

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده برق و رباتیک

گروه الکترونیک

پایان نامه کارشناسی ارشد مهندسی برق-الکترونیک

فشرده سازی تصاویر دندان

هدی شریفی

استاد راهنما :

دکتر امیدرضا معروضی

استاد مشاور:

دکتر هادی گرایلو

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

ماه و سال انتشار: بهمن ۹۲

تقدیم بہ پدر بزرگوار و مادر مہربانم

آن دو فرشتہ ای کہ از خوارتہ ایشان گذشتند، سخی نارابر جان خریدند و خود را سپہر لای مشکلات و ناملایات کردند تا من بہ جایابی کہ اکنون در آن ایستادہ ام برسم.

تقدیر و تشکر

پاس بی کران پروردگار یکتا را که هستی مان نشید و به طریق علم و دانش را بنمونان شد و به بهنشین حر و ان علم و دانش متخرمان نمود و خوشه چینی از علم و معرفت را روزیان ساخت.

از پدر و مادر عزیزم که بهواره برکتی و درستی من، قلم عنو کشیده و گریانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاور بی چشم داشت برای من بوده اند؛ از استاد با کالات و شایسته جناب آقای دکتر امیر رضا مهر و منی که در کمال سع و صدر با حسن خلق و فروتنی، از بیجگی در این عرصه بر من دریغ ننمودند و زحمت را بهمانی این پایان نامه را بر عهده گرفتند و از نزدیک مشاور آقای دکتر سلمان قراگوزلو کمال تشکر و قدردانی را دارم.

چکیده

از تصاویر پزشکی دیجیتال مانند اشعه X، به طور گسترده در تشخیص بیماری‌ها استفاده می‌شود. سهولت در ذخیره‌سازی و انتقال تصاویر پزشکی دیجیتال یکی از مزیت‌های حائز اهمیت این امر برای بیماران می‌باشد. واضح است که هرچه حجم تصاویر کم‌تر باشد، به فشرده‌سازی و انتقال آن‌ها سرعت بخشیده می‌شود و نیازمند فضای کم‌تری برای ذخیره شدن می‌باشد. نکته‌ای که نسبت به سایر جوانب ارجحیت دارد، حفظ کیفیت و ویژگی‌های مهم دندان در حد مناسب مخصوصاً در نرخ بیت‌های پایین است.

در این پایان نامه طی یک سیر تکاملی دو روش برای فشرده‌سازی تصاویر دندان مبتنی بر نواحی مطلوب در نرخ بیت‌های پایین پیشنهاد شده است. تبدیل موجک و کدگذاری SPIHT پایه و اساس هر دو روش پیشنهادی را تشکیل می‌دهند. در هر دو روش، ابتدا باید ناحیه مطلوب مشخص شود، برای این کار سه روش بررسی و در نهایت یکی از آن‌ها با انجام اصلاحات پیشنهادی مورد استفاده قرار گرفت. این سه روش عبارتند از: روش انتشار موج، روش مرز فعال، روش پیوسته کردن مرزها.

با اعمال فیلتر میانه برای حذف نویز و همچنین انجام تغییراتی در روش امتدادیابی در مرزهای آشکار سازی شده توسط عملگر کنی، عملکرد آن در یافتن نواحی پرشدگی دندان ارتقاء داده شد، بطوری‌که روش پیشنهادی قادر به پیوسته کردن مرزها و تعیین نواحی مطلوب با دقت مناسب است. در نهایت روش سوم انتخاب، و فشرده‌سازی بر اساس این روش انجام شد. در روش اول، ناحیه مطلوب و پس زمینه بصورت مجزا کدگذاری شدند ولی در روش ماکزیمم انتقال هر دوی آن‌ها در کنار هم پردازش شدند. میزان فشرده‌سازی در روش ماکزیمم انتقال بیش‌تر است اما خطای حداقل مربعات در ناحیه مطلوب در این روش بیش‌تر از روش اول است. سرعت پردازش در روش SPIHT بیش‌تر است، زیرا در مرحله کدگذاری ماکزیمم انتقال، چند بار رشته بیت تولید می‌شود که این کار باعث کندی نرم افزار در روش ماکزیمم انتقال است.

کلمات کلیدی: تبدیل موجک، تصویر دندان، فشرده‌سازی تصویر، کدگذار SPIHT، ناحیه مطلوب.

لیست مقالات مستخرج از این پایان نامه:

"Specifying Regions of Interest in Dental"; International Research Journal of Applied and Basic Sciences

"Compressing Dental Radiographic Images"; International Research Journal of Applied and Basic Sciences

فهرست مطالب

عنوان.....	صفحه.....
۱- فصل اول.....	۱
۱-۱- تعریف مسئله.....	۲
۱-۱-۱- ضرورت فشرده‌سازی مبتنی بر ناحیه مطلوب.....	۳
۲-۱- مروری بر کارهای گذشته.....	۴
۱-۲-۱- روش‌های مبتنی بر فرهنگ داده.....	۴
۲-۲-۱- روش‌های مبتنی بر تبدیل.....	۷
۳-۱- انتخاب طرح کلی.....	۸
۴-۱- ساختار پایان‌نامه.....	۱۱
۲- فصل دوم.....	۱۳
۱-۲- تبدیل موجک.....	۱۴
۱-۱-۲- کد گذاری زیرباند.....	۱۴
۲-۱-۲- تبدیل موجک.....	۱۶
۳-۱-۲- تبدیل موجک گسسته.....	۱۸
۴-۱-۲- تبدیل موجک مبتنی بر ناحیه.....	۲۰
۱-۴-۱-۲- پیچیدگی محاسباتی تبدیل RWT.....	۲۵
۲-۲- کد گذاری تصویر مبتنی بر موجک.....	۲۶
۱-۲-۲- کد گذاری مبتنی بر درخت صفر.....	۲۷
۲-۲-۲- کد گذاری مبتنی بر بلوک صفر.....	۳۰
۳-۲- کد گذاری ROI در حوزه موجک.....	۳۱
۱-۳-۲- مقیاس گذاری کلی.....	۳۲

۳۳	انتقال ماکزیمم.....۲-۳-۲
۳۶	فصل سوم.....۳-۳
۳۷	روش مورد بررسی اول.....۱-۳-۳
۳۷	مشکلات این روش.....۱-۳-۳
۳۸	روش مورد بررسی دوم.....۲-۳-۳
۴۱	روش پیشنهادی.....۳-۳-۳
۴۳	الگوریتم اتصال.....۱-۳-۳
۴۴	چگونه نواحی پرشدگی انتخاب شود؟.....۴-۳-۳
۴۵	نتایج تجربی برای تصاویر دندان.....۵-۳-۳
۵۰	فصل چهارم.....۴-۴
۵۰	ضرورت فشرده‌سازی.....۱-۴-۴
۵۰	تعریف فشرده‌سازی و انواع آن.....۲-۴-۴
۵۲	روش اول.....۳-۴-۴
۵۳	روش دوم (ماکزیمم شیفت).....۲-۴-۴
۵۸	فصل پنجم.....۵-۴
۵۸	نتیجه‌گیری.....۵۸-۴
۵۸	معیارهای ارزیابی.....۵۸-۴
۵۹	فشرده‌سازی.....۱-۵-۴
۶۴	نحوه بدست آوردن ROI به کمک مدل مار.....۲-۵-۴
۶۶	پیوسته کردن مرزهای گسسته.....۳-۵-۴
۶۷	نحوه بدست آوردن ROI به کمک پیوسته کردن مرز.....۴-۵-۴
۷۰	پیشنهادات برای کارهای آینده.....۵-۵-۴

فهرست شکل‌ها

- شکل ۱-۲: بانکی از فیلترهای میان گذر ۱۵
- شکل ۲-۲: گام‌های پردازشی تبدیل ۲۳
- شکل ۳-۲: تصویر فروشنده ۲۴
- شکل ۴-۲: درخت‌های جهت دار مکانی و بخش‌بندی مجموعه در SPIHT ۲۹
- شکل ۵-۲: نمایش بخش‌بندی درخت چهارگانه و کد گذاری در یک ناحیه بلوکی $16*16$ ۳۱
- شکل ۶-۲: روش‌های کد گذاری ROI در استاندارد JPEG2000 ۳۳
- شکل ۱-۳: الف) تصویر ورودی ب) تصویر خروجی از مدل مار ۴۰
- شکل ۲-۳: فلوجارت الگوریتم پیشنهادی (روش سوم) ۴۱
- شکل ۳-۳: نحوه اتصال یک مرز به مرز دیگر ۴۴
- شکل ۴-۳: الف) عکس اصلی ب) فیلتر میانه $15*15$ ۴۶
- شکل ۵-۳: الف) عکس اصلی ب) فیلتر میانه $15*15$ ۴۷
- شکل ۱-۴: الف) تصویر ورودی ب) ناحیه پس‌زمینه پ) ناحیه ROI ت) تصویر بازسازی SPIHT ۵۳
- شکل ۲-۴: الف) تصویر ورودی ب) ناحیه پس‌زمینه پ) ناحیه ROI ت) تصویر بازسازی Maxshift ۵۵
- شکل ۱-۵: الف) تصویر ورودی ب) ناحیه ROI ۶۰
- شکل ۲-۵: الف) تصویر ورودی ب) ناحیه ROI ۶۲
- شکل ۳-۵: الف) تصویر ورودی ب) ناحیه انتخاب شده توسط کاربر ۶۵
- شکل ۴-۵: الف) مرز گسسته ب) مرز پیوسته ۶۶
- شکل ۵-۵: الف) عکس اصلی ب) فیلتر میانه $15*15$ ۶۸
- شکل ۶-۵: مشخص شدن ناحیه مطلوب ۶۹

فهرست جداول

جدول ۴-۱: نتایج روش اول ۵۳

جدول ۵-۱: نتایج حاصل از مقایسه دو روش ۶۱

جدول ۵-۲: نتایج مقایسه دو روش فشرده‌سازی ۶۳

فصل اول

مقدمه

فصل اول

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 فشردن سازی نقش مثبت خود را نشان می‌دهد. البته شاید برخی بگویند که، درست است که امروزه حجم داده‌ها روز به روز افزایش می‌یابد، اما از آن طرف با پیشرفت تکنولوژی، ظرفیت حافظه‌ها نیز روز به روز در حال افزایش است. در پاسخ به این افراد می‌توان سه دلیل برای ضرورت فشردن سازی 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 فشردن سازی را بطور کلی می‌توان به دو دسته فشردن سازی بی‌اتلاف^۱ و فشردن سازی با-

اتلاف^۲ تقسیم‌بندی کرد. نرخ فشردن سازی^۳ نوع با اتلاف بسیار بیش‌تر از نوع بی‌اتلاف است، اما این

مزیت فشردن سازی با اتلاف به قیمت از دست دادن بخشی از داده‌ها تمام می‌شود. در حوزه داده‌های

تصویری همانند سایر حوزه‌های داده تاکنون روش‌های زیادی برای فشردن سازی ارائه شده‌است. 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 روش-

1 Lossless Compression

2 Lossless Compression

3 Compression Rate

هایی که تاکنون برای فشرده‌سازی تصاویر ارائه شده‌اند، همه‌منظوره 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 مزایای 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

همانطور که قبلاً گفته شد، اکثر روش‌های فشرده‌سازی متداول، کیفیت تصاویر را بطور سراسری کاهش می‌دهند، چرا که عمل فشرده‌سازی در این روش‌ها در سرتاسر تصویر با یک نرخ بیت یکسان صورت می‌گیرد. بنابراین استفاده از روش‌های متداول برای فشرده‌سازی تصاویر دندان مخصوصاً در نرخ بیت‌های پایین، نتیجه‌ای جز کاهش شدید کیفیت نواحی و ویژگی‌های مهم دندان و در نهایت کاهش دقت نخواهند داشت. برای رفع این مشکل می‌توان به جای فشرده‌سازی سرتاسر تصویر در یک نرخ بیت یکسان، به نواحی مهم نسبت به سایر نواحی اهمیت و نرخ بیت بیشتری اختصاص داد. این نوع فشرده‌سازی را مبتنی بر ناحیه مطلوب (ROI)^۲ می‌نامند.

1 Redundancy

2 Region Of Interest

۱-۲ مروری بر کارهای گذشته

تاکنون اکثر مطالعات و پژوهش‌ها حاوی مطالبی پیرامون فشردن سازی تصاویر چهره بوده است. در این پایان نامه سعی بر آن بوده که این پژوهش‌ها و تمرکز موضوعی در مورد تصاویر دندان اعمال شود. به دلیل وجود شباهت در زمینه ROI و عدم وجود مقالات جدید در مورد فشردن سازی دندان قسمت مروری بر کارهای گذشته با نگاهی به مقاله‌های استفاده شده در این پایان نامه [۲۴] به صورت زیر آورده شده است.

بطور کلی راهکارهای ارائه شده برای فشردن سازی مبتنی بر ROI را می‌توان به دو دسته تقسیم بندی کرد: فشردن سازی مبتنی بر فرهنگ داده^۱ یا کتاب رمز^۲ و فشردن سازی مبتنی بر تبدیل^۳. البته این دسته بندی خیلی کلی بوده و در مورد سایر حوزه‌های فشردن سازی 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 بر اساس 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 الگوریتم‌های خوشه بندی^۶ 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^۷ می‌نامند. در حالتی که از الگوریتم‌های خوشه بندی در مرحله آموزش استفاده شود، به ازای یک موقعیت بلوک

¹ Dictionary-Based Compression

² Code Book

³ Transform-Based Compression

⁴ Training Phase

⁵ Encoding

⁶ Clustering

⁷ Code Word

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 بردار است که بطور میانگین بیش‌ترین شباهت را به تمامی اعضای خوشه داشته باشد. این میزان شباهت معمولاً توسط یک معیار فاصله سنجیده می‌شود. در مرحله آزمایش^۱ با سنجش میزان شباهت هر بلوک با تک تک اعضای فرهنگ داده مربوطه و سپس ارسال اندیس^۲ متناظر با شبیه‌ترین عضو فرهنگ داده به واحد کدگشا^۳ یا ذخیره آن در حافظه عمل کدگذاری صورت می‌گیرد. در صورت استفاده از تبدیلات ماتریسی در مرحله آموزش، به ازای یک موقعیت بلوک مشخص، فرهنگ داده مربوطه مجموعه‌ای از بردارهای متعامد^۴ یک خواهد بود که یک سیستم پایه یک‌معامد را برای نمایش بردارهای متناظر با این موقعیت بلوکی در مجموعه آموزش تشکیل می‌دهند. در این حالت در فاز آموزش، هر بلوک بصورت یک ترتیب خطی از اعضای فرهنگ داده مربوطه نشان داده شده، و سپس با ارسال ضرایب چندی‌سازی شده^۵ این ترکیب خطی به واحد کدگشا یا ذخیره آن‌ها در حافظه عمل کدگذاری صورت می‌گیرد. البته ممکن است به منظور کاهش تزیاید اندیس‌های حاصل از حالت اول و یا ضرایب ترکیب‌های خطی حاصل از حالت دوم، یک مرحله اضافی به نام کدگذاری آنتروپی نیز بر روی این داده‌ها صورت گیرد تا در نهایت بتوان به نرخ فشرده-سازی 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،

¹ Testing Phase

² Index

³ Decoder

⁴ Orthonormal Vectors

⁵ Quantized Coefficients

بطوریکه به بلوک‌های مهم‌تر فرهنگ داده‌های بزرگ‌تر و در نتیجه نرخ بیش‌تری اختصاص داده می‌شود. در [۳] با تغییر شکل هندسی تصاویر به یک فرم معیار، تقسیم تصویر حاصل به بلوک‌های مجزا و در نهایت کدگذاری هر بلوک به روش چندی‌سازی برداری (VQ)^۱ عکس فشرده‌سازی صورت می‌گیرد. تبدیل تصاویر به فرم معیار و نیز استفاده از VQ این روش را قادر به بهره‌گیری از مزایای موجود در ویژگی‌های چهره یکسان افراد مختلف، و در نتیجه افزایش کارایی فشرده‌سازی مخصوصاً در نرخ‌های پایین ساخته است. از معایب این روش می‌توان به زمان طولانی مورد نیاز برای آموزش فرهنگ داده‌های VQ و نیز حافظه جانبی نسبتاً زیاد ۴۰ مگابایتی برای ذخیره‌سازی این فرهنگ داده‌ها در هر دو واحد کدگذار و کدگشا اشاره کرد. در [۴] یک روش مبتنی بر فرهنگ داده برای فشرده‌سازی تصاویر چهره ارائه شده است که کارایی خوبی مخصوصاً در نرخ بیت‌های پایین دارد. در این روش ابتدا مشابه [۳] شکل هندسی تصاویر به یک فرم معیار تغییر داده شده، و سپس بلوک‌های تصویر حاصل بصورت مجزا با استفاده از نمایش پراکنده کدگذاری می‌شوند. نمایش پراکنده نیازمند فرهنگ داده‌های فوق کامل^۲ است، که این فرهنگ داده‌ها در مرحله آموزش با استفاده از الگوریتم تجزیه برداری منفرد (SVD)^۳ تولید می‌شوند. استفاده از نمایش پراکنده باعث افزایش کارایی فشرده‌سازی این روش نسبت به [۳] شده است، اما در عین حال همچنان همان معایب روش [۳] را با خود به یدک می‌کشد. این روش در ادامه با حذف عناصر مزاحم بصری، با استفاده از یک تکنیک بلوک شکنی در [۵] توسعه داده شد. این تکنیک بلوک شکنی مبتنی بر فیلترهای محلی بوده و به عنوان مرحله پس‌پردازش انجام می‌گیرد. با استفاده از این تکنیک یکی از معایب روش [۴] تقریباً بر طرف شده و کارایی آن تا حدودی افزایش یافته است.

¹ Vector Quantization

² Overcomplete

³ Singular Value Decomposition

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 تزیاید و هم بستگی^۱ میان داده‌های تصویر در نمایش جدید حذف شده و یا کاهش یابد. این هدف عموماً با متراکم سازی 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 بیش‌تر را فراهم سازد. یک تبدیل ممکن است داده‌های تصویر را در حوزه فرکانس به نمایش گذارد مانند تبدیل کسینوسی (DCT)^۲ و تبدیل والش-هادامارد^۳، یا داده‌های تصویر را به فضای مکانی دیگر منتقل کرده و توصیف نماید مانند تحلیل مولفه اصلی (PCA)^۴.

علاوه بر این تبدیلاتی وجود دارند که قادرند اطلاعات تصویر را هم در حوزه مکانی 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 قابلیت خود را به رخ کشیده‌است. اکثر روش‌های مبتنی بر تبدیل ارائه شده در زمینه فشرده‌سازی تصاویر مختلف مانند 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 نرخ بیت‌های پایین، از تکنیک کدگذاری ROI در

¹ Correlation

² Discrete Wavelet Transform

³ Walsh-Hadamard Transform

⁴ Principal Component Analysis

⁵ Discrete Wavelet Transform

⁶ Local and Multiscale Analysis

حوزه موجک استفاده شده است. در ادامه به مروری بر چند روش ارائه شده در این دسته می‌پردازیم. در [۶-۸] یک روش فشرده‌سازی تصاویر مبتنی بر ناحیه مطلوب ارائه شده است که از تبدیل DWT بهره می‌گیرد. از طرفی تعیین دقیق ماسک ضرایب موجک ROI یکی از مشکلات این روش به حساب می‌آید. در این روش از الگوریتم بخش‌بندی مجموعه درخت‌های سلسله (SPIHT) [۹]، 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.

۱-۳ انتخاب طرح کلی

در بخش قبل به بررسی انواع روش‌های موجود برای فشرده‌سازی مبتنی بر ROI تصاویر پرداختیم و مشاهده کردیم که هر دو دسته روش‌های مبتنی بر فرهنگ داده و مبتنی بر تبدیل دارای مزایا و معایبی هستند. در حقیقت این روش‌ها قادرند با بهره‌گیری از شباهت موجود میان تصاویر در موقعیت‌های مکانی مشابه، تزاید اطلاعات میان تصاویر مجموعه‌ای بزرگ را به مقدار زیاد کاهش دهند. البته مسائل چالش برانگیزی همچون انتخاب الگوریتم مناسب برای آموزش فرهنگ داده‌ها و نیز تعیین اندازه مناسب بلوک‌ها، معمولاً طراحی این روش‌ها را مشکل و پیچیده می‌سازد. 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 نرخ بیت‌های مختلف و به تبع آن بازسازی تصاویر با کیفیت‌های مختلف به سادگی امکان پذیر نیست و کاری بس زمان‌بر و هزینه‌بر است، چرا که نیازمند محاسبه دقیق نرخ بیت مورد نیاز هریک از بلوک‌ها و سپس تکرار مجدد فاز آموزش برای دستیابی به فرهنگ داده‌های متناسب با نرخ بیت و دقت مورد نیاز است. می‌توان گفت روش‌های مبتنی بر فرهنگ داده روش‌های تطبیقی هستند، بدین معنی که در فاز

¹ Set Partitioning In Hierarchical Tree

آموزش با فراگیری دانش و اطلاعاتی راجع به یک جامعه آماری محدود از تصاویر 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 تطبیق یافته‌اند و روی تصاویر خارج از این مجموعه کارایی مناسب را نخواهند داشت. 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 بر طرف نموده است. از طرفی ممکن است در نرخ بیت‌های 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 نباشند، و بتوان با استفاده از آن‌ها تصاویر گوناگون را با کارایی مناسب فشرده و بازسازی کرد. همانطور که در بخش ۲-۱-۲ گفته شد، در میان تبدیلات مورد استفاده در این روش‌ها، استفاده از تبدیل DWT به دلیل ویژگی‌های کم- نظیر 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 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 کرده‌اند که تنها با در اختیار داشتن یک دنباله بیت بتوان تصویر را در مقیاس‌های مختلف بازسازی کرد. چنین الگوریتم‌هایی را اصطلاحاً الگوریتم‌های کدگذاری جاسازی شده¹ 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 چالش برانگیز است، عدم امکان تعیین دقیق ماسک ضرایب موجک ROI، و در نتیجه عدم امکان تخصیص نرخ‌بیتی دقیق به هریک از نواحی مطلوب و غیرمطلوب 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

¹ Embedded Coding

دانست، چرا که تبدیل DWT تنها متشکل از یک سری عملیات ضرب و جمع مربوط به فیلترها بوده و الگوریتم‌های کدگذاری ضرایب موجک نیز معمولا تنها متشکل از یک سری عملیات مقایسه‌ای و تصمیم‌گیری‌های منطقی هستند.

۴-۱ ساختار پایان‌نامه

در این پایان‌نامه بطور کلی هدف آن است که با ارائه روش یا روش‌هایی کارآمد برای فشرده‌سازی تصاویر دندان 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) بررسی روش‌های موجود در زمینه فشرده‌سازی تصاویر، دسته‌بندی روش‌ها و انتخاب طرح کلی فشرده‌سازی به عنوان پایه و مبنای روش‌های پیشنهادی.

(۲) انتخاب روش‌های مناسب و کارآمد برای تحقق هر یک از اجزا و بخش‌های طرح کلی و در صورت نیاز اصلاح طرح کلی برای تطبیق با مسئله فشرده‌سازی مبتنی بر ناحیه مطلوب تصاویر دندان.

این پایان‌نامه متشکل از چند بخش است. در بخش ۱-۲ یک دسته‌بندی کلی برای روش‌های موجود در زمینه فشرده‌سازی تصاویر گوناگون ارائه شده است، که قابل تعمیم به سایر حوزه‌های فشرده‌سازی تصویر نیز هست. سپس براساس این دسته‌بندی به بررسی مزایا و معایب برخی از روش‌های فشرده‌سازی تصاویر گوناگون پرداخته‌ایم. 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 فشرده‌سازی و پایه و مبنای روش‌های پیشنهادی برگزیدیم. در روش‌های مبتنی بر تبدیل عموماً اولین گزینه‌ای که برای انتخاب تبدیل به ذهن‌خطور می‌کند، تبدیل DWT است که معمولا 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 تبدیل یافته‌است. 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. الگوریتم‌های متنوعی برای کدگذاری ضرایب در حوزه موجک معرفی شده است. این الگوریتم‌ها را می‌توان به طور 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 آن -ها کدگذاری و ارسال بیش‌ترین ضرایب موجک مهم با کم‌ترین هزینه نرخ‌بیت 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)^۳، سعی در بهره‌گیری از هم‌بستگی بین باندهای^۴ ضرایب موجک را دارند. از معروف‌ترین الگوریتم‌های این دسته می‌توان به SPIHT اشاره کرد که در این پایان‌نامه از این کدگذار استفاده شده است. در فصل سوم به تشریح روش‌های پیشنهادی برای بدست آوردن ناحیه مطلوب خواهیم پرداخت. در فصل چهارم به تشریح روش‌های پیشنهادی در این پایان‌نامه برای بدست آوردن فشرده‌سازی مبتنی بر ناحیه مطلوب خواهیم پرداخت. در فصل پنجم نتایج شبیه‌سازی‌های انجام شده به نمایش گذاشته می‌شوند و مورد تحلیل و بررسی قرار می‌گیرند و در انتهای این فصل به نتیجه‌گیری و ارائه پیشنهادات خواهیم پرداخت.

¹ Zerotree-Based

² Zeroblock-Based

³ Spatial Orientation Tree

⁴ Inter-Subband Correlation

فصل دوم

مغایم پاپه

۲- فصل دوم

۱-۲ 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 مالت^۱ [۱۰] شروع شد. او با ارائه دیدگاهی جدید از تبدیل موجک، آن را بعنوان 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 فشرده سازی تصویر تا قبل از معرفی استاندارد JPEG بسیار محدود بود که علت آن را می‌توان نبود روشی کارآمد برای کدگذاری ضرایب تبدیل موجک دانست [۱۱]. با معرفی روش موجک درخت صفر جاسازی شده (EZW)^۳ توسط جوزف شاپیرو^۴ در [۱۲]، انقلابی در زمینه کدگذاری ضرایب موجک و فشرده‌سازی تصویر مبتنی بر تبدیل موجک بوجود آمد، به گونه ای 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 شده اند و این روند همچنان ادامه دارد.

۱-۱-۲ کد گذاری زیرباند

قبل از پرداختن به تبدیل موجک، بهتر است ابتدا با کدگذاری زیرباند^۵ که به نوعی از اجداد تبدیل موجک به شمار رفته و گاهی به آن تبدیل موجک اولیه نیز گفته می‌شود به اختصار آشنا شویم [۱۳]. 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

¹ Mallet

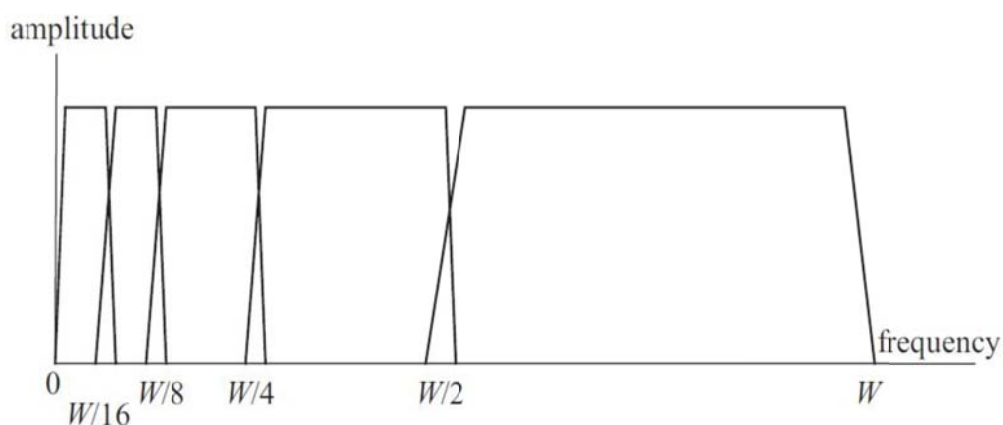
² Multiresolution Analysis

³ Embedded Zerotree Wavelet

⁴ Jussef Shapiro

⁵ Subband Coding

زیرباند، بخش‌بندی طیف سیگنال به چندین محدوده یا باند فرکانسی و سپس کد گذاری و انتقال هر باند به طور مجزا می‌باشد [۱۴]. 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 فیلتر بانک 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 اکتا^۱ (یعنی افزایش / کاهش عرض باندها به صورت توان صحیحی از ۲) طراحی شده‌اند. از آنجا که پهنای باند هر نسخه تصویر فیلتر شده کاهش یافته است، از نظر تئوری مطابق قضیه نایکوئیست می‌توان با نرخ کمتری از این تصاویر نمونه‌برداری کرد، و در نتیجه تعدادی زیر تصویر کاهش بعد یافته بدست آورد. در واحد کدگذار این زیر تصاویر پس از چندی‌سازی کدگذاری شده و ارسال می‌شوند. در واحد کدگشا، زیر تصویرهای دریافتی به ابعاد اولیه بازگردانده شده و از مجموعه‌ای از فیلترها به نام فیلتر بانک ترکیب^۲ عبور داده می‌شوند. این فیلتر بانک عمل درونیابی و ترکیب زیر تصاویرها به منظور بازسازی تصویر اولیه را انجام می‌دهد.



شکل ۱-۲: بانکی از فیلترهای میان‌گذر

¹ Octave Band

² Synthesis filter bank

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 استوار است. با این حال از آنجا که تصاویر طبیعی به طور محلی هموار بوده و می توان آن ها را با تقریب تکه ای-چندجمله ای^۱ مدل کرد، در صورت انتخاب مناسب تابع چند جمله ای می توان مشابه با ایده کدگذاری زیرباند به نوعی تحلیل فرکانسی دست یافت. از نظر ریاضی، چنانچه $X(t)$ تابعی مربعی-انتگرالپذیر^۲، تبدیل موجک آن معادل با تجزیه این تابع برحسب مجموعه ای از توابع پایه به صورت زیر است:

$$X_w(a,b) = \int_{-\infty}^{+\infty} x(t)\varphi_{a,b}(t)dt \quad (1-2)$$

که در آن $\varphi_{a,b}(t)$ به نام تابع پایه شناخته می شود. این تابع نسخه ای گسترش یافته و انتقال یافته از یک سیگنال میان گذر مانند $\varphi(t)$ است که موجک مادر^۳ نامیده شده و به صورت زیر تعریف می شود:

$$\varphi_{a,b} = \frac{1}{\sqrt{a}} \varphi\left(\frac{t-b}{a}\right) \quad (2-2)$$

که در آن a و b به ترتیب پارامترهای گسترش و انتقال نامیده می شوند. عرض تابع پایه با تغییر فاکتور گسترش (یا مقیاس) a به این شکل تغییر می کند که هر چه این فاکتور بزرگ تر شود، عرض زمانی تابع پایه نیز بزرگ تر و بنابراین عرض فرکانسی آن کم تر می شود. بنابراین به کمک این پارامتر می توانیم درجه تفکیک زمانی و فرکانسی را در تبدیل موجک تغییر دهیم، و همین ویژگی جالب

¹ Piecewise Polynomial

² Square-Integrable

³ Mother Wavelet

تبدیل موجک است که این تبدیل را برای تحلیل سیگنال‌هایی همچون تصاویر طبیعی که ویژگی-هایی با اندازه‌های مختلف دارند، مناسب می‌سازد.

متناظر با هر اندازه‌ی ویژگی^۱ یک تابع پایه 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 پایه 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.

۳-۱-۲ تبدیل موجک گسسته

تبدیل موجک تعریف شده در رابطه‌ی (۱-۲) سیگنال یک بعدی $x(t)$ را به تابع دو بعدی $X_w(a,b)$ نگاشت می‌کند و بنابراین موجب تولید تزايد زیادی می‌شود. سیگنال اولیه را می‌توان از روی تبدیل موجک و به ازای مقادیر گسسته‌ای از پارامترهای a و b بدست آورد [۱۵]. پارامتر a را می‌توان با انتخاب $a = a_0^m$ ، بطوریکه $a_0 > 1$ ، m عددی صحیح باشد، گسسته کرد. با افزایش مقدار a ، پهنای باند تابع پایه کاهش می‌یابد و بنابراین برای پوشش یک ناحیه فرکانسی مشخص، نیاز به گام‌ها یا در اصطلاح سلول‌های درجه تفکیک^۳ بیش‌تری است. بطور مشابه، گسسته سازی 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

¹ Feature Size

² Quality Factor

³ Resolution Cell

با انتخاب $b=nb_0a_0^m$ گسسته کرد. به ازای $a_0=2$ و $b_0=1$ انتخاب‌هایی برای تابع $\varphi(t)$ وجود دارند، بطوریکه توابع $\varphi_{m,n}(t)$ حاصل از آن‌ها پایه‌های یکامتعامدی را در فضای توابع مربعی-انتگرال پذیر تشکیل می‌دهند. این امر بدین معناست که هر تابع مربعی-انتگرال پذیر $x(t)$ را می‌توان بصورت یک ترکیب خطی از توابع پایه همانند زیر نمایش داد:

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

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

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

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 باند فرکانسی و نیز تبدیل موجک توان-دو^۱ است. در تبدیل موجک توان-دو مشابه کد گذاری زیر باند، برای تحلیل سیگنال از یک تکنیک مبتنی بر فیلتر بانک موجک که به الگوریتم هرمی نیز معروف است، استفاده شده است [۱۶]. یک فیلتر بانک از فیلترهایی تشکیل شده است که یک سیگنال را به باندهای فرکانسی مختلف تجزیه می‌کند [۱۷]. در الگوریتم هرمی، سیگنال اصلی توسط دو فیلتر موجک پایین‌گذر L و بالاگذر H به دو سیگنال فرکانس پایین و 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

¹ Dyadic Wavelet Transform

² Downsampling

³ Decomposition Level

اول، مجدداً با عبور از فیلترهای L و H و زیر نمونه‌برداری 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 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 مورد سیگنال‌های دوبعدی مانند تصاویر، چنانچه سطرها و ستون‌ها به عنوان سیگنال‌هایی یک بعدی در نظر گرفته شوند، به روشی مشابه فوق می‌توان تبدیل موجک گسسته دو بعدی را بدست آورد [۱۸]. در این روش ابتدا در جهت افقی از تک تک سطرهای تصویر توسط دو فیلتر L و H تبدیل موجک گسسته یک بعدی گرفته می‌شود. در نتیجه این عمل تصویر اصلی به دو زیر تصویر مستطیلی هم اندازه با محتوای فرکانسی متفاوت تجزیه می‌شود. تعداد سطرهای این زیر تصاویر با تصویر اصلی برابر، و تعداد ستون‌های آن‌ها به دلیل زیر نمونه‌برداری نصف تصویر اصلی است. حال این بار در جهت عمودی مجدداً با اعمال فیلترهای L و H روی تک تک ستون‌های این زیر تصاویر، از آن‌ها تبدیل موجک گسسته یک بعدی گرفته می‌شود. با این کار در نهایت چهار زیر تصویر حاصل خواهد شد. زیر تصویری که در نتیجه دوبار اعمال متوالی فیلتر L در جهت‌های افقی و عمودی روی تصویر اصلی حاصل شده است، اصطلاحاً زیر تصویر تقریب نامیده شده و با LL نشان داده می‌شود. به همین ترتیب زیر تصاویر LH، HL و HH به ترتیب زیر تصاویر جزئیات افقی، عمودی و قطری نامیده می‌شوند. اکنون یک تجزیه موجک یک سطحی از تصویر اصلی بدست آمده است. برای دستیابی به سطح دوم تجزیه موجک کفایت 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.

۲-۱-۴ تبدیل موجک مبتنی بر ناحیه

تبدیل موجک سیگنال داده شده را به زیر باندهای مختلف فرکانسی تجزیه می‌کند و امکان بهره‌مندی از قابلیت‌های تحلیل چند مقیاسی 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 متراکم سازی انرژی در زیر باندهای فرکانسی پایین دانست. در کنار مزایای فراوان تبدیل موجک، عدم امکان اعمال آن بر روی نواحی و یا اشیا 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 پیشنهاد شده- اند [۲۰-۲۲]. تبدیل موجک و فقی شکل (SWAT)^۱، برای اعمال تبدیل موجک به هر ناحیه با شکل دلخواهی پیشنهاد شده است [۲۰] و [۲۱]. 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 بازسازی شده 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 شی^۳ با تبدیل شی با شکل دلخواه به یک ناحیه مستطیلی، تبدیل موجک معمولی بر روی آن اعمال می‌شود [۲۲]. این روش از نظر محاسباتی کارآمد نیست، چرا که فرایند تجزیه در این روش نیازمند عملیات برون‌یابی^۴ و فیلترگذاری پایین‌گذر است. در [۱۹] روشی برای اعمال تبدیل موجک بر روی نواحی با شکل دلخواه پیشنهاد شده است، که علاوه بر

¹ Shape Adaptive Wavelet Transform

² Break-Ups

³ Object-based Wavelet Transform

⁴ Extrapolation

پیچدگی محاسباتی پایین، عملکرد بهتری نسبت به روش‌های پیشین از خود نشان داده شده است. این روش اصطلاحاً تبدیل موجک مبتنی بر ناحیه یا به اختصار RWT نامیده شده است.

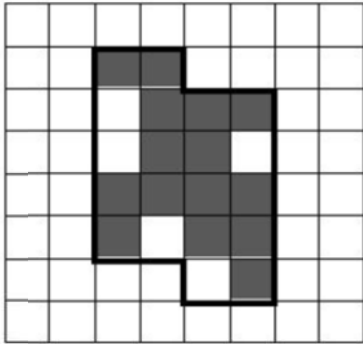
تبدیل RWT قابل اعمال بر روی هر ناحیه و یا شیء درون تصویر با شکل دلخواه است. در تبدیل RWT، عمل تجزیه به جای اعمال فیلترهای موجک به تصویر مستطیلی، با اعمال این فیلترها به ناحیه شیء توسعه یافته صورت می‌گیرد. در تبدیل RWT با حفظ شکل اصلی شیء در هر زیرباند امکان دستیابی به ماکزیمم ناهمبستگی^۱ نواحی تصویر فراهم است. از طرفی تبدیل معکوس RWT می‌تواند شکل اصلی ناحیه مورد نظر را بطور کامل و بی‌عیب بازسازی می‌کند. 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، بطور قابل ملاحظه‌ای نیازمند محاسبات کم‌تری است.

در شکل ۲-۲ الگوریتم تبدیل RWT نشان داده شده است. همانطور که مشاهده می‌شود در اینجا نیز همچون روش‌های تجزیه متداول، روشی تفکیک پذیر 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 توسعه یافته لحاظ می‌شوند. نواحی شیء اصلی و توسعه یافته به ترتیب در شکل ۲-۲ (الف) و شکل ۲-۲ (ب) نشان داده شده‌اند. پس از تعیین ناحیه شیء توسعه یافته، فیلتر بانک موجک شامل فیلترهای پایین‌گذر و بالاگذر در جهت افقی به این ناحیه اعمال می‌شوند. نمونه‌های فیلتر شده ناحیه شیء توسعه یافته طوری درون یک زیرباند قرار می‌گیرند که شکل اصلی شیء در زیرباند حفظ شود. شکل ۲-۲ (ج) نتیجه حاصل از تجزیه افقی را نشان می‌دهد. به روشی مشابه در جهت عمودی از بالا به پایین، تجزیه عمودی بر روی تصویر تجزیه شده افقی اجرا می‌شود. نتیجه حاصل از تجزیه با

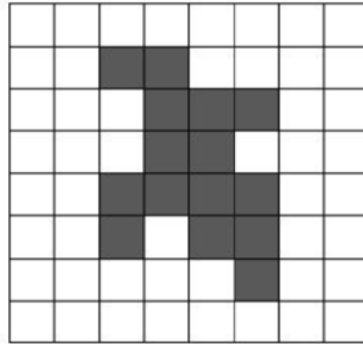
^۱ Decorrelation

چهار زیر باند LL، LH، HL و HH در شکل ۲-۲ (د) نشان داده شده است. باید توجه داشت که تبدیل RWT شکل اصلی شی تصویر را در هر زیرباند حفظ می‌کند، که ویژگی مهمی در تحلیل چند مقیاسی و کدگذاری کارآمد محسوب می‌شود، چرا که این روش مانع از ایجاد برخی اعوجاج‌ها میان پیکسل‌های تصویر شده و خود تشابهی شی را در هر زیرباند حفظ می‌کند.

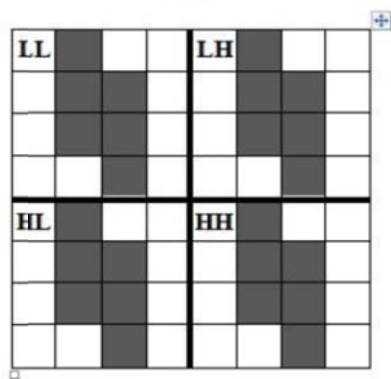
برای اعمال تبدیل RWT به تصویر حقیقی، تصویر قطعه بندی شده مورد نیاز است. شکل ۲-۳ (الف) تصویر "فروشنده" را نشان می‌دهد. این تصویر بصورت دستی به ناحیه شی و پس زمینه قطعه بندی شده، که در شکل ۲-۳ (ب) نشان داده شده است. "شی" شامل سر شخص و بخش بالای بدن اوست، و پس زمینه با مقدار صفر پر شده است تا لبه‌ها و مرز بین شی و پس زمینه بطور کامل قابل تشخیص باشند. شکل ۲-۳ (ج) تصویر تجزیه شده حاصل از تبدیل RWT را نشان می‌دهد. به منظور مقایسه، تصویر تجزیه شده حاصل از تبدیل SWAT نیز در شکل ۲-۳ (د) نشان داده شده است. مشاهده می‌شود که تبدیل RWT علاوه بر حفظ شکل اصلی شی تصویر، مقدار اطلاعات فرکانس بالای کمتری نسبت به تبدیل SWAT تولید کرده است. تصویر بازسازی شده با استفاده از تبدیل معکوس RWT در شکل ۲-۳ (ه) نشان داده شده است. میانگین مربعات خطا (MSE) میان شی اصلی و شی بازسازی شده برابر صفر است. بنابراین تبدیل معکوس RWT می‌تواند بطور کامل شی اصلی را بازسازی کند.



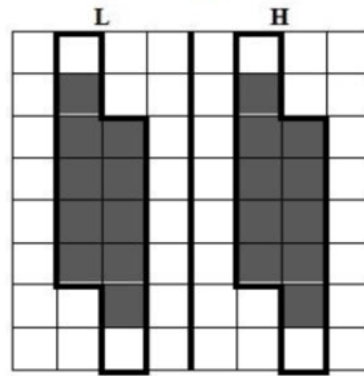
(ب)



(الف)



(د)



(ج)

شکل ۲-۲: گام‌های پردازشی تبدیل RWT (الف) شی اصلی. (ب) شی توسعه یافته. (ج) نتیجه تجزیه افقی. (د) نتیجه تجزیه عمودی اعمال شده روی نتیجه تجزیه افقی [۱۹].



(ب)



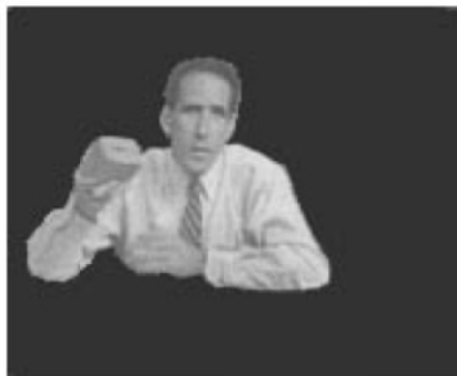
(الف)



(د)



(ج)



(ه)

شکل ۲-۳: 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 SWAT. (ه) تصویر بازسازی شده از معکوس RWT.

۲-۱-۴-۱ پیچیدگی محاسباتی تبدیل RWT

تبدیل RWT می‌تواند پیچیدگی محاسباتی تبدیل موجک را بطور قابل ملاحظه‌ای کاهش دهد. به منظور مقایسه پیچیدگی تبدیل RWT با تبدیل موجک متداول، تعداد عملیات مورد نیاز برای انجام این دو تبدیل ارزیابی می‌شوند. فرض کنید ابعاد تصویر $N * N$ 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 فیلتر بانک شامل فیلترهای پایین‌گذر و بالاگذر، 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 می‌تواند بطور قابل ملاحظه‌ای تعداد عملیات ضرب و جمع را کاهش دهد، اگرچه عمل تجزیه توسط RWT نیازمند $3/2 N^2$ عملیات اضافی مقایسه به منظور تعیین ناحیه شی در تصویر است (N^2 عملیات مقایسه برای تجزیه افقی و $1/2 N^2$ عملیات مقایسه برای تجزیه عمودی). با فرض اینکه ناحیه شی $1/K$ از تصویر را اشغال کرده باشد، تعداد عملیات ضرب و جمع مورد نیاز در RWT به $1/K$ تعداد این عملیات در تبدیل موجک متداول کاهش می‌یابد. بنابراین در تبدیل RWT تجزیه سطح اول تصویر می‌تواند با تعداد $2.P.N^2.1/K$ عملیات ضرب، $2.(P-1).N^2.1/K$ عملیات جمع، و $3/2 N^2$ عملیات مقایسه انجام شود.

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 باطری محدودیت دارند، نیازمند الگوریتم‌هایی برای کدگذاری تصویر 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-عوجاج نیز کارایی بالایی داشته باشند. چنین الگوریتم‌هایی باید قابلیت ساخت و بازسازی پیشرو^۱ تصاویر کد شده 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 می‌گیرد، چرا که در این حالت محدودیت ظرفیت کانال، محدودیت حافظه و نیز محدودیت در پیچیدگی محاسباتی عواملی قطعی خواهند بود. تبدیل موجک به نحو بسیار خوبی قادر به خوشه‌بندی انرژی در مکان و فرکانس است. 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 مهم با کم‌ترین هزینه نرخ بیت و زمان است. دستیابی به این هدف مستلزم ارائه روشی بهینه چه از نظر نرخ بیت 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

^۱ Progressive

^۲ Bit Plane-Based

گرفته می‌شود، چرا که ثابت شده است ضرایب موجک بزرگ‌تر، نقش و تاثیر بیشتری در کاهش اعوجاج تصویر فشرده دارند [۲۳].

۲-۲-۱ 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 پایین‌ترین زیر باند فرکانسی به عنوان گره‌های ریشه^۱ سرچشمه گرفته‌اند. آن‌گاه درخت صقر تعریف می‌شود، یک SOT که نسبت به یک آستانه مشخص (γ^n) دارای هیچ ضریب مهمی (بزرگ‌تری) نباشند. کدگذارهای موجک پیشنهاد شده در [۹]، [۱۲]، [۲۴] و [۲۵] از نوع مبتنی بر درخت صفر هستند. در میان این کدگذارها، الگوریتم SPIHT پیشنهاد شده توسط Said و Pearlman [۷]، بیش‌ترین شهرت و محبوبیت را دارد. الگوریتم SPIHT از کارایی بالایی در کدگذاری و پیچیدگی محاسباتی پایین برخوردار است، و در عین حال دنباله بیت‌های یکپارچه‌ای تولید می‌کند. کدگذاری SPIHT برای فشرده‌سازی و انتقال پیشرو بهینه^۲ طراحی شده است. فرآیند انتقال را می‌توان 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 به ما این اختیار را می‌دهد که از کدگذاری حسابی^۳ [۲۲] صرف نظر کنیم. در ادامه به توضیح مختصر نحوه عملکرد الگوریتم کدگذاری SPIHT می‌پردازیم. در الگوریتم SPIHT به غیر از پایین‌ترین زیر باند فرکانسی، ضرایب موجود در سایر زیر باندها به آرایه‌هایی 2×2 تحت عنوان گره‌های فرزند^۴ یک ضریب از یکی از زیر باندهای سطح تجزیه بالاتر دسته‌بندی می‌شوند. همچنین ضرایب پایین‌ترین زیر باند فرکانسی به آرایه‌هایی 2×2 تقسیم می‌شوند، اما همان‌طور که در شکل ۲-۴ (الف) و (ب) مشاهده می‌شود، ضریب

¹ Root Nodes

² Optimal Progressive Transmission

³ Arithmetic Coding

⁴ Offspring

گوشه بالا-چپ این آرایه‌ها هیچ فرزندی ندارند. بر اساس نمادهای تعریف شده در [۱۲]، مجموعه تمام فرزندان گره (i,j) را با $o(i,j)$ و مجموعه تمام نسل‌های^۱ گره (i,j) را با $D(i,j)$ (یک مجموعه از نوع "A") و مجموعه تمام نسل‌ها به جز فرزندان بی‌واسطه^۲ گره (i,j) را با $L(i,j)$ (یک مجموعه از نوع "B") نشان می‌دهیم. شکل ۲-۴ (ج) مجموعه‌های $o(i,j)$ ، $D(i,j)$ و $L(i,j)$ را در یک درخت SOT نشان می‌دهد. اطلاعات مهم در سه فهرست منظم ذخیره می‌شوند: فهرست پیکسل‌های غیر مهم^۳ (LIP)، فهرست مجموعه‌های غیر مهم^۴ (LIS) و فهرست پیکسل‌های مهم^۵ (LSP). در نخستین گام مقدار دهی 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 یک فهرست خالی خواهد بود.

فرآیند کدگذاری از پر ارزش‌ترین صفحه بیتی شروع شده و به سوی کم ارزش‌ترین صفحه بیتی پیش می‌رود. پیشروی کدگذاری در هر صفحه بیتی از طریق سه فهرست LIP، LIS و LSP صورت می‌گیرد. به بیان دیگر در هر صفحه بیتی، 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، یک بیت برای توصیف "اهمیت" آن استفاده می‌شود. اگر ضریبی مهمی نباشد در LIP باقی خواهند ماند و بیت بیش‌تری تولید نخواهد شد، در غیر اینصورت بیت علامت مربوطه تولید شده و آن ضریب به LSP جابجا می‌شود.

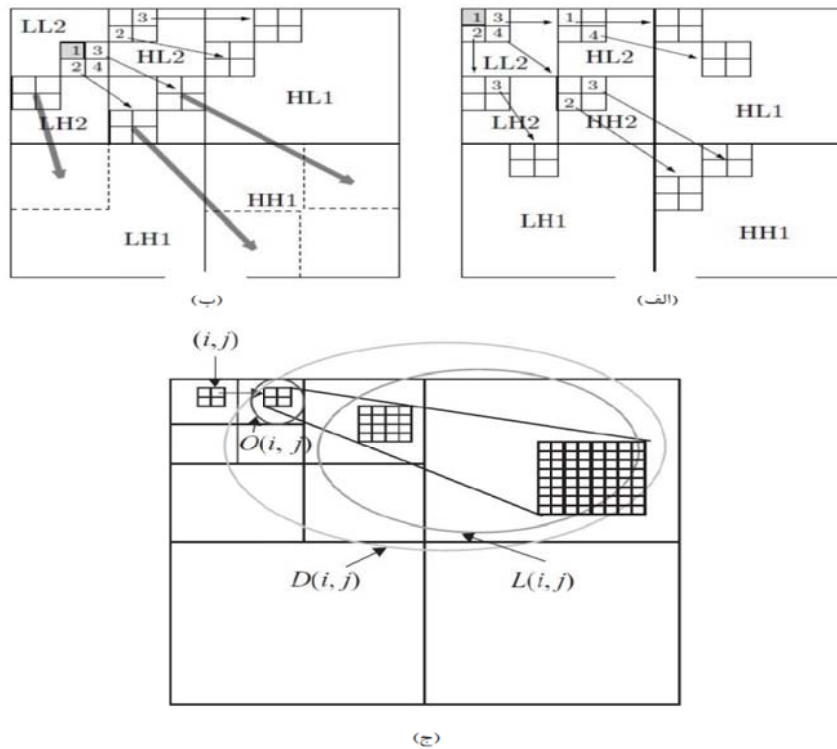
¹ Descendants

² Immediate Offsprings

³ List of Insignificant Pixels

⁴ List of Insignificant Sets

⁵ List of significant Pixels



شکل ۲-۴: درخت‌های جهت دار مکانی و بخش‌بندی مجموعه در SPIHT [۱۱ و ۲۳].

بطور مشابه، هر مجموعه موجود در LIS نیازمند یک بیت برای اطلاعات "اهمیت" است. مجموعه‌های غیر مهم در LIS باقی می‌مانند، در حالیکه مجموعه‌های مهم به زیر مجموعه‌هایی بخش بندی می‌شوند. یک مجموعه مهم نوع "A" به یک مجموعه نوع "B" و چهار ضریب فرزند بخش‌بندی خواهد شد، بطوریکه مجموع 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 LIS افزوده می‌شوند. از آنجایی که تمام مجموعه‌های غیر مهم اخیراً تولید شده به انتهای LIS افزوده می‌شوند، آن‌ها نیز در همان صفحه بیتی و به روشی یکسان پردازش خواهند شد تا زمانیکه هر زیر مجموعه 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.

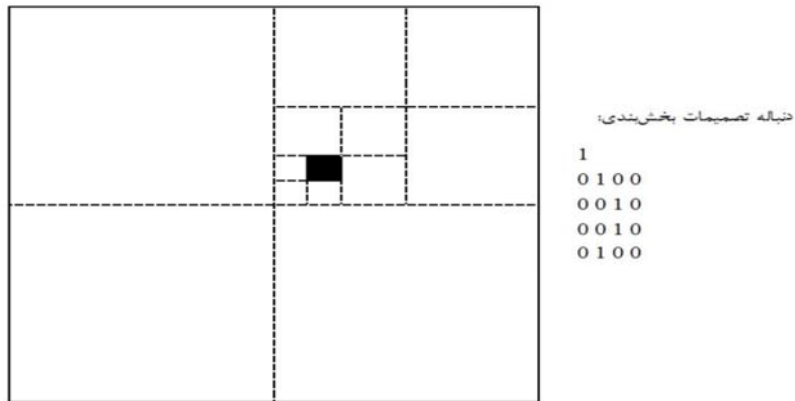
۲-۲-۲ کدگذاری مبتنی بر بلوک صفر

الگوریتم‌های مبتنی بر بلوک صفر تصویر تبدیل یافته را به بلوک‌هایی مجاور هم تقسیم کرده و آزمون اهمیت را بر روی این بلوک‌های منحصر بفرد اجرا می‌کنند. مشابه تعریف درخت صفر، در این-جا 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 جاسازی شده بخش‌بندی مجموعه (SPECK) [۲۷]، بخش‌بندی سلسله مراتبی بلوک زیرباند (SBHP) [۲۸]، EBCOT می-باشند. 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"^۳ 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 زمانی‌که ضرایب مهم پیدا شوند. در شکل ۲-۵ می‌توان عملیات بخش‌بندی درخت چهارگانه و کدگذاری را در یک ناحیه بلوکی 16×16 مشاهده کرد. مزیت کدگذاری بلوکی این است که بلوک‌های کوچک که نشان دهنده نواحی فرکانس بالا هستند، بطور مستقل از نواحی بزرگ‌تر با محتوای فرکانسی-مکانی پایین کدگذاری می‌شوند.

¹ Set Partitioning Embedded Block

² Subband Block Hierarchical Partitioning

³ Quadtree Partitioning



شکل ۲-۵: نمایش بخش‌بندی درخت چهارگانه و کدگذاری در یک ناحیه بلوکی 16×16 . در این جا آزمون اهمیت برای هر آرایه 2×2 یک ترتیب پیمایش Z را دنبال می‌کند. بیت "۱" در دنباله کدگذاری شده نشان دهنده ادامه بخش‌بندی یک گره معین است [۲۹].

۲-۳ کدگذاری ROI در حوزه موجک

تکنیک کدگذاری مبتنی بر ناحیه مطلوب (ROI) تصویر، به معنی فشرده‌سازی بی‌اتلاف و یا کم-اتلاف نواحی مطلوب درون یک تصویر، و فشرده‌سازی نواحی غیرمطلوب با اتلاف بیش‌تر است. بر اساس این ایده می‌توان علاوه بر دستیابی به یک نرخ فشرده‌سازی بالا، اطلاعات مهم تصویر را نیز بدون اتلاف و یا با اتلاف کم حفظ کرد [۳۰]. دو روش کدگذاری ROI در استاندارد JPEG2000 معرفی شده است: روش انتقال ماکزیمم (Maxshift) و روش مقیاس‌گذاری کلی [۳۱]. 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/

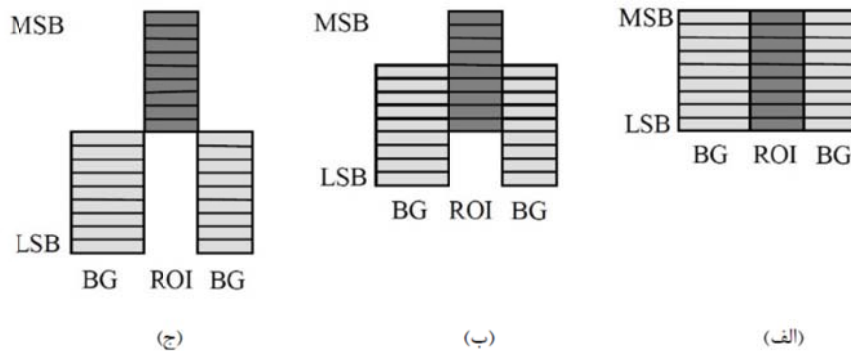
¹ Down-Shifting

² Up-Shifting

کدگشایی شود، دنباله بیت کدگذاری شده برش خورد یا فرایند کدگذاری / کدگشایی متوقف شود، نواحی ROI کیفیت بالاتری نسبت به BG خواهند داشت. شکل ۲-۷ نحوه جابجایی صفحه‌های بیتی را در روش‌های کدگذاری ROI مورد استفاده در استاندارد JPEG2000 نشان می‌دهد.

۲-۳-۱ مقیاس‌گذاری کلی

توصیفی از روش مقیاس‌گذاری کلی در شکل ۲-۶ (ب) به تصویر کشیده شده است. در این شکل مقدار مقیاس‌گذاری صفحه‌های بیتی با S نشان داده شده است. در روش مقیاس‌گذاری کلی، باید مقدار مقیاس‌گذاری و اطلاعات شکل نواحی ROI نیز به دنباله کدگذاری شده پیوست شوند. روش مقیاس‌گذاری کلی سه اشکال و نقص بزرگ دارد. اول این‌که در این روش نمی‌توان به راحتی با زیرباند‌های مختلف موجک به طرق مختلف برخورد کرد، که گاهی اوقات این مسئله یکی از خواسته‌های کاربران است. دوما این روش نیازمند کدگذاری و ارسال اطلاعات شکل نواحی ROI به واحد کدگشا است. 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 مشکل بخش دو استاندارد JPEG2000 [۳۲]، شکل‌های ROI را به مستطیل و بیضی محدود کرد، چرا که این شکل‌ها را می‌توان با تعداد بیت‌های کم‌تری 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.



شکل ۲-۶: روش‌های کدگذاری ROI در استاندارد JPEG2000 (الف) بدون مقیاس‌گذاری کلی، (ب) روش مقیاس‌گذاری کلی، $S=4$ ، (ج) روش ماکزیمم، $S=9$ [۳۰].

۲-۳-۲ انتقال ماکزیمم

یک راه حل بسیار موثر برای جبران نقایص مقیاس‌گذاری کلی، استفاده از روش انتقال ماکزیمم است. روش ماکزیمم را می‌توان حالت خاصی از روش مقیاس‌گذاری کلی در نظر گرفت، که در آن مقیاس‌گذاری آن قدر بزرگ است که هیچ هم‌پوشانی میان صفحه‌های بیتی ROI و BG^۱ وجود ندارد. بنابراین در روش ماکزیمم مقدار مقیاس‌گذاری 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 مقیاس‌گذاری، تمامی بیت‌های با ارزش به ROI در صفحات بیتی بالاتری نسبت به تمام بیت‌های با ارزش مربوط به BG قرار خواهند داشت.

شکل ۲-۶ (ج) انتقال صفحه بیتی را در روش ماکزیمم نشان می‌دهد. بنابراین در روش ماکزیمم شکل ROI برای واحد کدگشا مفهوم بوده، و لذا کدگذاری ROI با شکل دلخواه قابل پشتیبانی است. همچنین این روش می‌تواند به راحتی با زیرباندهای موجک مختلف به طرق مختلف رفتار کند. محدودیت بزرگ روش ماکزیمم این است که نمی‌تواند به راحتی اهمیت نسبی میان نواحی ROI و

^۱ Back Ground

BG را با تنظیم مقدار مقیاس‌گذاری 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 نمی‌تواند دریافت شود، تا زمانی‌که همه جزئیات ضرایب ROI بطور کامل کدگشایی شده باشند، حتی اگر این جزئیات نوین تصادفی نامحسوس و غیر قابل مشاهده باشند.

فصل سوم

تخصیص ناحیہ مطلوب

۳- فصل سوم

تصاویر پزشکی دیجیتال مانند اشعه X، بطور گسترده در تشخیص بیماری استفاده می‌شود. به علت حجم بالای تصاویر، فشردگی تصاویر جهت انجام سریع و انتقال مناسب و کاهش در فضای ذخیره شده‌ی تصاویر پزشکی ضروری و مورد نیاز است. تکنیک‌های فشردگی استفاده شده در صورتیکه فشردگی تصاویر پزشکی دیجیتال به عنوان ناحیه مطلوب برای تشخیص بیماری در قیاس با کل تصویر گرفته شده معمولاً کوچک باشد، بسیار مهم است. تکنیک‌های فشردگی بدون تلفات، فشردگی را بدون از دست دادن اطلاعات و با نرخ فشردگی پایین داده انجام می‌دهند و تکنیک‌های فشردگی با از بین رفتن اطلاعات، فشردگی را با سرعت بالا و با از بین رفتن مقدار ناچیزی از اطلاعات انجام می‌دهند. استفاده از تکنیک‌های بدون اتلاف در تصویر پزشکی برای انتقال و ذخیره‌سازی مفید نمی‌باشد و در تکنیک‌های با اتلاف امکان از بین رفتن اطلاعات حیاتی که برای تشخیص بیماری ضروری است، وجود دارد. برای حداکثر ساختن فشردگی، در این پایان‌نامه تکنیک مبتنی بر ناحیه مطلوب پیشنهاد و بررسی می‌شود.

ابتدا ناحیه مطلوب از پس‌زمینه باید جدا شود که این کار به راحتی با لبه‌گیری از تصویر، اتصال لبه‌های موجود در تصویر به منظور رفع شکستگی لبه‌ها در مرز بین پس‌زمینه و ناحیه مطلوب صورت می‌گیرد. واضح است که مرزهای مورد نظر و مطلوب، مرزهای نزدیک به ناحیه پرشدگی اطلاق می‌گردد.

پایگاه داده متشکل از ۲۵۰ عکس با فرمت png و tif است. که سایز آن‌ها بین ۱۸۰۰*۲۵۰۰ می‌باشد. حدود ۵۴ تا از آن دارای کیفیت بسیار پایین (نویز نمک و لفل) می‌باشد. این پایگاه داده با جمع‌آوری عکس از پزشکان و بیماران بدست آمده است.

برای شروع، ابتدا یک پایگاه داده باید تهیه می‌شود. این پایگاه با جمع آوری عکس‌های رادیوگرافی دندان و سپس اسکن کردن آن‌ها تهیه شد. پایگاه داده متشکل از ۵۴ عکس با کیفیت بسیار پایین که دارای نویز نمک و لفل است و از آن‌ها برای داده آموزشی استفاده شد. مابقی شامل ۲۲۰ عکس است که از کیفیت بسیار بالایی برخوردار است. این عکس‌ها با فرمت png, tif و سایز آن‌ها بین ۱۸۰۰*۲۵۰۰ و با حجم ۳ مگابایت می‌باشند.

۳-۱ روش مورد بررسی اول

برای بدست آوردن ROI از روش انتشار موج استفاده می‌شود [۲۵]. در این روش جداسازی هر دندان به طور مستقل انجام می‌گیرد. چون در عکس‌ها دندان‌ها خیلی سفیدتر از بقیه نقاط هستند، بین آن‌ها لثه یا فضای خالی وجود دارد که باعث تیره‌گی می‌شود. پس ناحیه تاریک بین دندان‌ها را به عنوان مرز تشخیص می‌دهیم. لذا نیازمند الگوریتمی هستیم که بر همین اساس دندان‌ها را مرزبندی کند، یعنی خطی از بین دندان‌ها (فاصله فک بالا و پایین) عبور دهد که در آن پارامتر شدت روشنایی را به پارامتر سرعت انتشار تبدیل کند. به این معنی که پیکسل‌ها اجزایی از زمینه باشند که سرعت انتشار امواج (مثلا امواج صوتی) در آن‌ها متفاوت است. هرچایی که شدت روشنایی بیشتر باشد (سفیدتر) سرعت انتشار کم‌تر است. موج از مسیرهای مختلف منتشر می‌شود تا از کوتاه‌ترین مسیر به نقطه مورد نظر برسد.

۳-۱ مشکلات این روش

(۱) این روش بر روی عکس‌هایی با سایز حداکثر ۸۰۰*۸۰۰ قابل اجرا است، ولی سایز عکس‌هایی که در دسترس بود حدودا ۱۸۰۰*۲۵۰۰ بود، عملیات پردازش بسیار سنگین و زمان‌بر می‌شود.

۲) در این روش ابتدا یک ماسک اولیه تعریف می‌شود، سپس هر دندان بطور جدا تقسیم‌بندی می‌شود، اما این روش برای پرکردگی مناسب نیست، زیرا از آنجا که اندازه هر پرکردگی و فرم آن در هر دندان متفاوت است، ماسک مشخصی برای پرکردگی وجود ندارد.

۲-۳ روش مورد بررسی دوم

لازم به ذکر است که مدل مار عبارت است از یک مرز فعال که می‌تواند تحت تاثیر انرژی‌های داخلی و خارجی به سمت ویژگی‌های مطلوب در تصویر که معمولاً لبه‌ها هستند حرکت کند [۲۴]. نیروهای داخلی مرز را پیوسته نگاه داشته و از پیش‌روی بیش از حد مرز آن جلوگیری می‌کنند، در حالی که نیروهای خارجی مرز را به سمت ویژگی‌های مورد نظر در تصویر سوق می‌دهند.

برای بدست آوردن ROI ابتدا روش مار استفاده می‌شود. با مد نظر قرار دادن ویژگی‌های کنتراست تصویر، برنامه کامپیوتری برای پیاده‌سازی انواع مختلف مدل‌های مار که در مقالات مختلف بیان شده‌اند، نوشته شده است. با این وجود به علت یکنواختی کنتراست تصویر و به تبع آن، گنگ بودن مرزها، نتایج خوبی در پی‌نداشت و با ۳ مشکل عمده به شرح ذیل مواجه بود:

(۱) تصاویر موجود دارای نویز فلفل و نمک بود، بنابراین مار مستقیماً به هر پیکسل

نویزی واکنش نشان می‌داد. برای رفع این مشکل از فیلتر میانه استفاده شد.

(۲) کندروی کد نرم افزار در پیش‌روی ناحیه پوسیدگی (نواحی سفید در تصویر).

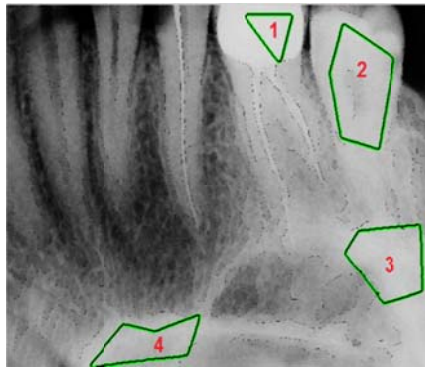
(۳) عدم تشخیص لبه‌ها و عبور از روی آن.

برای رفع مشکل سوم، از ارتقاء لبه استفاده شد. عملگر کنی از با ترشولد ۸ بر روی تصویر اعمال

شد، سپس تصویر این مرحله با تصویر اصلی جمع شد، در نتیجه مرزها واضح‌تر شدند. این‌طور بنظر

می‌رسید که دیگر مار بموقع متوقف شود، اما بعد از اجرای کد، باز هم مار از روی مرزها عبور کرد که در شکل ۳-۲ نشان داده شده است. بخاطر عدم پیوستگی و فراوان بودن مرزها، بعضی اوقات، مار به اشتباه روی لبه زائد متوقف می‌شد، و در جاهای دیگر به دلیل ناپیوستگی مرز، مدل مار به آرامی از این ناپیوستگی عبور می‌کرد که اگر انرژی درونی آن کاهش می‌یافت احتمال متوقف شدن مار بیش‌تر بود.

برای رفع مشکلات بالا، تصمیم گرفته شد، که سطح انرژی پایین آورده شود. بدین منظور یک ناحیه جدید با مرزهای واضح و اختلاف شدت روشنایی جدید در کنار عکس مورد نظر اضافه شد، (به اندازه $1/3$ طول عکس پیکسل‌هایی با شدت روشنایی صفر اضافه شد، سپس در این ناحیه‌ای که اضافه شده، یک مستطیل سفید قرار داده شد که تعداد پیکسل‌های آن بستگی به عرض تصویر اولیه دارد). سپس این ناحیه همراه با ROI دندان به مدل مار داده شد. بعد از میانگین‌گیری، مدل مار ROI دندان با شدت روشنایی بیش‌تری در نظر می‌گرفت و با این کار مار متوقف می‌شد. این ایده بر روی پایگاه داده عکس اجرا شد و کم‌تر از ۲۰ درصد به جواب مطلوب رسید، زیرا بعضی از عکس‌ها آن‌قدر بی‌کیفیت بود که این روش هم نمی‌توانست مار را متوقف کند. هم‌چنین مدل مار برای شروع نیازمند یک ماسک اولیه است که این ماسک باید توسط کاربر یا بصورت هوشمند انتخاب شود. اما ناحیه پرشدگی منحنی و سایز مشخصی ندارد، در نتیجه در انتخاب ماسک اولیه با مشکل مواجه می‌شود. هم‌چنین سرعت پردازش در مدل مار بسیار پایین است. سرانجام، با مشکلاتی که مطرح شد مدل مار با شکست مواجه شد. نتایج در شکل ۳-۱ و ۳-۲ نشان داده شده است.

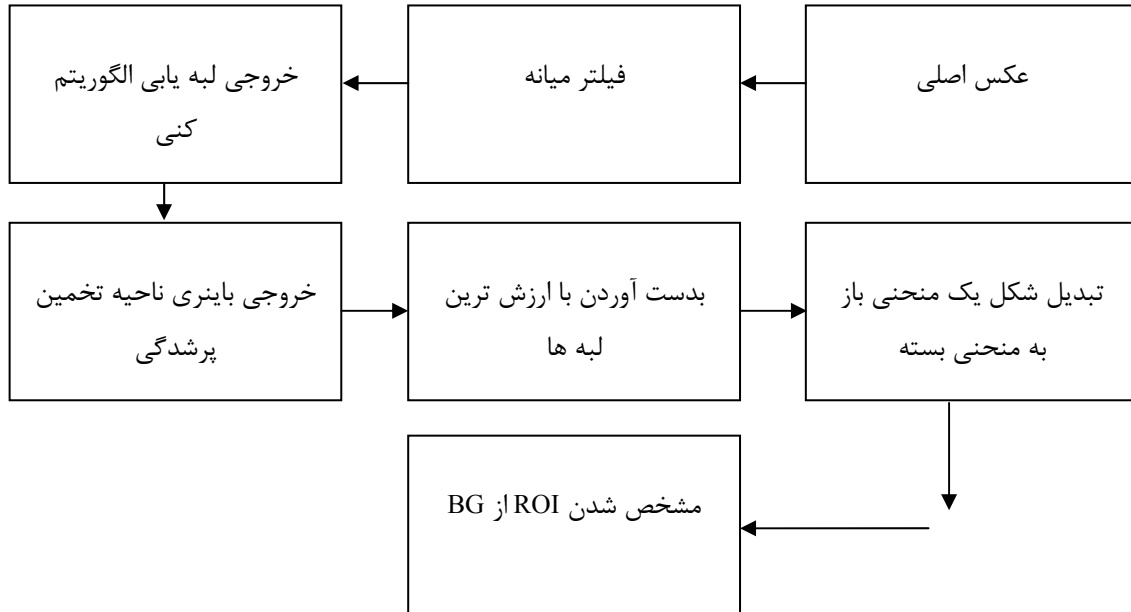


الف

شکل ۳-۱: الف) تصویر ورودی ب) تصویر خروجی از مدل مار

در شکل ۳-۲ ناحیه ۱ فقط پرشدگی است و نواحی ۲، ۳، ۴ کاملاً اشتباه تشخیص داده شده است، و جزء نواحی پرشدگی نیست. برای مثال ناحیه ۴ به دلیل کیفیت پایین تصویر و فقدان اسکنر مخصوص عکس‌های رادیوگرافی، در قسمت پایین عکس روشنایی بیش از حدی بوجود آمده که هیچ ارتباطی به ناحیه پرشدگی ندارد. اما مدل مار به دلیل واضح نبودن مرزها، از مرزهای باریک که شکاف بسیار کوچکی داشت عبور می‌کرد و با در نظر گرفتن تعداد دفعات انتخابی توسط کاربر، مار تمام این نواحی را در ابتدای فرآیند در نظر گرفته و در طول فرآیند این نواحی را رشد می‌دهد، که کاملاً اشتباه است. به همین علت قصد بر آن شد که مرزها بسته (پیوسته) شوند، به بیان ملموس‌تر راستا و جهتی انتخاب شود و با توجه به زاویه اولیه، این مرز ناپیوسته به مرز مورد نظر متصل شود که در بخش بعدی توضیح داده می‌شود.

۳-۳ روش پیشنهادی



شکل ۳-۲: فلوجارت الگوریتم پیشنهادی (روش سوم)

قرار بر برگیری نواحی مطلوب از عکس شد. مشاهدات حاکی از این است که ماده پرشدگی استفاده شده توسط پزشکان دارای روشنایی بالا و قابل توجهی است. برای انجام این کار ابتدا با بهره‌گیری از عملگر canny مرتبه ۸، تصویر لبه‌یابی می‌شود. با وجود ایجاد لبه‌های زیاد، تمام لبه‌ها مفید نمی‌باشند، بالطبع باید شماری از لبه‌های غیر مفید حذف شوند. با این وجود، این پرسش مطرح می‌شود که چگونه این امر امکان پذیر است. آنچه واضح و مبرهن است مرزهای نزدیک به ناحیه پرشدگی در شمار بهترین مرزها می‌باشند. لذا به منظور تمایز نواحی مفید از دیگر نواحی تصویر اختلاف بر پایه شدت روشنایی تشدید می‌شود. برای انجام این امر ابتدا تصویر با فیلتر میانه 50×50 مات می‌شود. پس از آن از تصحیح گاما بهره می‌برند. به موجب این امر نواحی پرشدگی تا حدودی واضح‌تر خواهند شد که در تصویر ۳-۵ نشان داده شده‌است. پس از آن بر روی تصویر اولیه فیلتر میانه 15×15 زده شد تا باعث از بین رفتن نویز نمک و فلفل شود. پس از انجام

مراحل فوق، فواصل مورد نظر بدست آورده می‌شود. سپس مجموع مقادیر پیکسل مربوط به یک مرز محاسبه می‌شود. این مقادیر مطابق است با همان فاصله از ناحیه مطلوب.

اگر فاصله کوچک باشد بدین مفهوم می‌باشد که مرز مورد بررسی جزء مرزهای مفید است. علاوه بر این مقدار آستانه به صورت تجربی و با آزمایش بر روی عکس‌های مختلف تعیین شده است. لکن با تکرار آزمایش بر روی عکس‌های مختلف مشکلی پدید آمد. این مشکل بیان‌گر این بود که بعضی از مرزها در نزدیکی ناحیه مطلوب بسیار بلند و پیوسته‌اند. اگر مجموع مقادیر پیکسل‌های این مرزهای طولیل محاسبه شود، مقدار قابل توجهی بدست می‌آید که برحسب مقایسه با آستانه چنین مرز مفیدی غیر قابل اغماض است.

در جهت حل این مشکل مقادیر بدست آمده (فاصله) را معکوس می‌کنیم. اما با این کار چون در خروجی تابع مقادیری از جمله صفر نیز وجود داشت، با معکوس نمودن آن ابهام ایجاد می‌شود. پس برای رفع این مشکل، مقادیر خروجی با عدد ۱ جمع شده و پس از آن معکوس می‌شود. به سبب این کار اعداد بزرگ کوچک می‌شوند و در نتیجه مرزهای نزدیک به ناحیه مطلوب که طولیل نیز بودند، باقی مانده و حذف نمی‌شوند. حاصل تلاش‌های فوق بدست آمدن با ارزش‌ترین لبه‌ها است که مرزهای نزدیک پرشدگی می‌باشند. به کمک این لبه‌ها شماری از نواحی که شدت روشنایی بیشتری داشتند ایجاد می‌شوند. اکنون باید دانست که آیا تمامی این نواحی جز نواحی پرشدگی هستند یا خیر؟ به همین منظور از نواحی باقی مانده مرکز سطح گرفته می‌شود. در نتیجه یک سطح به یک نقطه تبدیل می‌شود. حال اگر نقطه مذکور در بالای $2/3$ ستون‌های تصویر واقع شده باشد، غیر قابل قبول است و ملزم به حذف آن‌ها هستیم. لازم به ذکر است که دندان‌ها در زیر $2/3$ ارتفاع تصویر واقع شده‌اند. در مرحله کنی لبه‌های زیادی استخراج شد که پیوسته نیستند برای رفع ناپیوستگی الگوریتمی به شرح زیر پیشنهاد شد:

۳-۳-۱ الگوریتم اتصال

در ابتدای امر از میان یک ماتریس بزرگ $1800 * 2500$ ، بر حسب تصادف، ماتریس $200 * 200$ انتخاب می‌شود. یک ماتریس بر حسب لبه‌ها و دومی دارای یک پیکسل ۱ و مابقی صفر است. از برهم کنش آن‌ها و محاسباتی خاص، فاصله اقلیدسی پیکسل‌ها با پیکسل مورد نظر که برایش مقدار ۱ تعریف شده بود، محاسبه می‌شود. پس از آن نزدیک‌ترین فاصله انتخاب شده و انتهای مرز آن با یک خط به پیکسلی که نزدیک‌ترین فاصله را دارد متصل می‌شود. که با در نظر گرفتن این شرط به عنوان تنها شرط، با مشکلات زیر مواجه شدیم:

(۱) انتهای یک مرز به خودش متصل شود.

(۲) انتهای یک مرز به مرزهایی که در پشت آن با فاصله کمی قرار گرفته‌اند، متصل می‌-

شوند.

گام بعدی مشخص شدن جهت و راستا می‌باشد. به این دلیل که به صورت اشتباه به مرز دیگر یا حتی به خودش متصل نشود، یک ماتریس $200 * 200$ از مرزها برگزیده می‌شود. پس از آن یک ماتریس با ابعاد کوچک‌تر ($10 * 10$) در انتهای مرز قرار داده می‌شود. لازم به ذکر است که این ماتریس کوچک، دلخواه است. زاویه انتهای مرز در راستای x محاسبه و θ_0 نام گذاری می‌شود. برای اصلاح زاویه θ_0 ، یک خط عمود بر انتهای مرز کشیده می‌شود. اگر راستایش با θ_0 هم‌جهت باشد، θ_0 بدون تغییر باقی می‌ماند. در غیر این صورت θ_0 برابر است با θ_0 به اضافه 180 . θ_1 زاویه مرز مورد نظر با مهم‌ترین مرزهای اطرافش (مرزهای ناحیه پرشدگی). پس از آن ماتریس سوم این‌گونه تعریف می‌شود که دارای یک پیکسل یک و مابقی صفر باشد. پس از برهم کنش ماتریس سوم بر ماتریس اول فاصله اقلیدسی دیگر پیکسل‌ها با پیکسلی که مقدارش یک بود محاسبه می‌شود. این فاصله I نامیده می‌شود.

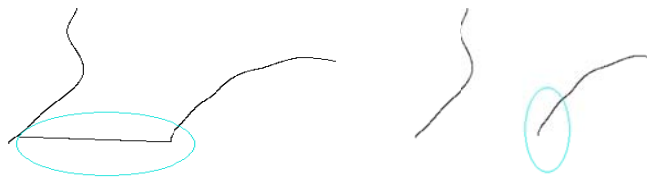
زاویه تمام پیکسل‌های تصویر نسبت به نوک انتهای مرز با رابطه (۱-۳) بدست آورده شده، و سپس فاصله نهایی (r_f) که جهت نیز در آن موثر است، به کمک رابطه (۲-۳) محاسبه می‌شود.

$$\theta_0 = \arctan\left(\frac{dy}{dx}\right) * \frac{180}{\pi} + 180 \quad (1-3)$$

$$r_f = r + [r * (1 + \cos(\theta_1 - \theta_0))^2] \quad (2-3)$$

dx و dy میزان تغییرات مختصات پیکسل مورد نظر در راستای y و x نسبت به انتهای لبه می‌باشند.

بدین ترتیب کوتاه‌ترین فاصله مشخص شده، و با خط راست به انتهای مرز مورد نظر متصل می‌شود که در شکل ۳-۴ نشان داده شده است. در نتیجه، انتهای مرز به مرزهای پشت سر، یا به خودش متصل نمی‌شود.



ب

الف

شکل ۳-۴: نحوه اتصال یک مرز به مرز دیگر

خروجی این قسمت، ناحیه‌ای شامل چندین منحنی با اندازه‌های متفاوت از هم هستند، ولی همه این نواحی مربوط به پرشدگی نیست.

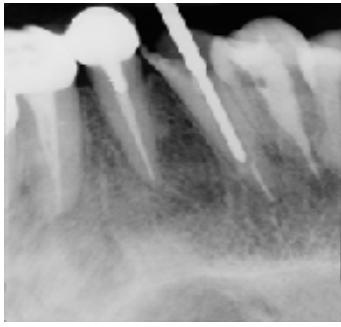
۳-۴ چگونه نواحی پرشدگی انتخاب شود؟

حاصل فرآیندهای ذکر شده، شماری ناحیه سفید است، اما همه آن‌ها پرشدگی نیست. واضح است که نواحی پرشدگی شدت روشنایی بالایی دارند. با علم بر این که در آزمایش بر روی عکس‌های مختلف، میانگین نواحی پرشدگی معادل ۱۸۰ یا بیش‌تر از آن است. در ابتدای امر، میانگین نواحی

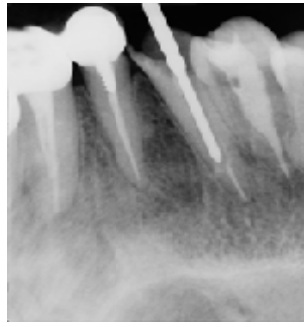
گرفته می‌شوند و میانگین‌هایی که بیش‌تر از ۱۸۰ باشند، قابل پذیرش‌اند. در برخی موارد نواحی‌ای یافت می‌شوند که میانگین آن‌ها اختلاف ناچیزی با ۱۸۰ دارند. برای این نواحی شرط دیگری نیز اعمال می‌شود و در نواحی‌ای که از پرشدگی بودن آن‌ها اطمینان حاصل شده، تا شعاع معینی هاله ایجاد کرده و میانگین آن‌ها و ناحیه داخلی‌شان محاسبه شده، این عمل برای تمامی نواحی سفید تکرار می‌گردد. اگر مقدار اختلاف میانگین آن‌ها برابر با حدود اختلاف نواحی پرشدگی باشد، آن نواحی پذیرفته شده‌اند.

۳-۵ نتایج تجربی برای تصاویر دندان

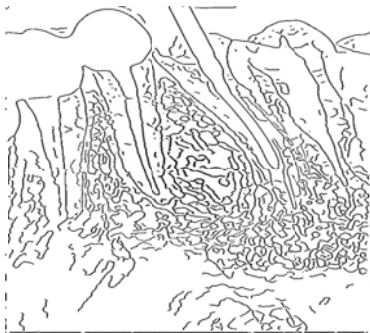
در عکس‌های زیر روند یافتن ROI که همان نواحی پرشدگی می‌باشد نشان داده شده است. در طی اجرای نرم افزار، در هر مرحله بصورت مجزا از آن عکس گرفته می‌شود.



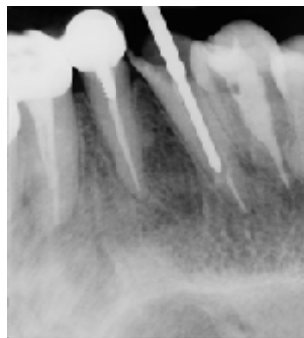
الف



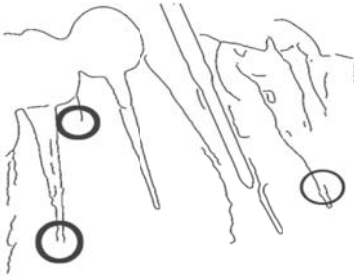
ب



ج



د



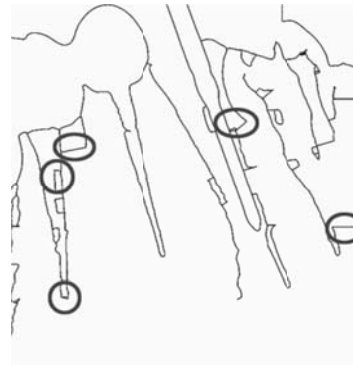
ه



و



ز



ح

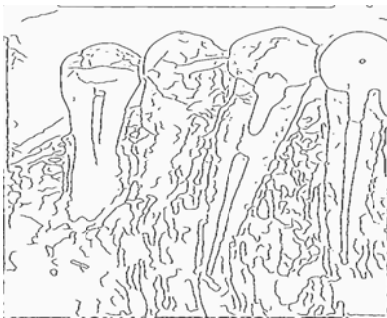
شکل ۳-۴: الف) عکس اصلی ب) فیلتر میانه ۱۵*۱۵ پ) فیلتر میانه ۵۰*۵۰ ت) خروجی لبه‌یابی
 الگوریتم canny بر روی عکس ب. ث) خروجی باینری ناحیه تخمین پرشدگی ج) با ارزش‌ترین لبه‌ها در نزدیکی
 محدوده‌ی پرشدگی ث. چ) لبه‌ها از شکل یک منحنی باز به منحنی‌های بسته ح) ناحیه مطلوب نهایی به کمک
 آستانه‌گیری



ب



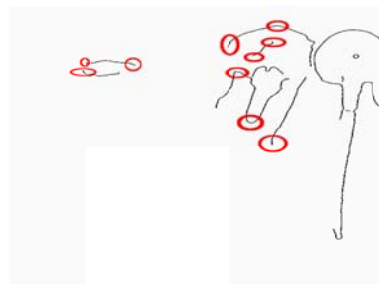
الف



ت



ب



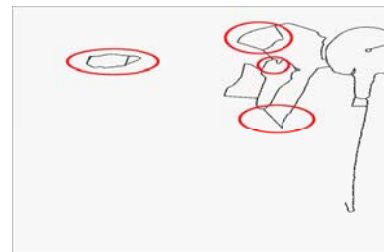
ج



ث



ح



چ

شکل ۳-۵: الف) عکس اصلی ب) فیلتر میانه 15×15 پ) فیلتر میانه 50×50 ت) خروجی لبه‌یابی
 الگوریتم canny بر روی عکس ب. ث) خروجی باینری ناحیه تخمین پرشدگی ج) با ارزش‌ترین لبه‌ها در
 نزدیکی محدوده‌ی پرشدگی ث. چ) لبه‌ها از شکل یک منحنی باز به منحنی‌های بسته ح) ناحیه مطلوب
 نهایی به کمک آستانه‌گیری

فصل چهارم

فشرده سازی

۴- فصل چهارم

۴-۱ ضرورت فشرده‌سازی

تصاویر پزشکی معمولاً حجم زیادی از داده‌ها را تولید می‌کنند. انتقال این حجم عظیم اطلاعات به بیمارستان‌های دیگر، عملاً بر روی شبکه‌های کم سرعت موجود غیرممکن است، زیرا پهنای باندی که در دسترس عموم قرار دارد، توان انتقال این حجم داده را ندارد و شبکه‌هایی که توان انتقال بالایی را دارند، گران و استفاده آن برای بسیاری از افراد عملی نمی‌باشد.

از طرف دیگر ذخیره کردن این حجم داده نیز عملاً غیرممکن می‌باشد و حافظه‌های جانبی که امروزه در بازار موجود و برای عموم قابل استفاده می‌باشد، گنجایش حداکثر چند ترابایت را دارند. حتی اگر حافظه‌های جانبی توان ذخیره کردن این حجم داده را داشته باشند، تمام پهنای باند گذرگاه داده، درگیر این انتقال خواهد شد. این اتلاف منابع، محدود به حافظه جانبی و پهنای باند نمی‌باشد، بلکه منابع دیگر مانند حافظه اصلی و واحد پردازش نیز به هدر می‌رود.

به همین دلیل نیازمند روش‌هایی هستیم که با کم‌ترین حجم داده، بیشترین حجم اطلاعات را منتقل و یا ذخیره کنیم.

۴-۲ تعریف فشرده‌سازی و انواع آن

فشرده‌سازی تصویر، داده مورد نیاز برای نمایش یک تصویر با شباهت نزدیک به تصویر اصلی را توسط حذف اطلاعات اضافی و تکراری کاهش می‌دهد. فشرده‌سازی تصویر یکی از اقسام فشرده‌سازی داده است که در آن حجم، تصویر و زمان انتقال تصاویر را برای اهداف مختلف کاهش می‌دهند. فشرده‌سازی به دو دسته فشرده‌سازی با از دست دادن اطلاعات و فشرده‌سازی بدون از دست دادن اطلاعات، تقسیم می‌گردد. در دسته اول اطلاعات تصویر بازیابی شده با تصویر اصلی‌اندکی در نمایش

جزئیات متفاوت می‌باشد. در این روش امکان دستیابی به درصد بالایی از فشرده‌سازی وجود دارد. این‌گونه روش برای کاربردهایی که داده فشرده شده باید بعد از بازیابی دقیقاً مانند داده اولیه باشد، (مانند فایل‌های متنی) مناسب نیستند.

در روش فشرده‌سازی بدون از بین رفتن اطلاعات، اطلاعات بازیابی شده دقیقاً با اطلاعات اولیه برابر است، اما در این روش میزان فشرده‌سازی داده محدود می‌باشد. با توجه به این‌که تجهیزات تصویربرداری پزشکی اغلب دیجیتال می‌شوند، فشرده‌سازی این تصاویر برای ذخیره و انتقال ضروری است. فشرده‌سازی تصاویر پزشکی ضروری است به طوری که تصاویر بتوانند با سرعت بالا حتی از طریق پهنای باند پایین‌تر با قابلیت اعتماد بالا برای تشخیص پزشکی در مکان‌های دور ارسال شوند. چالشی که با فشرده‌سازی تصاویر پزشکی روبرو است، با این‌که نرخ فشرده‌سازی بالایی از تصاویر مطلوب است، قابلیت استفاده از تصاویر بازسازی شده بستگی به برخی مشخصات مهم از تصاویر اولیه دارد که برای تشخیص صحیح می‌بایست حفظ شوند.

در فشرده‌سازی نکته حائز اهمیت این است که ناحیه مطلوب (پرشدگی) تا حد امکان کم‌تر فشرده شود، به این دلیل که پزشکان توانایی حداکثر بهره‌گیری را از عکس مورد نظر داشته باشند. از طرف دیگر برای کاهش حجم این عکس‌ها از تکنیک ROI استفاده شده است و در سایر نواحی که شامل پس‌زمینه و یا دیگر قسمت‌های دندان که دیگر جز ناحیه مطلوب نیست، بیت کم‌تری به آن اختصاص داده شود. پس حتی المقدور پس‌زمینه فشرده شد. اما کیفیت آن به نسبت قبل کم‌تر خواهد شد. برای فشرده‌سازی از کدگذار SPIHT و روش Maxshift استفاده شده است.

قبل از انتخاب دو روش زیر، از نرم افزار فشرده‌سازی jpeg-2000 استفاده شد. متأسفانه نتیجه مطلوب نبود. این نرم افزار عکس را با کیفیت بسیار پایین فشرده کرد. هم‌چنین در بعضی از قسمت‌های تصویر خروجی، نواحی‌ای که شامل نقاط مطلوب نبود، بعنوان ناحیه مطلوب در خروجی نشان می‌داد.

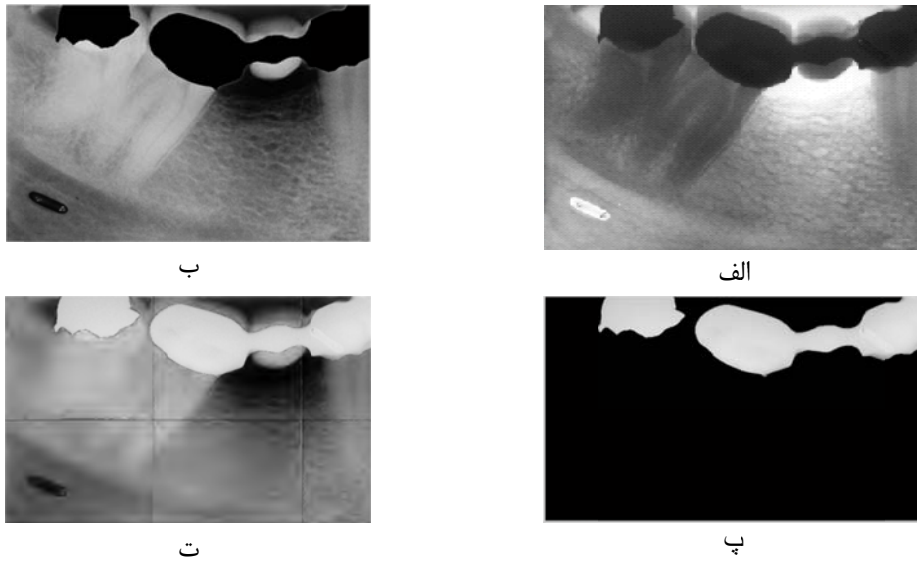
۳-۴ روش اول

در این مبحث، عکس به کمک تبدیل موجک و کدگذاری SPIHT فشرده می‌شود. از آنجایی که در تصاویر دندان، ناحیه مطلوب از اهمیت بسزایی برخوردار است، تا حد امکان کم‌تر فشرده می‌شود تا جزئیات آن از بین نرود. اما بخش پس‌زمینه را که اطلاعات مهمی در آن نیست بیش‌تر فشرده می‌کنند و بالطبع کیفیت بسیار پایینی خواهد داشت.

در ادامه، روندی که برای فشرده‌سازی انجام شده‌است، بصورت کامل شرح داده می‌شود:

ابتدا عکس به ماتریس مربعی تبدیل شد، عکس‌های اولیه در حدود $1800 * 2500$ بودند، با این کار حجم عکس بصورت چشم‌گیری افزایش یافت و عمل پردازش آن بسیار طولانی شد. هنگامی که عکس با کدگذار SPIHT به بیت استریم تبدیل شد، مشکل کمبود حافظه بوجود آمد. برای حل این مشکل عکس به مربع‌های $32 * 32$ تبدیل شد، اما ناپیوستگی‌هایی در عکس بازسازی شده بوجود آمد. برای رفع این مشکل، عکس به 6 مربع $1024 * 1024$ تبدیل و بر روی هر یک تبدیل موجک گرفته و با SPIHT کدگذاری شد و بیت استریم بدست آمده بصورت مجزا ذخیره شد. این فرایند بر روی هر دو ناحیه اعمال شد. در نهایت هر 6 قالب به همراه اطلاعات سائز اصلی عکس در یک فایل ذخیره شد. برای کدگشایی، فایل نهایی بازخوانی و کلیه فرآیندهای طی شده بالعکس تکرار شد. بعنوان مثال، در شکل ۴-۱، قسمت الف، تصویر ورودی را که شامل 4 ناحیه پرشدگی می‌باشد نشان می‌دهد. قسمت ب، شامل پس‌زمینه است که جای نواحی مطلوب در آن خالی است، قسمت پ، نواحی مطلوب (پرشدگی) را نشان می‌دهد. در نهایت با تقسیم کردن عکس به 6 مربع مجزا و انجام عملیاتی که در بخش ۳-۴ ذکر شد، عکس قسمت ت حاصل می‌شود.

نتایج بر روی پایگاه داده عکس به شرح زیر است:



شکل ۴-۱: الف) تصویر ورودی ب) ناحیه پس‌زمینه پ) ناحیه ROI ت) تصویر بازسازی SPIHT

جدول ۴-۱: نتایج روش اول

نتایج بدست آمده	ROI	BG
فشرده‌سازی (درصد)	۱.۴۵۴۴	۹۳.۱۴۹۴
خطا حداقل مربعات	۰.۱۰۲۰	۱۴۶.۸۵۷۶
سیگنال به نویز	۵۸.۰۴۶۰	۲۶.۴۶۱۸

در مثال بالا نرخ بیت ۰.۴ و فشرده‌سازی کل ۹۴.۶ در صد بدست آورده شد.

۲-۴ روش دوم (ماکزیمم شیفت)

تفاوت این روش با روش قبل به شرح زیر است:

در تمام مراحل (فیلتر کردن، لحاظ کردن ROI و کدگذاری) بخش مطلوب و پس‌زمینه را از هم

جدا نمی‌شوند بلکه در کنار هم پردازش می‌شوند.

ابتدا بر روی عکس ورودی عمل تبدیل موجک اعمال می‌شود. خروجی این قسمت رند شده و

سپس با استفاده از کدگذار SPIHT و با نرخ بیت پایین (کیفیت کم) کدگذاری می‌شود. خروجی این

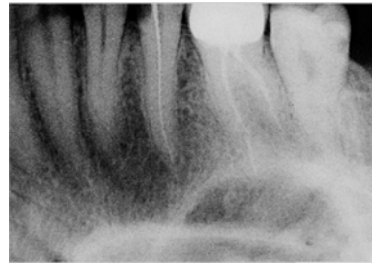
قسمت C_1 نامیده می‌شود. حال برای افزایش کیفیت ROI یک بیت استریم دیگر تولید می‌شود، بنابراین بیت استریم مرحله قبل کدگشایی می‌شود که C_2 خروجی این قسمت شامل یک ماتریس هم ابعاد با عکس اولیه است و تقریبی از ماتریس C_1 است و تفاضل این دو ماتریس را C_T می‌نامیم.

در ماتریس C_T ، بیش‌ترین مقداری که در ناحیه BG بود انتخاب و با S نمایش داده می‌شود. سپس مقادیر مربوط به ROI در 2^S ضرب شده و افزایش داده می‌شوند. با این کار اختلاف شدیدی در خروجی تبدیل موجک بین دو ناحیه ROI و پس‌زمینه ایجاد می‌شود. سپس ماتریس C_T با نرخ بیت بالا کد می‌شوند، چون در مرحله قبل به ناحیه ROI مقادیر بیش‌تری نسبت داده شده بود، در موقع کد کردن ROI از اولویت بیش‌تری برخوردار بوده‌است. در واقع داده‌های حاصل از بیت استریم ۲، به ROI اختصاص دارد. بعد از این مرحله بیت استریم‌های حاصل ذخیره شده اند و موقع بازسازی عکس مراحل بالا باید انجام شوند. نمونه‌ای از تصاویر پردازش شده در شکل ۴-۲ نشان داده شده است. در قسمت الف، تصویر ورودی را که شامل ۱ ناحیه پرشدگی می‌باشد، نشان می‌دهد. قسمت ب، شامل پس‌زمینه است که جای نواحی مطلوب در آن خالی است، قسمت پ، نواحی مطلوب (پرشدگی) را نشان می‌دهد. در نهایت با انجام عملیاتی که در بخش ۴-۴ ذکر شد، عکس قسمت ت حاصل می‌شود.

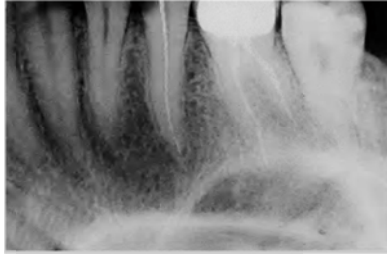
نمونه‌ای از تصاویر پردازش شده در شکل ۴-۲ نشان داده شده است.



ب



الف



ت



پ

شکل ۴-۲ : الف) تصویر ورودی ب) ناحیه پس زمینه پ) ناحیه ROI ت) تصویر بازسازی Maxshift

فصل پنجم

نتیجہ

۵- فصل پنجم

نتیجه گیری

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 نرخ بیت نیز وجود دارد که خیلی زیاد مورد استفاده قرار می‌گیرد. منظور از نرخ بیت تعداد بیت‌هایی است که برای ذخیره‌سازی هر پیکسل مصرف می‌شود. چنانچه فرض کنیم f تصویر اصلی با ابعاد $M*N$ پیکسل و نرخ بیت b ، و f^c تصویر فشرده شده با ابعاد $M*N$ پیکسل و نرخ بیت $B.R$ باشد، نرخ فشرده‌سازی بصورت زیر تعریف خواهد شد:

$$C.R = \frac{S^f}{S^{f^c}} \quad (۱-۵)$$

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 فشرده شده f^c است. با معلوم بودن ابعاد تصویر f و نرخ بیت آن داریم:

$$S^f = N M b \quad (۲-۵)$$

بنابراین نرخ بیت تصویر فشرده شده را می‌توان از رابطه زیر بدست آورد:

$$B.R = \frac{b}{C.R} \quad (۳-۵)$$

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

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

$$PSNR(f, f^{dec}) = 10 \log_{10} \left(\frac{MN (2^b - 1)^2}{\|f^{dec} - f\|^2} \right) [dB] \quad (4-5)$$

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 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.

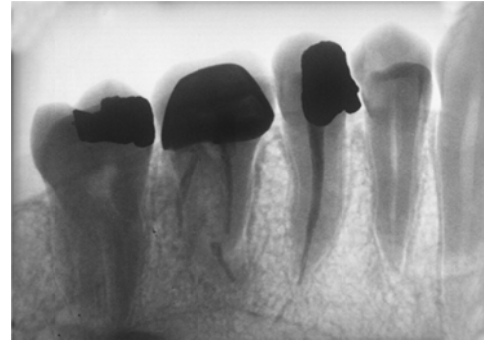
۵-۱ فشرده سازی

روش SPIHT و ماکزیمم شیفت بر روی پایگاه داده عکس اجرا شدند. با توجه به نتایج شکل ۵-۱ و جدول ۵-۱ مشاهده می شود که میزان فشرده سازی در روش ماکزیمم شیفت ۹۶.۰۳٪ و در روش SPIHT ۹۵.۸۸٪ است، که میزان فشرده سازی در روش ماکزیمم شیفت بیشتر است. اما خطای حداقلی مربعی ناحیه مطلوب در روش SPIHT ۰.۰۷۳۳ و در روش ماکزیمم شیفت ۰.۷۶۳۲ است، که این نتایج بیان گر دقت کم روش ماکزیمم شیفت است. سرعت پردازش در روش SPIHT

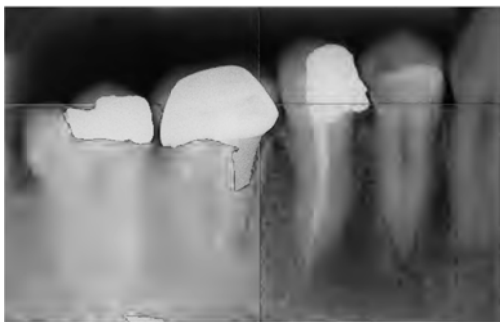
۳۵.۲۴ ثانیه و در روش ماکزیمم شیفت ۶۹.۹۰۷ ثانیه است، که این نتایج بیانگر سرعت بیشتر پردازش در SPIHT است، زیرا در مرحله کدگذاری چند بار رشته بیت تولید می‌شود که این کار باعث کندی کد است.



ب



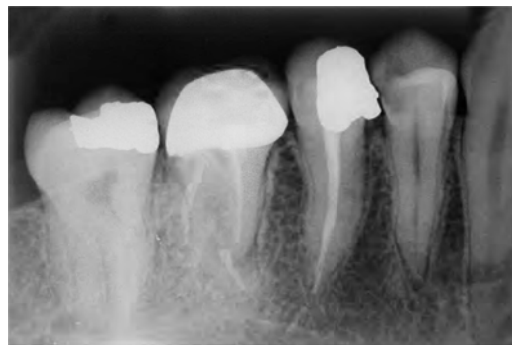
الف



ت



پ



ث

شکل ۵-۱: الف) تصویر ورودی ب) ناحیه ROI پ) ناحیه BG ت) تصویر بازسازی SPIHT ث) تصویر بازسازی Maxshift

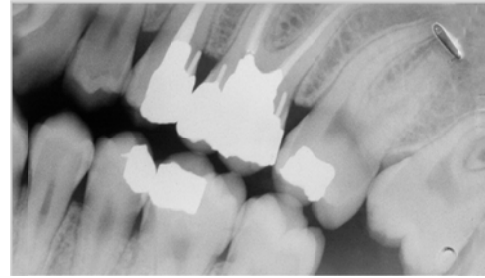
جدول ۵-۱: نتایج حاصل از مقایسه دو روش

	SPIHT	MAXSHIFT
Total CR(Coefficient)	۲۴.۳۵۰	۲۵.۱۹۸۵
Total CR(%)	۹۵.۸۹۴۴	۹۶.۰۳۱۵
Total bpp	۰.۳۲۸۴	۰.۳۱۷۵
ROI bpp	۴.۵۷۷۶	۴.۵۷۷۶
BG bpp	۰.۰۲۷۱	۰.۰۲۷۱
ROI mse	۰.۰۷۳۳	۰.۷۶۳۲
ROI psnr	۵۹.۴۸۰۵	۴۹.۳۰۴۳
BG mse	۶۷.۷۸۷۷	۴۱.۳۴۴۸
BG psnr	۲۹.۸۱۹۳	۳۱.۹۶۶۶
Time(s)	۳۵.۲۴۶۷	۶۹.۹۰۷۹

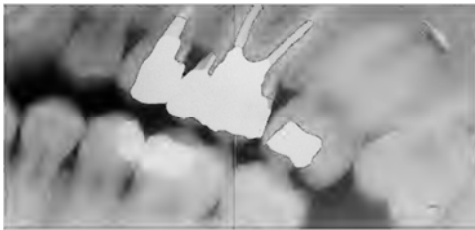
در نگاه اول در دو روش ذکر شده، تصاویر خروجی این‌گونه مشاهده می‌شود که تصویر حاصل از ماکزیمم شیفت بسیار با کیفیت‌تر است، اما در واقع چنین نیست. در روند فشرده‌سازی ناحیه مطلوب نکته حائز اهمیت این است که این ناحیه دارای کیفیت قابل قبولی باشد و اطلاعات آن حتی المقدور از بین نرود، اما در روش ماکزیمم‌شیفت، ناحیه مطلوب در اطراف مرزهایش دارای محو‌شدگی و نویز است.



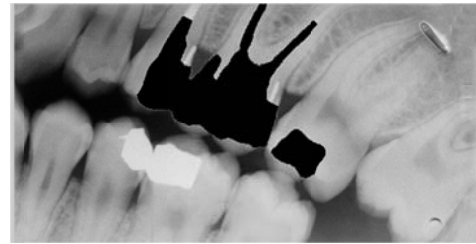
ب



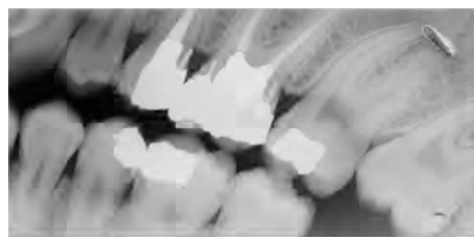
الف



ت



پ



ث

شکل ۵-۲: الف) تصویر ورودی ب) ناحیه ROI پ) ناحیه BG ت) تصویر بازسازی SPIHT ث) تصویر بازسازی

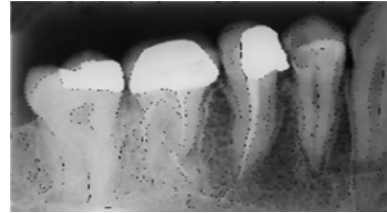
Maxshift

جدول ۵-۲ نتایج مقایسه دو روش فشرده‌سازی

	Spiht	Maxshift
Total CR	۱۷.۰۹۸۴	۱۷.۴۴۷۸
Total CR(%)	۹۴.۱۵۱۵	۹۴.۲۶۸۶
Total bpp	۰.۴۶۷۹	۰.۴۵۸۵
ROI bpp	۴.۷۰۰۷	۴.۷۰۰۷
BG bpp	۰.۰۰۳۲	۰.۰۰۳۲
ROI mse	۰.۱۶۱۱	۱۰.۴۳۵۹
ROI psnr	۵۶.۰۶۰۱	۳۷.۹۴۵۵
BG mse	۶۲۲.۸۹۵۵	۱۴۶.۱۹۷۸
BG psnr	۲۰.۱۸۶۷	۲۶.۴۸۱۸
Time(s)	۲۱.۲۳۱۹	۵۷.۷۴۳۳

برای مقایسه این دو روش بیت یکسانی به هر کدام تعلق گرفت و مطابق جدول عمل فشرده‌سازی انجام شد. همان‌طور که ملاحظه می‌شود و مشهود است، هر کدام در یک قسمت بهتر عمل کرده‌اند. به صورت کلی علی‌رغم اینکه در Maxshift فشرده‌سازی بیش‌تر صورت گرفته‌است، نتیجه مطلوبی حاصل نمی‌شود. زیرا در کنار نواحی مطلوب که با رنگ سفید از دیگر نواحی تمایز یافته، اعوجاج و نویز حاکم است.

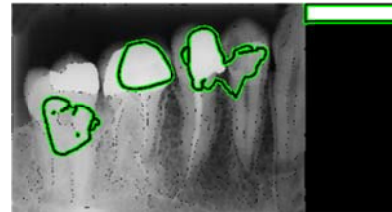
۵-۲ نحوه بدست آوردن ROI به کمک مدل مار



الف



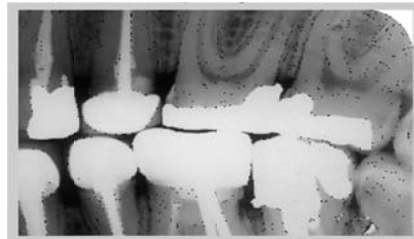
ب



پ



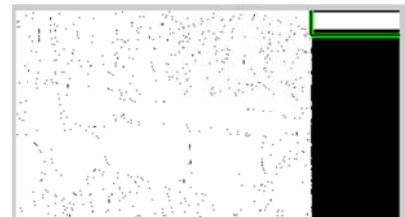
ث



الف



ب



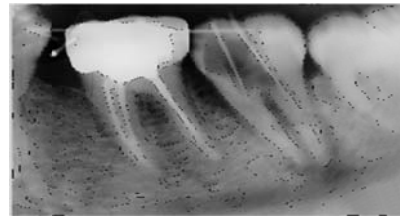
پ



ث



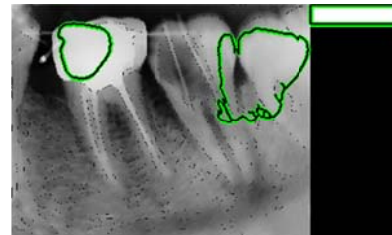
ب



الف



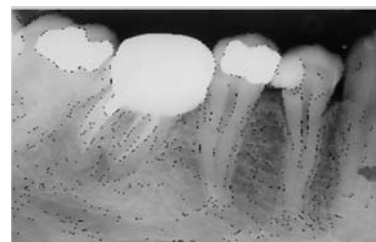
ت



پ



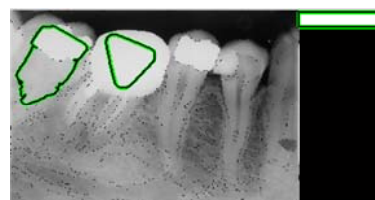
ب



الف



ت



پ

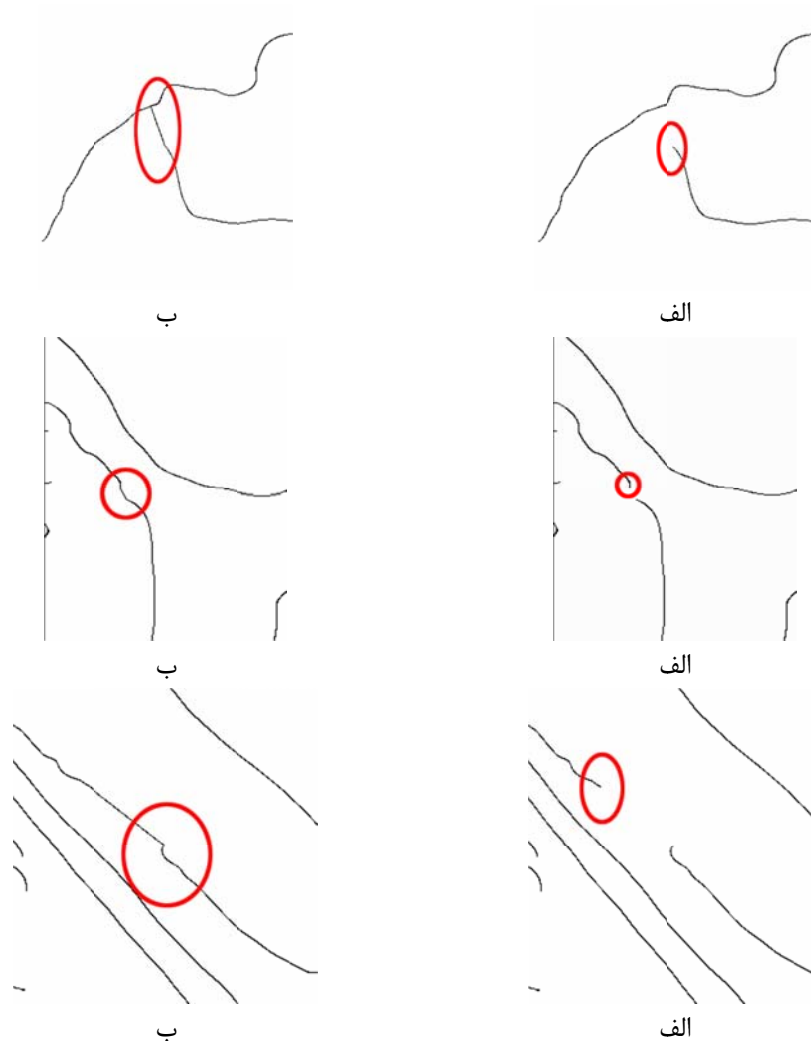
شکل ۳-۵: الف) تصویر ورودی (ب) ناحیه انتخاب شده توسط کاربر (پ) فرآیند اجرای کد (ت) ناحیه مطلوب نهایی

در این تصاویر تعداد دفعات گسترش مار بین ۵۰۰ تا ۷۰۰ انتخاب شد، برای ارزیابی بهتر مدل مار، تمرکز بر روی قسمت‌هایی از تصویر که مد نظر پزشک نیست (ناحیه غیر پرشدگی) گذاشته شده- است. انتظار می‌رفت که مدل مار قابلیت تشخیص این نواحی از نواحی ROI را دارا باشد که اینگونه

نشود. همان طور که در تصاویر مشهود می باشد این نواحی در خروجی نیز باقی مانده است. نکته دیگر که اگر تعداد تکرار دفعات انجام شده مدل مار را کم تر در نظر می گرفتیم مثلا ۵۵۰، در بعضی از عکس ها موجب این می شد که ناحیه مطلوب بصورت ناقص در نظر گرفته شده است.

۳-۵ پیوسته کردن مرزهای گسسته

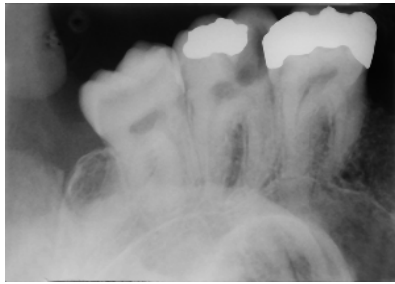
عملگر کنی قادر به ردیابی کامل مسیر نبود، پس باید برای ردیابی کامل مرزها، جهت و راستا مشخص می شد، که بصورت اشتباه به مرز دیگر یا حتی به خودش متصل نشود (توضیحات کامل در فصل ۳ توضیح داده شد). این تصاویر نحوه اتصال مرزها را در طی فرآیند اجرای کد نشان می دهد.



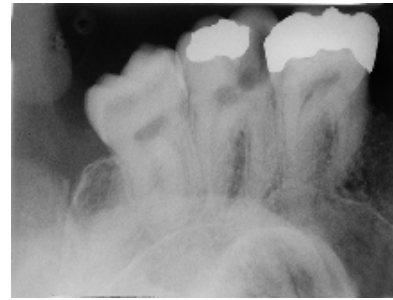
شکل ۴-۵: الف) مرز گسسته ب) مرز پیوسته

۵-۴ نحوه بدست آوردن ROI به کمک پیوسته کردن مرز

قدم نخست جداسازی ناحیه مطلوب دندان از سایر نواحی است که با رنگ سفید نمایش داده می‌شود. همانگونه که در فصل ۳ به این موضوع پرداخته شد، انجام این کار مستلزم پیوسته کردن مرزها و هم‌چنین به دست آوردن راستا و جهت مناسب است. با اعمال این ایده بر روی تمام عکس‌ها تنها در حدود ۶۵٪ جواب مطلوب (کاملاً صحیح) بدست آمد و در مابقی موارد یا جواب صحیح حاصل نشد و یا قسمتی از ROI مشخص نمی‌شد. برای مثال در شکل ۵-۵ در قسمت ب یکی از نواحی ROI انتخاب نشده، با وجود اینکه این ناحیه جز ناحیه مطلوب نیز بوده است. پس ناحیه مطلوب در شکل ۵-۵ در قسمت ت و چ قابل تشخیص و واضح است.



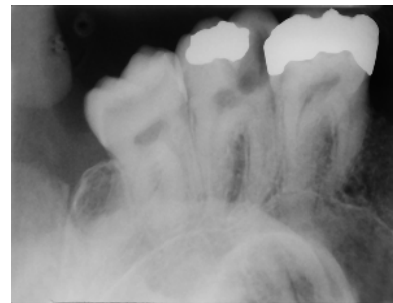
الف



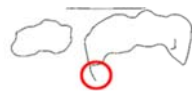
ب



پ



ت



ث



ج



ث

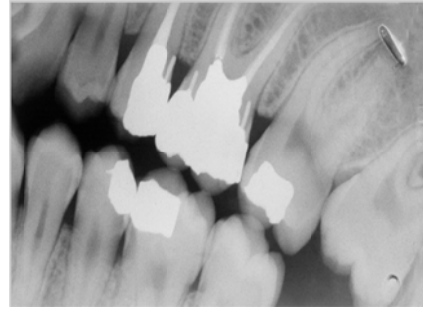


ح

شکل ۵-۵: الف) عکس اصلی ب) فیلتر میانه ۱۵*۱۵ پ) فیلتر میانه ۵۰*۵۰ ت) خروجی لبه‌یابی
 الگوریتم canny بر روی عکس ب. ث) خروجی باینری ناحیه تخمین پرشدگی ج) با ارزش‌ترین لبه‌ها در نزدیکی
 محدوده‌ی پرشدگی ث. چ) لبه‌ها از شکل یک منحنی باز به منحنی‌های بسته ح) ناحیه مطلوب نهایی به کمک
 آستانه‌گیری



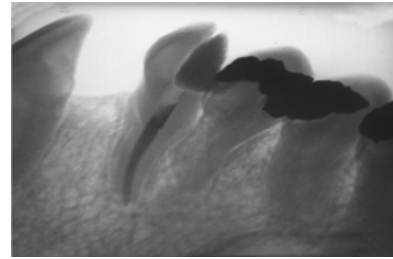
ب



الف



ب



الف



ب



الف

شکل ۵-۶: الف) تصویر ورودی ب) ناحیه مطلوب

۵-۵ پیشنهادات برای کارهای آینده

(۱) انجام تحقیق در مورد سایر عوارض دندان، نظیر تشخیص پوسیدگی، بیماری‌های لثه

و ...

(۲) بکارگیری روش‌های مبتنی بر فرهنگ داده نظیر VQ برای کد کردن ناحیه ROI و

مقایسه کیفیت و سرعت و نرخ فشرده‌سازی با روش پیشنهادی.

(۳) بهبود دقت تشخیص ROI با بکارگیری ویژگی‌های ترکیبی لبه و بافت دندان.

مراجع

- [1] Wallace G.K., "The JPEG Still Picture Compression Standard", *Communications ACM*, vol. 34, pp. 30-44, 1991.
- [2] Christopoulos C., Skodras A., and Ebrahimi T., "The JPEG2000 Still Image Coding System: An Overview", *IEEE Trans Consum Electron*, vol. 46, pp. 1103– 27, 2000 .
- [3] Elad M., Goldenberg R., and Kimmel R., "Low Bit-Rate Compression of Facial Images", *IEEE Trans Image Process*, vol. 16, pp. 2379–83, 2007.
- [4] Bryt O. and Elad M., "Compression of Facial Images Using the K-SVD Algorithm", *J Vis Commun Image R*, vol. 19, pp. 3445–62, 2008.
- [5] Bryt O. and Elad M., "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.
- [6] 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.
- [7] Taubman D., "High Performance Scalable Image Compression with EBCOT", *IEEE Trans. Image Processing*, vol. 9, pp. 1158 –1170, 2000.
- [8] Somasundaram K. and Palaniappan N., "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. 589-594, 2011.
- [9] Said A. and Pearlman WA., "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.

- [10] Mallat S., "A Theory of Multiresolution Signal Decomposition: The Wavelet Representation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11:7, pp. 674-693, 1989.
- [11] Shapiro J.M., "Embedded Image Coding Using Zero Trees of Wavelet Coefficients", *IEEE Trans Signal Process*, vol. 41, pp. 3445-62, 1993.
- [12] Usevitch B.E., "A Tutorial on Modern Lossy Wavelet Image Compression: Foundation of JPEG2000", *IEEE Signal Process*, vol. 18:5, pp. 22-35, 2001.
- [13] Crochiere R.E., Weber S.A., and Flanagan J.L., "Digital Coding of Speech in Sub Bands", *Bell Syst. Tech. J.*, vol. 55, pp. 1069-1085, 1967.
- [14] Daubechies I., "The Wavelet Transform, Time Frequency Localization and Signal Analysis", *IEEE Trans. Inf. Theory*, vol. 36:5, pp. 961-1005, 1990.
- [15] Burrus C.S., Gopinath R.A, and Guo H. , *Introduction to Wavelets and Wavelet Transform.*: Prentice Hall, 1999.
- [16] Strang G. and Nguyen T., *Wavelets and Filter Banks*, 2nd ed.: Wellesley Cambridge Press, 1997
- [17] Vetterli M. and Kovacevic J., *Wavelets and Subband Coding.*: Prentice Hall, 1999.
- [18] Egger O., Fleury P., and Ebrahimi T., "Shape Adaptive Wavelet Transform for Zerotree Coding", *Proc. European Workshop Image Anal. Coding*, vol. 1, pp. 201- 208, 1996.
- [19] Kim J.H., Lee J.Y., Kang E.S., and Ko S.J., "Region-Based Wavelet Transform for Image Compression", *IEEE Trans on Circuits and Systems—II: Analog and Digital Signal Processing*, vol. 45, no. 8, 1998.
- [20] O. Egger, T. Ebrahimi, and M. Kunt, "Arbitrarily-Shaped Wavelet Packets for Zerotree Coding", *Proc. ICASSP*, vol. IV, pp. 2335-8, 1996.
- [21] Aono T., Ito N., Katata H., and Kusao H., "Object Based Wavelet Transform Tool for Video", *ISO/IEC JTC1/SC29/WG11, MPEG-4 Proposal*, 1995.

[22] Salomon D., *Data compression – the complete reference*, 2nd ed. Verlag: Springer, 2000.

[21] Aono T., Ito N., Katata H., and Kusao H., "Object Based Wavelet Transform Tool for Video", *ISO/IEC JTC1/SC29/WG11, MPEG-4 Proposal*, 1995.

[22] Salomon D., *Data compression – the complete reference*, 2nd ed. Verlag: Springer, 2000.

[23] Mohsen Akrami, "Face Images Compression Based On Region Of Interest" *Thesis*, winter 2013.

[24] R. Loganathan, and Y.S. Kumaraswamy, "An Improved Active Contour Medical Image Compression Technique with Lossless Region of Interest," 3rd International Conference on Trendz in Information Sciences and Computing (TISC), Chennai, India, pp. 128-132. Dec. 2011

[25] H. Chen, A.K. Jan, "Tooth Contour Extraction for Matching Dental Radiographs," *Proceeding of the 1/th International Conference on Pattern Recognition*, vol .3, pp.522-525. Aug. 2004

Abstract

Digital medical images like X-Ray are widely used in diagnosis. One of the considerable advantages of digital medical images for patients is ease of storing and transmission of them. Obviously, less volume of images cause to faster compression and transmission. Furthermore, they need less space for storing. The most significant issue is attempting to preserve quality of important features of tooth adequately, especially in low bit-rates.

In this thesis two ROI-based methods are proposed in a progressive manner to compress dental images in low bit-rates. Wavelet transform and SPIHT encoding are the basis of both presented methods. In both methods, at first, region of interest should be specified, to do this we investigated three methods and finally chose and modified one of them. These three methods are: wave propagation method, active contour method and joining up boundaries method.

Its performance in finding filled regions has been improved by applying median filter to eliminate noise and perform some changes in tracing method on detected edges by canny operator, so that the proposed method is able to join up boundaries and determine regions of interest with adequate accuracy. Finally third method was chose, and used to compress. In first method, region of interest and background were encoding separately but in maxshift method they were processed together. In maxshift method compression rate is higher but least squares error in ROI in this method is higher than first method. Processing speed in SPIHT method is higher, because in encoding step of maxshift, the bit stream is generated several times so it cause to low speed in maxshift method.

keywords: wavelet transform, tooth image, image compression. SPIHT encoding, region of interest.



Shahrood University of Technology

Faculty of Electrical and Robatic

Dental Image Compression

Hoda Sharifi

Supervisor:

Dr. Omidreza Marouzi

Dr. Hadi Grailu

Date: winter 2014