.۱۵ dB تا ۱.۴۹ dB در روش پیشنهادی سوم.

همانطور که در شکل ۴-۲ و شکل ۴-۳ مشاهده می‌شود، کارایی روش‌های پیشنهادی با افزایش نرخ‌بیت به تدریج کاهش پیدا می‌کند. یک علت این امر کاهش کارایی کدگذاری WBTC با افزایش نرخ‌بیت است. در کدگذاری WBTC با افزایش نرخ‌بیت، تعداد صفحه‌های بیت‌ی بیشتری از ضرایب باید ارسال شود، و هرچه به صفحه‌های بیت‌ی پایین‌تر نزدیک می‌شویم، تعداد ضرایب مهم بیشتر و بیشتر می‌شود. این امر منجر به کاهش همبستگی درون‌باندی ضرایب خواهد شد، و بنابراین تعداد بلوک‌های صفر کاهش و تعداد بلوک‌های مهم افزایش می‌یابد. در نتیجه تعداد بیت بیشتری بایستی صرف تقسیم بلوک‌های مهم به منظور یافتن ضرایب مهم گردد. هرچه اندازه بلوک‌ها در WBTC بزرگ‌تر در نظر گرفته شود، تعداد بیت بیشتری باید صرف شکسته شدن بلوک‌ها گردد. علت دوم نیز کاهش کارایی تکنیک بازآرایی بلوک‌های ضرایب با افزایش نرخ‌بیت است. همانطور که در بخش ۳-۶ ذکر شد، برای بدست آوردن آرایه‌های ترتیب اهمیت بلوک‌های ضرایب در فاز شبه-آموزش از تکنیک ساده میانگین‌گیری بلوک‌های ضرایب روی کل پایگاه داده و سپس مرتب کردن بلوک‌ها بر حسب متوسط اندازه استفاده شده است. احتمال درست بودن ترتیب اهمیت هر بلوک در این روش وابسته به اندازه ضرایب موجک درون بلوک مربوطه در فاز شبه-آموزش است. بطوریکه هر چه اندازه ضرایب درون یک بلوک در فاز شبه-آموزش بزرگ‌تر باشد احتمال درست بودن ترتیب اهمیت بدست آمده برای بلوک

مربوطه بالاتر است، چرا که ضرایب موجک بزرگتر متناظر با کلیات تصویر چهره هستند و این تکنیک ساده میانگین گیری نیز بسیار وابسته به کلیات چهره و یکسان بودن آن‌ها در کل پایگاه داده است. بنابراین در نرخ بیت‌های بالا، هنگامی که الگوریتم WBTC به سراغ کدگذاری ضرایب موجک کوچکتر می‌آید، با توجه به احتمال بالای خطا در ترتیب قرارگیری این ضرایب، الگوریتم WBTC مجبور به صرف زمان و نرخ بیت بیشتری برای جستجو و یافتن این ضرایب شده و در نتیجه کارایی آن کاهش می‌یابد. البته در نرخ بیت‌های بالا، کارایی روش‌های پیشنهادی در ناحیه چهره نسبت به سرتاسر تصویر شاهد افت بسیار کمتری است. علت این امر اختصاص حداقل ۴۰ درصد از نرخ بیت کل به کدگذاری ناحیه چهره است، که نسبت به کل تصویر ناحیه نسبتاً کوچکی به شمار می‌رود. با افزایش نرخ بیت، میزان نرخ بیت اختصاص یافته به ناحیه چهره نیز افزایش یافته و در نرخ بیت‌های بالا به میزان قابل توجهی می‌رسد.

جدول ۴-۱ نتایج حاصل از بازسازی یک تصویر از پایگاه داده را با استفاده از روش‌های مختلف در چندین نرخ بیت نشان می‌دهد. در این جدول علاوه بر نمایش تصاویر بازسازی شده حاصل از روش‌های مختلف به منظور مقایسه کیفیت دیداری^۱، با نمایش مقدار کمیت بیشینه سیگنال به نویز (PSNR) سرتاسر تصویر و ناحیه چهره امکان مقایسه کمی نتایج حاصل از روش‌های مختلف نیز فراهم شده است. همانطور که مشاهده می‌شود، استاندارد JPEG2000 در نرخ بیت‌های کمتر از ۰.۱۰۴ bpp کیفیت دیداری ضعیفی را ارائه می‌دهد. این ضعف در نرخ بیت ۰.۰۴۴ bpp به اوج خود رسیده، بطوریکه در این نرخ بیت ویژگی‌های چهره بطور کلی از دست رفته‌اند. روش SPIHT نسبت به JPEG2000 عملکرد بهتری داشته و تا نرخ بیت ۰.۰۸۸ bpp کیفیت دیداری ویژگی‌های چهره را در حد

¹ Visual Quality

جدول ۴-۱: مقایسه کمی و کیفی یک تصویر بازسازی شده با استفاده از روش های مختلف در چندین نرخ بیت.

	تصویر اصلی	تصویر بازسازی شده			
		نرخ بیت ۰.۱ bpp	نرخ بیت ۰.۰۸ bpp	نرخ بیت ۰.۰۶ bpp	نرخ بیت ۰.۰۴ bpp
JPEG2000					
		PSNR(dB)	PSNR(dB)	PSNR(dB)	PSNR(dB)
		سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره
		۲۴.۹۳	۲۷.۳۶	۲۳.۳۵	۲۶.۰۵
SPIHT					
		PSNR(dB)	PSNR(dB)	PSNR(dB)	PSNR(dB)
		سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره
		۲۶.۶۲	۲۸.۵۴	۲۵.۸۶	۲۷.۶۸
[۱۱]					
		PSNR(dB)	PSNR(dB)	PSNR(dB)	PSNR(dB)
		سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره	سرتاسر تصویر ناحیه چهره
		۲۹.۳۴	۲۷.۴۸	۲۶.۷۹	۲۵.۹۹

قابل قبولی حفظ کرده است. اما با کاهش بیشتر نرخ بیت، کارایی SPIHT در ناحیه چهره افت کرده، بطوریکه در نرخ بیت ۰.۰۴ bpp ویژگی های چهره یا به زحمت قابل تشخیص هستند و یا از دست رفته اند. در روش [۱۱] به کمک تکنیک کدگذاری ROI، کیفیت دیداری ناحیه ویژگی های چهره در نرخ بیت های پایین بهبود داده شده، بطوریکه تا نرخ بیت ۰.۰۶ bpp ویژگی های چهره به خوبی قابل تشخیص هستند. در نرخ بیت های کمتر از ۰.۰۶ bpp کارایی این روش نیز در ناحیه چهره افت می کند. در نرخ بیت ۰.۰۴ bpp، اگرچه ممکن است ویژگی های چهره قابل تشخیص به نظر برسند، اما احتمال

ادامه جدول ۴-۱

	تصویر اصلی	تصویر بازسازی شده							
		نرخ بیت ۰.۱ bpp		نرخ بیت ۰.۰۸ bpp		نرخ بیت ۰.۰۶ bpp		نرخ بیت ۰.۰۴ bpp	
پیشنهادی ۱									
		PSNR(dB)		PSNR(dB)		PSNR(dB)		PSNR(dB)	
		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره	
		۲۹.۵۷		۲۸.۱۲		۲۶.۲۱		۲۵.۰۴	
پیشنهادی ۲									
		PSNR(dB)		PSNR(dB)		PSNR(dB)		PSNR(dB)	
		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره	
		۳۰.۱۰		۲۹.۰۷		۲۷.۴۲		۲۵.۴۶	
پیشنهادی ۳									
		PSNR(dB)		PSNR(dB)		PSNR(dB)		PSNR(dB)	
		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره		سرتاسر تصویر ناحیه چهره	
		۳۰.۴۱		۲۹.۵۸		۲۸.۳۹		۲۶.۵۹	

اینکه سیستم‌های تشخیص هویت با این تصویر دچار مشکل شوند نیز بالا است. در روش پیشنهادی اول با بهره‌گیری از تکنیک کدگذاری ROI ارائه شده در [۱۱] و استفاده از تکنیک بازآرایی بلوک‌های ضرایب و کدگذاری WBTC توانسته‌ایم نتایج حاصل از روش [۱۱] را مخصوصاً در نرخ بیت‌های پایین تا حدودی بهبود دهیم. این بهبود بصورت کمی و کیفی در جدول ۴-۱ قابل مشاهده است. نتایج حاصل از روش پیشنهادی دوم، کیفیت دیداری بهتری را هم در ناحیه چهره و هم در سرتاسر تصویر در مقایسه با روش [۱۱] و روش پیشنهادی اول نشان می‌دهند. کدگذاری و فشرده‌سازی کاملاً مجزای

نواحی تصویر با استفاده از تبدیل RWT یکی از عوامل این کیفیت دیداری بالا است. این ویژگی همچنین این امکان را در اختیارمان قرار می‌دهد که در نرخ‌بیت‌های خیلی پایین با افزایش سهم ناحیه چهره از نرخ‌بیت کل و کاهش سهم سایر نواحی، بتوانیم به کیفیت دیداری بالاتری دست یابیم. در روش پیشنهادی سوم با استفاده ترکیبی از تبدیل RWT و تکنیک کدگذاری ROI ارائه شده در [۱۱]، کیفیت دیداری ناحیه چهره در مقایسه با سایر روش‌ها بهبود قابل توجهی داشته است. این بهبود در نرخ‌بیت‌های خیلی پایین قابل توجه‌تر است، بطوریکه در نرخ‌بیت ۰.۰۴ bpp ویژگی‌های چهره به خوبی قابل تشخیص هستند. از طرفی در این روش همانطور که در جدول ۴-۱ مشاهده می‌شود، عناصر مزاحم بصری اطراف مرز سر و شانه با لایه‌گذاری پس‌زمینه تضعیف شده‌اند. البته در روش پیشنهادی دوم نیز به روشی مشابه این عمل صورت گرفته است. این عناصر مزاحم بصری که در اثر نشت ضرایب موجک بوجود می‌آیند در نتایج حاصل از روش‌های SPIHT، روش [۱۱] و روش پیشنهادی اول و مخصوصاً در نرخ‌بیت‌های پایین شدت بیشتری دارند. در جدول ۴-۲ تا جدول ۴-۵ با نمایش چند تصویر بازسازی‌شده از پایگاه داده با استفاده از روش‌های مختلف در نرخ‌بیت‌های ۰.۱، ۰.۰۸، ۰.۰۶ و ۰.۰۴ bpp، امکان مقایسه بیشتر کیفیت دیداری نتایج حاصل از روش‌های مختلف فراهم آمده است.

جدول ۴-۲: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۱.

تصاویر بازسازی شده در نرخ بیت ۰.۱ bps	تصاویر مرجع [۱۱]	تصاویر اصلی				
پیشنهادی ۳	پیشنهادی ۲	پیشنهادی ۱	SPIHT	JPEG2000		

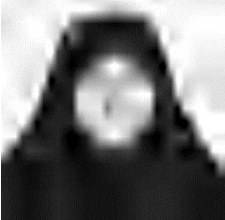
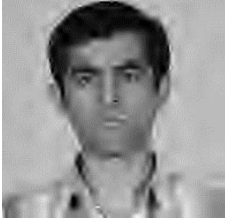
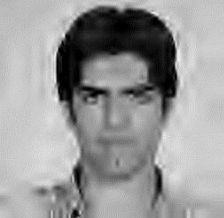
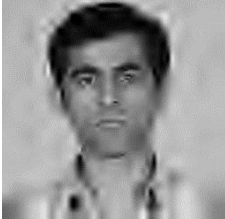
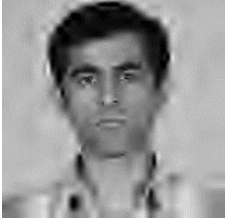
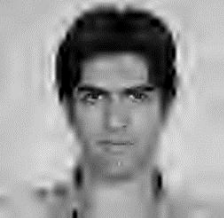
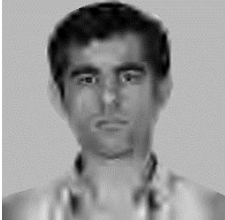
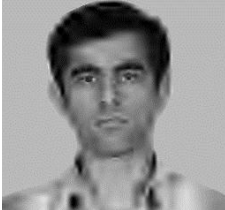
جدول ۳-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۰۸.

تصاویر اصلی				
	JPEG2000	SPIHT	مرجع [۱۱]	پیشنهادی ۱
				پیشنهادی ۲
				پیشنهادی ۳

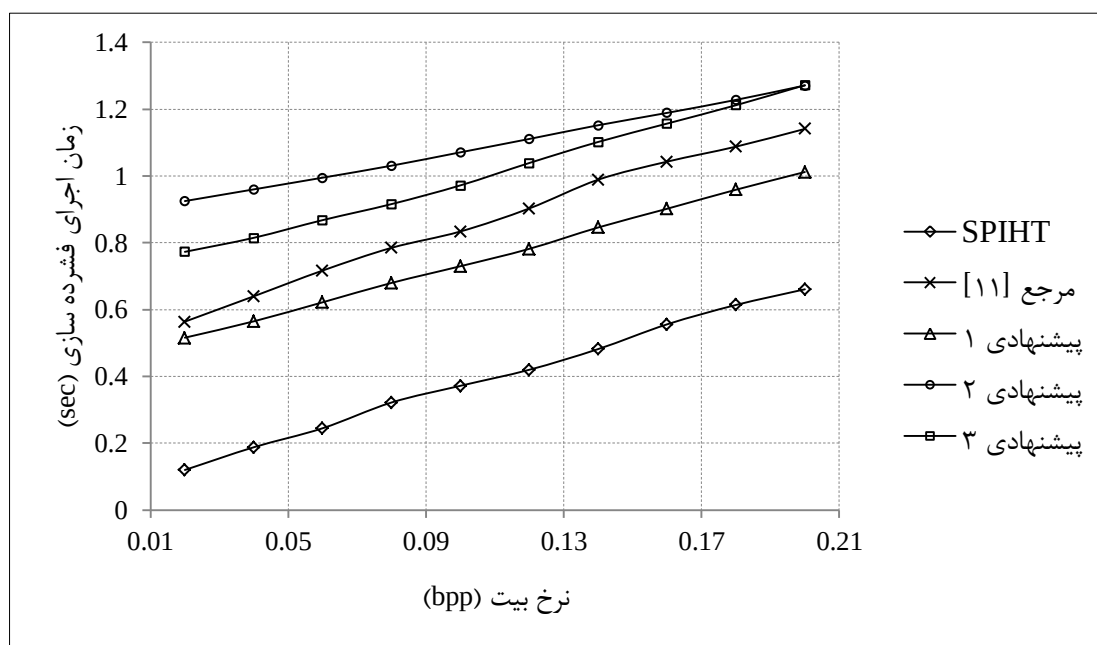
جدول ۴-۴: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۰۶.

تصاویر اصلی				
	JPEG2000	SPIHT	مرجع [۱۱]	پیشنهادی ۱
				پیشنهادی ۲
				پیشنهادی ۳

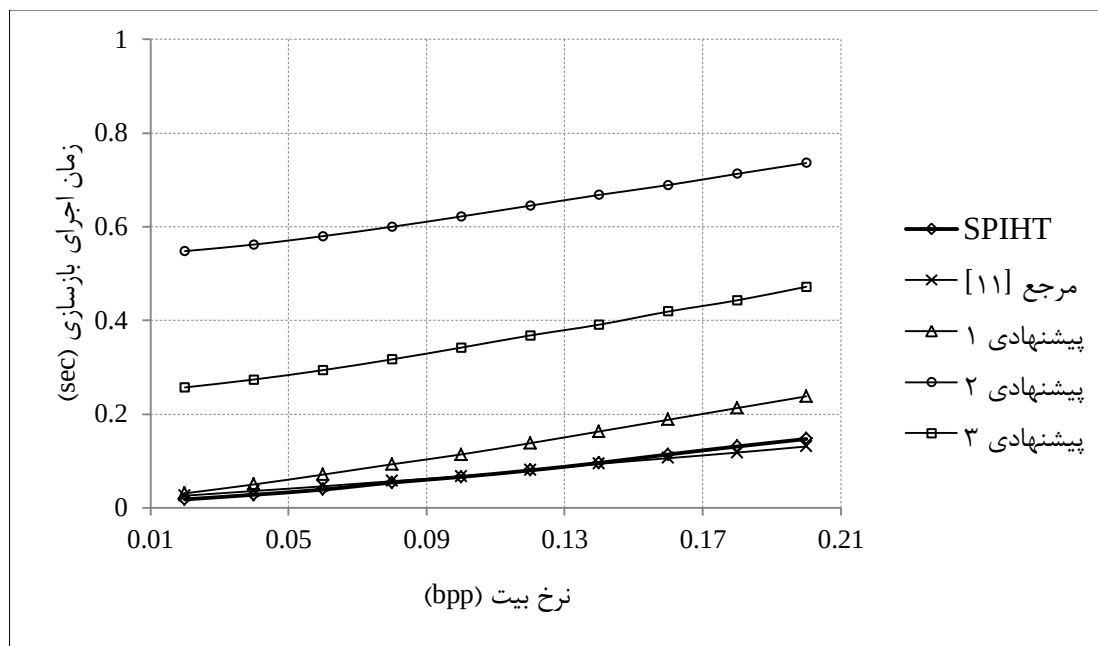
جدول ۴-۵: مقایسه کیفیت دیداری چند تصویر بازسازی شده با استفاده از روش های مختلف در نرخ بیت ۰.۰۴.

تصاویر بازسازی شده در نرخ بیت ۰.۰۴ bpp	تصویر اصلی				
	JPEG2000				
	SPIHT				
	مرجع [۱۱]				
	پیشنهادی ۱				
	پیشنهادی ۲				
	پیشنهادی ۳				

شکل ۴-۴ و شکل ۵-۴ به ترتیب منحنی‌های زمان اجرای فشرده‌سازی و بازسازی تصاویر را برای روش‌های مختلف نشان می‌دهند. این منحنی‌ها با میانگین‌گیری روی ۴۰ تصویر پایگاه داده بدست آمده‌اند. همانطور که انتظار می‌رود، روش SPIHT بیشترین سرعت را در فشرده‌سازی و بازسازی تصاویر نسبت به سایر روش‌ها دارد، چرا که در این روش کدگذاری و کدگشایی تصویر تنها در یک مرحله صورت گرفته و لذا پیچیدگی محاسباتی پایینی دارد. روش [۱۱] تنها شامل دو مرحله کدگذاری تصویر به روش SPIHT و یک مرحله کدگذاری Maxshift است، که هر دو تکنیک پیچیدگی محاسباتی پایینی دارند. همین مسئله باعث شده است که روش [۱۱] نیز در شکل ۴-۴ و شکل ۵-۴ سرعت بالایی را در فشرده‌سازی و بازسازی تصاویر نشان دهد. در روش پیشنهادی اول با استفاده از الگوریتم کدگذاری WBTC توانسته‌ایم سرعت فشرده‌سازی را نسبت به روش [۱۱] کاهش دهیم، چرا که الگوریتم WBTC در فاز کدگذاری پیچیدگی محاسباتی پایین‌تری نسبت به الگوریتم SPIHT دارد. اما با توجه به افزایش پیچیدگی محاسباتی الگوریتم WBTC در فاز کدگشایی با افزایش



شکل ۴-۴: منحنی‌های نرخ-زمان اجرای فشرده‌سازی برای روش‌های مختلف.

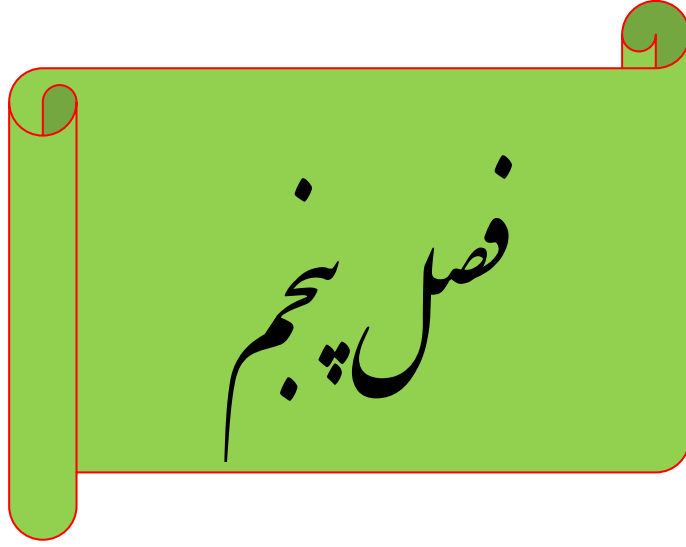


شکل ۴-۵: منحنی‌های نرخ-زمان اجرای بازسازی برای روش‌های مختلف.

نرخ‌بیت، سرعت بازسازی روش پیشنهادی اول نسبت به روش [۱۱] با افزایش نرخ‌بیت کاهش می‌یابد. با این حال باز هم روش پیشنهادی اول، در نرخ‌بیت‌های پایین روشی بسیار سریع محسوب می‌شود، چرا که قادر است در این نرخ‌بیت‌ها یک تصویر 256×256 پیکسلی را در حداکثر 0.73 ثانیه فشرده، و در حداکثر 0.12 ثانیه بازسازی کند. در روش پیشنهادی دوم با افزایش پیچیدگی محاسباتی، سرعت فشرده‌سازی و بازسازی تصاویر نسبت به روش [۱۱] و روش پیشنهادی اول کاهش یافته است. همانطور که در بخش‌های قبل گفته شد، علت این امر فشرده‌سازی نواحی بصورت مجزا، اضافه شدن تعداد مراحل بازآرایی بلوک‌های ضرایب بسته به تعداد نواحی مطلوب و نیز اضافه شدن مراحل کدگذاری و کدگذاری مرز نواحی مطلوب در روش پیشنهادی دوم است. همچنین به این عوامل باید عدم امکان پیاده‌سازی ماتریسی تبدیل RWT را نیز افزود. البته در این روش امکان فشرده‌سازی و بازسازی نواحی مختلف تصویر بصورت موازی وجود دارد، که در صورت استفاده از این قابلیت سرعت فشرده‌سازی و بازسازی بسته به تعداد نواحی چند برابر خواهد شد. در روش پیشنهادی سوم با تلفیق

ایده‌های استفاده شده در روش‌های پیشنهادی اول و دوم، توانسته‌ایم علاوه بر دستیابی به کارایی نرخ-اعوجاج بالاتر نسبت به روش پیشنهادی دوم، سرعت فشرده‌سازی و بازسازی تصاویر را نیز نسبت به روش دوم افزایش دهیم. باید توجه داشت که حدود ۰.۴۱ ثانیه از مدت زمان اجرای فشرده‌سازی در روش‌های پیشنهادی، مربوط به تکنیک جداسازی نواحی مطلوب و غیرمطلوب است؛ و لذا با استفاده از تکنیک‌هایی سریع‌تر می‌توان سرعت فشرده‌سازی را در روش‌های پیشنهادی بهبود داد. جدول ۴-۶ متوسط زمان اجرای تکنیک‌های استفاده شده در روش‌های پیشنهادی را نشان می‌دهد. این مدت زمان‌ها با میانگین‌گیری از زمان اجرای تکنیک‌ها روی ۴۰ تصویر پایگاه داده بدست آمده‌اند. جدول ۴-۶: متوسط زمان اجرای تکنیک‌های استفاده شده در روش‌های پیشنهادی روی ۴۰ تصویر پایگاه داده.

متوسط زمان اجرا (sec)	تکنیک استفاده شده
۰.۴۱۴	جداسازی نواحی مطلوب و غیرمطلوب
۰.۰۷۴	اعمال تبدیل RWT روی ناحیه چهره با ۸ سطح تجزیه موجک
۰.۰۸۵	اعمال تبدیل معکوس RWT روی ناحیه چهره تبدیل‌یافته با ۸ سطح تجزیه موجک
۰.۰۰۶	استخراج مرز پیش‌زمینه و کدگذاری آن با فرض $n = 13$
۰.۰۱۵	کدگذاری و بازسازی مرز پیش‌زمینه و تولید ماسک پیش‌زمینه با فرض $n = 13$



نتیجہ گیری

۵-۱ نتیجه گیری

در این پایان نامه طی یک سیر تکاملی سه روش برای فشرده سازی مبتنی بر ناحیه مطلوب تصاویر چهره پرسنلی در نرخ بیت های پایین پیشنهاد شد، که تبدیل DWT و الگوریتم کدگذاری WBTC پایه و اساس هر سه روش پیشنهادی را تشکیل می دهند. همچنین در این پایان نامه تکنیک هایی برای بهبود کارایی کدگذاری WBTC، جداسازی نواحی مطلوب و غیر مطلوب و کدگذاری مرز نواحی ارائه شدند. سپس کارایی روش های فشرده سازی پیشنهادی چه از نظر نرخ-اعوجاج و چه از نظر سرعت فشرده سازی و بازسازی در نرخ بیت های مختلف با روش هایی همچون استاندارد JPEG2000، روش SPIHT و یک روش مبتنی بر ناحیه مطلوب ([۱۱]) مقایسه شد.

در روش اول از تکنیک کدگذاری ROI ارائه شده در [۱۱] برای بهبود کیفیت ناحیه چهره در نرخ بیت های پایین استفاده شده است. تلفیق این تکنیک با تکنیک بازآرایی بلوک های ضرایب موجک و الگوریتم کدگذاری WBTC باعث شده است که کیفیت ناحیه چهره در نرخ بیت های پایین نسبت به روش [۱۱] بهبود یابد، و همچنین کارایی آن در نرخ بیت های پایین از استاندارد JPEG2000 و روش SPIHT بالاتر باشد. در این روش بازسازی تصاویر در نرخ بیت های پایین با سرعت بالایی انجام گرفته و سرعت فشرده سازی نیز در حد مناسبی است. بنابراین این روش مناسب کاربردهایی است که نیازمند فشرده سازی و بازسازی سریع و نیز ارسال و دریافت سریع تصاویر پرسنلی در نرخ بیت های پایین هستند، که از آن جمله می توان سیستم های تشخیص هویت بلادرنگ را مثال زد.

در روش دوم از تکنیک جدیدی برای فشرده سازی مبتنی بر ROI تصاویر چهره پرسنلی بهره گرفتیم. در این روش با استفاده از تبدیل موجک مبتنی بر ناحیه (RWT) [۱۴]، فشرده سازی نواحی مطلوب و غیرمطلوب بطور کاملاً مجزا صورت گرفته است. بنابراین در این روش می توان هر ناحیه را

دقیقا با نرخ بیت دلخواه و مورد نیاز فشرده و بازسازی کرد. یک عیب این روش نیاز تبدیل RWT و معکوس آن به ماسک نواحی مطلوب است، که در نتیجه کدگذاری مرز نواحی مطلوب و ارسال آن به واحد کدگشا ضروری خواهد بود. در این روش برای کدگذاری مرز نواحی چهره و سر و شانه به ترتیب از تکنیک‌هایی مبتنی بر اختلاف فاصله رئوس چندضلعی و چندجمله‌ای درونیاب بهره گرفتیم. از طرفی سرعت فشرده‌سازی و بازسازی تصاویر در این روش نسبت به روش اول کمتر است، چرا که پیچیدگی محاسباتی در این روش افزایش یافته است، و علت آن را می‌توان فشرده‌سازی نواحی بصورت مجزا، اضافه شدن تعداد مراحل بازآرایی بلوک‌های ضرایب بسته به تعداد نواحی مطلوب و نیز اضافه شدن مراحل کدگذاری و کدگشایی مرز نواحی مطلوب دانست. البته در این روش امکان فشرده‌سازی و بازسازی نواحی مختلف تصویر بصورت موازی وجود دارد، که در صورت استفاده از این قابلیت سرعت فشرده‌سازی و بازسازی بسته به تعداد نواحی چند برابر خواهد شد. نتایج شبیه‌سازی بیانگر کارایی بالاتر این روش در ناحیه چهره و در نرخ بیت‌های پایین در مقایسه با روش پیشنهادی اول، روش [۱۱]، استاندارد JPEG2000 و نیز SPIHT است.

روش سوم در حقیقت تلفیقی از روش‌های اول و دوم است. در این روش با بهره‌گیری توأم از تبدیل RWT و ایده کدگذاری ROI ارائه شده در [۱۱]، توان استیم علاوه بر دستیابی به کارایی بالاتر نرخ-عوجاج در نرخ بیت‌های پایین نسبت به روش‌های اول و دوم، سرعت فشرده‌سازی و بازسازی تصاویر را نیز نسبت به روش دوم افزایش دهیم. در این روش نیز مشابه روش اول، عمل فشرده‌سازی/بازسازی سرتاسر تصویر تا یک نرخ بیت مشخص بطور همزمان صورت می‌گیرد. سپس برای افزایش کیفیت نواحی مطلوب، با اعمال تبدیل RWT روی این نواحی در تصویر باقیمانده و کدگذاری آن‌ها با نرخ‌بیتی دیگر، اهمیت و نرخ‌بیت بیشتری به این نواحی اختصاص داده می‌شود. این روش در مقایسه با روش‌های پیشنهادی اول و دوم و به تبع آن روش [۱۱]، JPEG2000 و SPIHT

کارایی بالاتری را در ناحیه چهره و بخصوص در نرخ بیت های پایین، در شبیه سازی های انجام شده نشان داده است.

نتایج حاصل کار این پایان نامه را می توان بطور خلاصه بصورت زیر بیان کرد:

(۱) بطور کلی می توان گفت روش های فشرده سازی مبتنی بر فرهنگ داده در صورتی که به درستی طراحی شده باشند، در فشرده سازی پایگاه های بسیار بزرگ تصاویر چهره نسبت به روش های مبتنی بر تبدیل کارا تر خواهند بود. در حقیقت این روش ها قادرند با بهره گیری از شباهت موجود میان تصاویر چهره در موقعیت های مکانی مشابه، تزايد اطلاعات میان تصاویر مجموعه ای بزرگ را به مقدار زیادی کاهش دهند. البته مسائل چالش برانگیزی همچون انتخاب الگوریتم مناسب برای آموزش فرهنگ داده ها و نیز تعیین اندازه مناسب بلوک ها، معمولا طراحی این روش ها را مشکل و پیچیده می سازد. همچنین این روش ها نیازمند زمانی نسبتا طولانی برای آموزش فرهنگ داده ها روی مجموعه ای بزرگی از تصاویر، و نیز حافظه جانبی نسبتا زیادی برای ذخیره سازی این فرهنگ داده ها در هر دو واحد کدگذار و کدگشا هستند. با کمی سبک سنگین کردن مزایا و معایب روش های مبتنی بر فرهنگ داده متوجه خواهیم شد که معایب این روش ها وزن بیشتری دارند، و لذا به نظر استفاده از روش های مبتنی بر تبدیل به صرفه تر می آید.

(۲) در میان تبدیلات مورد استفاده در روش های مبتنی بر تبدیل، استفاده از تبدیل DWT به دلیل ویژگی های کم نظیر آن از قبیل قابلیت فوق العاده در متراکم سازی انرژی، رایج تر است.

(۳) با استفاده از تبدیل DWT و الگوریتم های کدگذاری تصویر مبتنی بر موجک می توان کدگذاری و فشرده سازی تصاویر را در هر نرخ بیت دلخواهی متوقف کرده و دقیقا به نرخ بیت مورد نیاز دست یافت.

- (۴) در میان الگوریتم‌های کدگذاری تصویر مبتنی بر موجک، الگوریتم WBTC با بهره‌گیری همزمان از همبستگی درون‌بندی و بین‌بندی ضرایب موجک، علاوه بر کارایی بالا سرعت بالایی را نیز در کدگذاری و کدگشایی، مخصوصاً در نرخ‌بیت‌های پایین ارائه می‌دهد.
- (۵) تبدیل موجک مبتنی بر ناحیه (RWT) با قابلیت اعمال بر روی نواحی و یا اشیاء تصویر با شکل دلخواه، امکان کدگذاری و فشرده‌سازی مجزای نواحی مطلوب تصویر را فراهم کرده است. یکی از معایب این تبدیل عدم امکان پیاده‌سازی آن بصورت ماتریسی به منظور افزایش سرعت اجرا است. همچنین این تبدیل و معکوس آن برای اعمال بر روی ناحیه مطلوب نیازمند ماسک ROI هستند. بنابراین کدگذاری مرز ناحیه مطلوب در واحد کدگذار و ارسال آن به واحد کدگشا الزامی بوده، که این مسئله کارایی فشرده‌سازی را کمی کاهش می‌دهد.
- (۶) با بلوک‌بندی ماتریس ضرایب حاصل از تبدیل موجک تصویر، و سپس بازآرایی بلوک‌های ضرایب بر حسب اهمیت و مطابق با الگوی جستجوی مورد استفاده در الگوریتم کدگذاری WBTC، می‌توان کارایی کدگذاری WBTC را بهبود بخشید؛ چرا که در این حالت زمان و نرخ‌بیت کمتری صرف یافتن ضرایب مهم یا همان ضرایب بزرگتر خواهد شد.

۵-۲ پیشنهادات

همانطور که در بخش ۱-۲ گفته شد، مطالعات و تحقیقات انجام شده تاکنون در زمینه فشرده‌سازی تصاویر چهره با تکیه بر نواحی مطلوب بسیار محدود است. بنابراین این پایان‌نامه می‌تواند نقطه شروعی برای تحقیقات و کارهای آینده در زمینه فشرده‌سازی مبتنی بر ناحیه مطلوب تصاویر چهره باشد. در ادامه چند پیشنهاد که می‌تواند در این زمینه موثر باشد بیان خواهد شد.

- ۱) استفاده از تکنیکی مناسب‌تر در فاز شبه-آموزش برای بدست آوردن آرایه‌های ترتیب اهمیت، چرا که هر چه بتوان این آرایه‌های ترتیب را با دقت بالاتری بدست آورد، کارایی تکنیک بازآرایی ضرایب و متعاقبا کارایی کدگذاری و فشرده‌سازی افزایش خواهد یافت. به عنوان مثال می‌توان همانند روش‌های [۴-۶] از یک مرحله هم‌ترازی و انطباق تصاویر در فاز شبه-آموزش و نیز فشرده‌سازی و بازسازی تصاویر بهره برد. این مرحله هم‌ترازی با نگاشت ویژگی‌های چهره به موقعیت‌های مکانی مشابه آن‌ها را از نظر هندسی به یک فرم معیار تغییر شکل می‌دهد. تبدیل تصاویر به یک فرم معیار این روش‌ها را قادر ساخته است که به خوبی از شباهت تصاویر چهره در موقعیت‌های مکانی مشابه بهره ببرند.
- ۲) استفاده از تکنیکی برای تبدیل تصویر چهره به یک تصویر با محتوای فرکانس پایین [۴۰]، زیرا در این حالت بیشترین انرژی تصویر در زیرباندهای فرکانسی پایین متمرکز شده و در نتیجه فشرده‌سازی تصویر با سهولت و کارایی بالاتری انجام می‌گیرد.
- ۳) تطبیق کدگذاری WBTC و یا سایر الگوریتم‌های کدگذاری در حوزه موجک، برای در دست گرفتن و کدگذاری نواحی و یا اشیاء تصویر با شکل دلخواه.
- ۴) بهره‌گیری از تزايد و همبستگی اطلاعات موجود در تقارن چهره با استفاده از تکنیک‌هایی همچون تطبیق بلوکی.

- [1] G. K. Wallace, "The JPEG Still Picture Compression Standard," *Communications of the ACM*, vol. 34, pp. 30-44, 1991.
- [2] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 Still Image Coding System: An Overview," *IEEE Trans Consum Electron*, vol. 46, pp. 1103–27, 2000.
- [3] W. Zhao, R. Chellappa, and P.J. Philips, "Face Recognition: A Literature Survey," *ACM Comput Surv*, vol. 35, pp. 399-458, 2003.
- [4] M. Elad, R. Goldenberg, and R. Kimmel, "Low Bit-Rate Compression of Facial Images," *IEEE Trans Image Process*, vol. 16, pp. 2379–83, 2007.
- [5] O. Bryt and M. Elad, "Compression of Facial Images Using the K-SVD Algorithm," *J Vis Commun Image R*, vol. 19, pp. 3445–62, 2008.
- [6] O. Bryt and M. Elad, "Improving the K-SVD Facial Image Compression Using a Linear Deblocking Method," in *25th Convention of Electrical and Electronics Engineers*, Israel, 2008, pp. 533-7.
- [7] H. Zheng, Y. Lu, and X. Feng, "Improved Compression Algorithm Based on Region of Interest of Face," in *Proceedings of the 16th International Conference on Artificial Reality and Telexistence*, 2006, pp. 345-8.
- [8] Z. Qiuyu and W. Suozhong, "Color Personal ID Photo Compression Based on Object Segmentation," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, China, 2005.
- [9] D. Taubman, "High Performance Scalable Image Compression with EBCOT," *IEEE Trans. Image Processing*, vol. 9, pp. 1158 –70, 2000.
- [10] J. Hou, W. Chen, and J. Yang, "Formal Photograph Compression in JPEG2000 Coding," in *International Symposium on Intelligence Information Processing and*

Trusted Computing, China, 2010.

- [11] K. Somasundaram and N. Palaniappan, "Adaptive Low Bit Rate Facial Feature Enhanced Residual Image Coding Method Using SPIHT for Compressing Personal ID Images," *International Journal of Electronics and Communications (AEÜ)*, vol. 65, pp. 589-94, 2011.
- [12] A. Said and W.A. Pearlman, "A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Trans Circ Syst Video Technol*, vol. 6, pp. 243-9, 1996.
- [13] A.A. Moinuddin, E. Khan, and M. Ghanbari, "Efficient Algorithm for Very Low Bit Rate Embedded Image Coding," *IET Image Processing*, vol. 2, pp. 59–71, 2008.
- [14] J. H. Kim, J. Y. Lee, E. S. Kang, and S. J. Ko, "Region-Based Wavelet Transform for Image Compression," *IEEE Trans on Circuits and Systems—II: Analog and Digital Signal Processing*, vol. 45, no. 8, pp. 1137-40, 1998.
- [15] S. Mallat, "A Theory of Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11:7, pp. 674-93, 1989.
- [16] M. Ghanbari, *Standard Codecs-Image Compression to Advanced Video Coding*, 3rd ed.: The Institution of Engineering and Technology, 2011.
- [17] J.M. Shapiro, "Embedded Image Coding Using Zero Trees of Wavelet Coefficients," *IEEE Trans Signal Process*, vol. 41, pp. 3445–62, 1993.
- [18] B. E. Usevitch, "A Tutorial on Modern Lossy Wavelet Image Compression: Foundation of JPEG2000," *IEEE Signal Process*, vol. 18:5, pp. 22–35, 2001.
- [19] R. E. Crochiere, S. A. Weber, and J. L. Flanagan, "Digital Coding of Speech in Sub Bands," *Bell Syst. Tech. J.*, vol. 55, pp. 1069–85, 1967.
- [20] I. Daubechies, "The Wavelet Transform, Time Frequency Localization and Signal Analysis," *IEEE Trans. Inf. Theory*, vol. 36:5, pp. 961–1005, 1990.

- [21] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transform.*: Prentice Hall, 1999.
- [22] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, 2nd ed.: Wellesley-Cambridge Press, 1997.
- [23] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding.*: Prentice Hall, 1999.
- [24] O. Egger, P. Fleury, and T. Ebrahimi, "Shape Adaptive Wavelet Transform for Zerotree Coding," *Proc. European Workshop Image Anal. Coding*, vol. 1, pp. 201-8, 1996.
- [25] O. Egger, T. Ebrahimi, and M. Kunt, "Arbitrarily-Shaped Wavelet Packets for Zerotree Coding," *Proc. ICASSP*, vol. IV, pp. 2335-8, 1996.
- [26] T. Aono, N. Ito, H. Katata, and H. Kusao, "Object Based Wavelet Transform Tool for Video," *ISO/IEC JTC1/SC29/WG11, MPEG-4 Proposal*, 1995.
- [27] D. Salomon and G. Motta, *Handbook of Data Compression*, 5th ed.: Springer-Verlag, 2010.
- [28] S. T. Hsiang, Highly Scalable Subband/Wavelet Image and Video Coding, 2002, Ph.D Thesis.
- [29] A. S. Lewis and G. Knowles, "Image Compression Using the 2-D Wavelet Transform," *IEEE Trans. Image Process.*, vol. 1, pp. 244–50, 1992.
- [30] Z. Xiong, K. Ramachandran, and M. T. Orchard, "Space-Frequency Quantization for Wavelet Image Coding," *IEEE Trans. Image Process.*, vol. 6, pp. 677–93, 1997.
- [31] I. H. Witten, R. Neal, and J. G. Cleary, "Arithmetic Coding for Data Compression," *Commun. ACM*, vol. 30, no. 6, pp. 520–40, 1987.
- [32] W. A. Pearlman, A. Islam, N. Nagaraj, and A. Said, "Efficient Low Complexity Image Coding with Set-Partitioning Embedded Block Coder," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, pp. 1219-35, 2004.

- [33] C. Chrysafis, A. Said, A. Drukarev, A. Islam, and W. A. Pearlman, "SBHP- A Low Complexity Wavelet Coder," in *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, vol. 4, 2000, pp. 2035-8.
- [34] S. T. Hsiang and J. W. Woods, "Embedded Image Coding Using Zeroblocks of Subband/Wavelet Coefficients and Context Modeling," in *IEEE Int. Symp. Circuits and Systems*, vol. 3, 2000, pp. 662-5.
- [35] Y. Liang and W. Liu, "A New JPEG2000 Region-of-Interest Coding Method: Generalized Partial Bitplanes Shift," *Proc. of SPIE*, vol. 5637, pp. 365-71, 2005.
- [36] C. A. Christopoulos, J. Askelf, and M. Larsson, "Efficient Methods for Encoding Regions of Interest in the Up-Coming JPEG2000 Still Image Coding Standard," *IEEE Signal Process. Let.*, vol. 7, pp. 247-9, 2000.
- [37] Final Committee Draft, "JPEG2000 Part II," ISO/IEC JTC 1/SC 29/WG 1 (ITU-T SG8), ISO/IEC FCD15444-2, 2000.
- [38] Final Committee Draft, "JPEG2000 Part I," Version 1.0 ISO/IEC JTC 1/SC 29/WG 1 (ITU-T SG8), ISO/IEC FCD15444-1, 2000.
- [39] F. N. Fritsch and R. E. Carlson, "Monotone Piecewise Cubic Interpolation," *Society for Industrial and Applied Mathematics*, vol. 17, no. 2, pp. 238-246, 1980.
- [40] Z. He, "Peak Transform for Efficient Image Representation and Coding," *IEEE Trans on Image Processing*, vol. 16, no. 7, pp. 1741-54, 2007.

Abstract

Today, face identification/verification systems have many applications in areas such as personal identity, traffic control, and information security. These systems usually need a large database of face images. Because of memory limitations and need to have high retrieval speed, we need to use an efficient compression method to compress these face images. In this task, we should preserve important face features especially in low bit rate applications.

In this thesis, in an evolutionary scheme, we propose three ROI-based facial image compression methods for low/very low bit rate applications. All these methods are based on wavelet transform and encoding the coefficients using WBTC coding scheme. We have improved coding performance of WBTC encoder using our proposed block reordering technique which works in coefficients importance order manner. By this technique less time and bit rate is consumed to find important coefficients in encoding process.

In the first method, quality of interested regions is improved by computing the residual image in wavelet domain, and encoding interested regions by Maxshift method. In the second proposed method, by using RWT transform, we compress interested and non-interested regions separately in desired bit rate. Third method is somehow a combined version of previous two proposed methods. This method uses RWT transform and residual image coding techniques and reaches to higher coding performance.

Simulation results show that all proposed methods have higher compression efficiency than JPEG2000, SPIHT, and a ROI-based method especially in low/very low bit rate, since by reducing bit rate, the efficiency and impact of WBTC and block-reordering techniques increase.

Keywords: Wavelet Transform, Facial Image, Image Compression, WBTC Encoder, Region of Interest (ROI).



Shahrood University of Technology
Faculty of Electrical and Robotic Engineering

Facial Images Compression Based on Region of Interest

Mohsen Akrami

Supervisors:

Dr. Hadi Grailu

Dr. Omidreza Maruzi

Advisor:

Dr. Hossein Marvi

FOR THE DEGREE OF MASTER OF SCIENCE

January 2013