





دانشکده مهندسی برق و رباتیک - گروه الکترونیک

پایان نامه کارشناسی ارشد مهندسی برق - الکترونیک

پایان نامه

افزایش وضوح تصاویر ویدیویی بدون تخمین دقیق حرکت

محمود امیری

استاد راهنما

دکتر علیرضا احمدی فرد

استاد مشاور

دکتر وحید ابوالقاسمی

بهمن ۱۳۹۵

تقدیم بہ خانوادہ می عزیزم

بہ خاطر محبت ما، پشتیبانی ما و فداکاری ما نشان

تقدیر و تشکر

این پایان‌نامه حاصل دو و نیم سال همکاری با چند تن از اساتید خوب دانشگاه صنعتی شاهرود می‌باشد. این همکاری صادقانه و پشتیبانی آن‌ها بود که به من در رسیدن به این مرحله کمک کرد.

در درجه نخست، می‌خواهم از استاد راهنمایم، جناب آقای دکتر علیرضا احمدی فرد که یک الگو و معلم بی‌نظیر است تشکر نمایم. این امتیاز بسیار بزرگی برای بنده بود که بتوانم تحت راهنمایی‌های ایشان کار کنم. همچنین لازم است از استاد مشاورم جناب آقای وحید ابوالقاسمی که با بخشندگی زیاد ایده‌هایش را با من به اشتراک می‌گذاشت تشکر نمایم.

از تمامی اساتید دانشگاه صنعتی شاهرود به‌خصوص دکتر خسروی، دکتر مروی، دکتر گرایلو کمال تشکر را دارم.

همچنین لازم می‌دانم از پدر و مادر و همسرم بخاطر دلگرمی و پشتیبانی که از من داشتند کمال تشکر را داشته باشم.

تعهد نامه

این جانب محمود امیری دانشجوی دوره کارشناسی ارشد رشته برق دانشکده برق دانشگاه صنعتی شاهرود نویسنده پایان نامه افزایش وضوح تصاویر ویدیویی بدون تخمین دقیق حرکت تحت راهنمایی دکتر علیرضا احمدی فرد متعهد می شوم.

- تحقیقات در این پایان نامه توسط این جانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش های 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

امضای 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.

چکیده:

فرا تفکیک پذیری فنی در پردازش تصویر است که سعی در بازسازی تصویری با وضوح بالا (HR^1) از روی تصویر یا تصاویر وضوح پایین (LR^2) ورودی دارد. این پایان نامه بر روی فرا تفکیک پذیری چند فریمی تمرکز دارد. روش های مبتنی بر درون یابی افزایش ابعاد تصویر را به وسیله ی نمونه برداری مجدد انجام می دهند. در این روش ها فقط تعداد پیکسل های تصویر افزایش می یابد و جزئیاتی که در نتیجه ی فرآیند تصویر برداری از دست رفته است، بازسازی نمی شود. روش های فرا تفکیک پذیری برای حل مشکلات روش های درونیابی توسعه داده شده اند. در این روش ها سعی بر این است که علاوه بر افزایش تعداد پیکسل های تصویر ورودی جزئیات از دست رفته طی فرآیند تصویر برداری نیز بازسازی شود. روش های فرا تفکیک پذیری کلاسیک روی ویدیوهای واقعی نتیجه ی خوبی تولید نمی کنند، زیرا این روش ها نیازمند تخمین دقیق حرکت می باشند. اخیرا روش هایی در فرا تفکیک پذیری رشته ویدیویی معرفی شده اند که می توانند بدون تخمین دقیق حرکت نتایج مؤثری تولید نمایند. روش فرا تفکیک پذیری مبتنی بر فیلتر NLM^3 تخمین دقیق حرکت را با مفهوم جدیدی به نام حرکت فازی جایجا کرده است. در این پایان نامه یک روش جدید فرا تفکیک پذیری مبتنی بر احتمالات ارائه می گردد. در روش پیشنهادی، با معرفی مفهوم ثبت احتمالاتی تصویر، حرکت فازی معرفی شده در کارهای پیشین بهبود داده شده است. روش پیشنهادی بر پایه ی الگوریتم برچسب گذاری آزاد⁴ می باشد. در روش پیشنهادی با استفاده از سایر فریم های ویدیویی می توانیم جزئیات فرکانس بالای از دست رفته ی تصویر را بازسازی کنیم. بر پایه روش ارائه شده در این پایان نامه به کاربر این امکان داده می شود که بدون پرداخت هزینه برای خرید سخت افزار

¹ High-Resolution

² Low-Resolution

³ Non-local means

⁴ Relaxation labelling

بهتر، بر مشکلات موجود در سیستم‌های تصویربرداری غلبه کند. نتایج تجربی نشان می‌دهد که روش پیشنهادی می‌تواند اثرات مصنوعی را حذف کند و لبه‌های تیزی نیز تولید کند. با انجام آزمایش بر روی تصاویر متفاوت نشان می‌دهیم که روش پیشنهادی از نقطه نظر معیار PSNR بهتر از سایر روش‌های معرفی شده در این پایان نامه عمل می‌کند.

واژه‌های کلیدی: فراتفکیک پذیری، برچسب گذاری آزاد، ثبت احتمالاتی تصویر، تصاویر ویدیویی

لیست مقالات مستخرج از پایان نامه:

- ❖ Mahmood Amiri, Alireza Ahmadifard, Vahid Abolghasemi, “A Fast Approach for Single Image Super Resolution via Dictionary Learning”, the second international conference on signal processing and intelligent systems, Amirkabir University of technology, December 2016.
- ❖ Mahmood Amiri, Alireza Ahmadifard, Vahid Abolghasemi “A Probabilistic Framework for Dense Image Registration Using Relaxation labelling”, the second international conference on signal processing and intelligent systems, Amirkabir University of technology, December 2016.

فهرست مطالب

۱ فصل اول : معرفی	۱
۲ فصل دوم: مرور روش‌های گذشته.....	۱۳
۲-۱-۱ فراتفکیک پذیری تک فریمی	۱۴
۲-۱-۲ روش‌های SR مبتنی بر پایگاه داده خارجی.....	۱۴
۲-۱-۲ استفاده از روش‌های خودیادگیری	۱۵
۲-۲ فراتفکیک پذیری چند فریمی.....	۱۶
۲-۲-۱ روش‌های بر پایه‌ی حوزه‌ی فرکانس.....	۱۶
۲-۲-۲ روش‌های بر پایه‌ی درون‌یابی.....	۱۷
۲-۲-۳ روش‌های SRR بر پایه‌ی تنظیم کننده	۱۸
۲-۲-۴ روش‌های بر پایه‌ی یادگیری.....	۲۰
۲-۲-۵ روش‌های مناسب برای بازسازی ویدیوهای واقعی.....	۲۲
۳ فصل سوم: تئوری	۲۵
۴ فصل چهارم : ثبت احتمالاتی تصویر	۳۷
۴-۱ روش پیشنهادی.....	۴۰
۴-۱-۱ استخراج چگال ویژگی با استفاده از بلاک‌های تصویر	۴۱

۴۲ پیدا کردن نقاط کاندید	۲-۱-۴
۴۳ لبه یابی	۳-۱-۴
۴۴ Relaxation labelling اعمال	۴-۱-۴
۴۸ Relaxation labelling تفسیر نتایج	۵-۱-۴
۴۹ نتایج تجربی	۲-۴
۵۹ فصل پنجم : فراتفکیک پذیری احتمالاتی	۵
۶۴ روش پیشنهادی	۱-۵
۶۵ ثبت احتمالاتی تصویر	۱-۱-۵
۶۷ Relaxation labelling اعمال	۲-۱-۵
۶۹ بازسازی تصویر وضوح بالا	۳-۱-۵
۷۰ نتایج تجربی	۲-۵
۸۱ نتیجه گیری	۶
۸۱ پیشنهادهایی برای کارهای آتی:	
۸۲ مراجع	

فهرست شکل‌ها

- شکل ۱-۱: دیاگرام ارائه شده از تشکیل تصویر و بازسازی چند فریمی در یک سیستم تصویربرداری دیجیتال. ۷.....
- شکل ۲-۱: نمایش ساده‌ای از مسئله‌ی فراتفکیک‌پذیری مبتنی بر حرکت. ۸.....
- شکل ۳-۱: (الف) یک تصویر واقعی، (ب) نتیجه‌ی درون‌یابی و (ج) نتیجه‌ی اجرای الگوریتم فراتفکیک‌پذیری. ۹.....
- شکل ۱-۴: نقشه‌ی اختلاف تصویر Venus. ۵۱.....
- شکل ۲-۴: نقشه‌ی اختلاف تصویر Tsukuba. ۵۲.....
- شکل ۳-۴: نقشه‌ی اختلاف تصویر Teddy. ۵۳.....
- شکل ۴-۴: نقشه‌ی اختلاف تصویر Cones. ۵۴.....
- شکل ۵-۴: نتایج بازسازی رشته‌ی بزرگ‌نمایی. ۵۵.....
- شکل ۶-۴: نتایج بازسازی رشته‌ی جابجایی. ۵۶.....
- شکل ۷-۴: نتایج بازسازی رشته‌ی چرخش. ۵۷.....
- شکل ۸-۴: نتایج بازسازی رشته‌ی از شکل انداختن. ۵۸.....
- شکل ۱-۵: مقایسه تابع وزن دهی روش پیشنهادی و روش فراتفکیک‌پذیری احتمالاتی [۴۳]. ۶۲.....
- شکل ۲-۵: مقایسه PSNR فریم به فریم رشته‌های ویدیویی تست. ۷۳.....
- شکل ۳-۵: مقایسه بصری فریم اول رشته‌ی ویدیویی Foreman. ۷۴.....
- شکل ۴-۵: مقایسه بصری فریم هشتم رشته‌ی ویدیویی Foreman. ۷۵.....
- شکل ۵-۵: مقایسه بصری فریم سیزدهم رشته‌ی ویدیویی Foreman. ۷۶.....
- شکل ۶-۵: مقایسه بصری فریم بیست و دوم رشته‌ی ویدیویی Foreman. ۷۷.....

شکل ۵-۷: مقایسه بصری فریم هشتم رشته‌ی ویدیویی Miss-America ۷۸

شکل ۵-۸: مقایسه بصری فریم بیستم رشته‌ی ویدیویی Miss-America ۷۹

شکل ۵-۹: مقایسه بصری فریم اول رشته‌ی ویدیویی Suzie ۸۰

فهرست جدول‌ها

جدول ۴-۱: مقایسه PSNR روش پیشنهادی با سایر روش‌ها..... ۵۰

جدول ۴-۲: مقایسه خطای نقشه‌ی اختلاف به دست آمده توسط روش پیشنهادی با سایر روش‌ها..... ۵۰

جدول ۵-۱: مقایسه PSNR روش پیشنهادی با سایر روش‌ها..... ۷۲

۱ فصل اول : معرفی

وضوح بصری^۱ معیاری برای سنجش توانایی سیستم تصویربرداری (دوربین) است که جزئیات درون تصویر را اندازه‌گیری می‌کند. به عبارت دیگر وضوح تصویر کوچک‌ترین جزئیاتی است که در یک تصویر قابل مشاهده است. وضوح یک موضوع بنیادی در قضاوت کیفیت سیستم پردازش تصویر می‌باشد. در ساده‌ترین حالت وضوح تصویر به عنوان کوچک‌ترین جزئیات قابل اندازه‌گیری در نمایش بصری تعریف می‌شود. در حوزه پردازش تصاویر دیجیتال و ویدیو مفهوم وضوح به یکی از سه تعریف ذیل مربوط می‌شود [۱]:

وضوح فضایی: هر تصویر از کوچک‌ترین عنصر به نام پیکسل ساخته شده است. وضوح فضایی به تعداد پیکسل‌ها در یک تصویر مربوط می‌شود و با معیار پیکسل بر طول اندازه‌گیری می‌شود. هرچه وضوح فضایی بالاتر باشد تعداد پیکسل‌ها در تصویر بیشتر خواهد بود. وضوح فضایی بالاتر درک بالاتری از جزئیات در یک تصویر را به کاربر می‌دهد.

وضوح روشنایی: به آن وضوح سطح خاکستری نیز گفته می‌شود. وضوح روشنایی به تعداد سطوح روشنایی یا سطوح خاکستری برای ارائه یک پیکسل گفته می‌شود. وضوح روشنایی با تعداد سطوح کوانتیزاسیون افزایش پیدا می‌کند. یک تصویر خاکستری معمولاً به ۲۵۶ سطح کوانتیزه می‌شود و هر سطح با ۸ بیت بیان می‌شود. برای یک تصویر رنگی هر کانال رنگ با ۸ بیت نمایش داده می‌شود که ۲۴ بیت برای ارائه هر سطح رنگ نیاز است. تعداد سطوح کوانتیزاسیون به نرخ نمونه‌برداری فضایی مربوط می‌شود.

وضوح زمانی: این وضوح نرخ فریم‌ها یا تعداد فریم‌های یک‌رشته ویدئویی را در ثانیه

¹ Optical Resolution

نشان می‌دهد. نرخ معمول برای یک ویدیو در حدود ۲۵ فریم در ثانیه می‌باشد [۱].

در این پایان‌نامه واژه افزایش وضوح به وضوح فضایی مربوط می‌شود. همان‌طور که قبلاً نیز بیان شد وضوح فضایی به‌طور ویژه تعداد کل پیکسل‌ها در یک تصویر را نشان می‌دهد. به‌عنوان مثال برای یک تصویر دیجیتال 300×300 تعداد کل پیکسل‌ها ۹۰۰۰۰ می‌باشد که در حدود ۰.۱ مگا پیکسل است. واضح است که هر چه تعداد پیکسل‌ها بالاتر باشد جزئیات بیشتری در تصویر قابل‌نمایش است. قبل از اینکه جزئیات موجود در فرا تفکیک‌پذیری بیان شود ابتدا راه‌های سخت‌افزاری که سبب افزایش وضوح یک تصویر می‌شود مطرح می‌شود. این روش‌ها عبارت‌اند از: (۱) کاهش ابعاد پیکسل‌ها (۲) افزایش اندازه چیپ [۲]. وضوح یک تصویر توسط حس‌گرهای موجود در دستگاه تصویربرداری محدود می‌شود. حس‌گرهای موجود در دوربین‌ها در یک آرایه‌ی دوبعدی به‌صورت مرتب چیده شده‌اند. ابعاد پیکسل‌ها و تعداد حس‌گرها در واحد سطح، وضوح فضایی یک وسیله تصویربرداری را مشخص می‌کنند. از طرفی محدودیت‌های سخت‌افزاری برای کاهش ابعاد حس‌گرها، وضوح فضایی را محدود می‌کند و با بالا رفتن تعداد حس‌گرها قیمت وسیله تصویربرداری نیز بالاتر می‌رود. راه‌حل دیگر برای افزایش وضوح، استفاده از روش‌های پردازش تصویر می‌باشد. این روش‌ها به‌طور ویژه به روش‌های فراتفکیک‌پذیری مربوط می‌شوند. فراتفکیک‌پذیری به فرآیندی گفته می‌شود که سعی در بازسازی تصویری وضوح‌بالا از یک یا چند تصویر وضوح‌پایین دارد. به عبارتی در فراتفکیک‌پذیری تلاش می‌شود اطلاعات فرکانس بالایی که در طول فرآیند تصویربرداری از بین رفته‌اند، بازیابی شوند. روش‌های نرم‌افزاری از لحاظ اقتصادی مقرون‌به‌صرفه می‌باشند و امکان تولید تصویری با وضوح‌بالتر را فراهم می‌آورند. به عبارتی در این روش‌ها می‌توانیم از محدوده‌ی توانایی سیستم تصویربرداری بالاتر رویم. این روش در بسیاری از زمینه‌ها همانند تصویربرداری ماهواره‌ای، هوایی، پردازش تصاویر پزشکی، بهبود تصاویر متن، بهبود تصاویر و ویدیو، خواندن پلاک خودرو، تشخیص افراد از روی عنبیه [3] کاربرد دارد.

مسئله فراتفکیک‌پذیری را نباید با روش‌های مشابه همانند درون‌یابی^۱ و بازیابی^۲ اشتباه گرفت. در مقایسه با روش‌های بهبود وضوح روش فراتفکیک‌پذیری نه‌تنها کیفیت تصویر را افزایش می‌دهد بلکه باعث افزایش وضوح فضایی یک تصویر نیز می‌شود. در روش‌های درون‌یابی برخلاف روش‌های فراتفکیک‌پذیری، جزئیات فرکانس بالا بازیابی نمی‌شوند [۴]. همچنین در روش‌های بازیابی ابعاد تصویر ورودی و خروجی یکسان است و تنها کیفیت تصویر خروجی افزایش پیدا می‌کند. در فراتفکیک‌پذیری علاوه بر اینکه کیفیت تصویر خروجی افزایش پیدا خواهد کرد ابعاد آن نیز نسبت به تصویر ورودی افزایش پیدا خواهد کرد.

معمولاً تصویری که به دست کاربر می‌رسد حالت رضایت بخشی از جزئیات صحنه موردنظر را به دلیل کامل نبودن سیستم و شرایط تصویربرداری نشان نمی‌دهد. این تصویر یک نسخه تنزل یافته از صحنه طبیعی است. این خرابی‌ها به‌واسطه‌ی فاکتورهای مثل ماتی^۳، نویز و روی هم افتادگی فرکانسی^۴ اتفاق می‌افتد. چنین خرابی‌هایی در یک سیستم تصویربرداری به علت مشکلاتی که در زیر مطرح می‌شود به وجود می‌آید.

۱- حرکت^۵ میان حس‌گر دوربین و شیء

۲- لنز و اپتیک دوربین

۳- اتمسفر

۴- کم بودن نرخ نمونه‌برداری

فاکتوری شبیه حرکت نسبی دوربین و صحنه، فوکوس نادقیق، توربولنس اتمسفر، تابع نقطه

¹ Interpolation method

² Restoration

³ Blur

⁴ Aliasing

⁵ Motion

گستر^۱، می‌توانند خرابی‌هایی در تصویر ایجاد کنند که به آن ماتی در طول فرآیند تصویربرداری گفته می‌شود. حذف اثر ماتی^۲ یکی از روش‌های بهبود کیفیت تصویر است. اگر شرایط تصویربرداری شناخته‌شده باشد حذف ماتی آسان‌تر و با دقت بالاتری انجام می‌گیرد.

فاکتور دیگری که در افزایش وضوح تصویر تأثیر دارد نرخ نمونه‌برداری می‌باشد. بر طبق تئوری نمونه‌برداری نایکوئیست [۵] حداقل نرخ نمونه‌برداری باید دو برابر بالاترین فرکانس موجود باشد. اگر فرکانس نمونه‌برداری کمتر از دو برابر فرکانس موجود باشد، تمام مؤلفه‌های فرکانسی بزرگ‌تر از نصف فرکانس نمونه‌برداری، به‌عنوان فرکانس‌های پایین در سیگنال بازسازی در نظر گرفته می‌شوند. این مسئله به‌عنوان زیرنمونه‌برداری^۳ شناخته می‌شود که در بسیاری از حس‌گرهای تصویربرداری اتفاق می‌افتد. به دلیل پدیده زیرنمونه‌برداری، مؤلفه‌ی فرکانس بالا با مؤلفه‌های فرکانس پایین همپوشانی خواهد داشت و باعث تنزل وضوح تصویر خروجی خواهد شد. به این پدیده روی هم افتادگی فرکانسی گفته می‌شود. روی هم افتادگی فرکانسی باعث از بین رفتن اطلاعات جزئی تصویر و تولید اثرات مصنوعی در بازسازی سیگنال می‌شود. برای کاهش این اثرات مصنوعی باید از روش‌های حذف اثر روی هم افتادگی فرکانسی^۴ استفاده شود.

در سالیان اخیر نیاز به الگوریتمی که تصویر وضوح‌بالا را از چندین تصویر وضوح پایین بازسازی کند و بتواند کاستی‌های سیستم تصویربرداری را رفع کند افزایش یافته است. تصاویر دیجیتال امروزه در همه‌جا حضور دارند. با داشتن چنین تصویرهایی به فرم دیجیتال این امکان به کاربر داده می‌شود که بتواند آن‌ها را با توجه به نیاز خود تغییر دهد. پردازش تصاویر دیجیتال در بهبود ویژگی‌ها و استخراج

¹ Point spread function

² De-bluring

³ Under-sampling

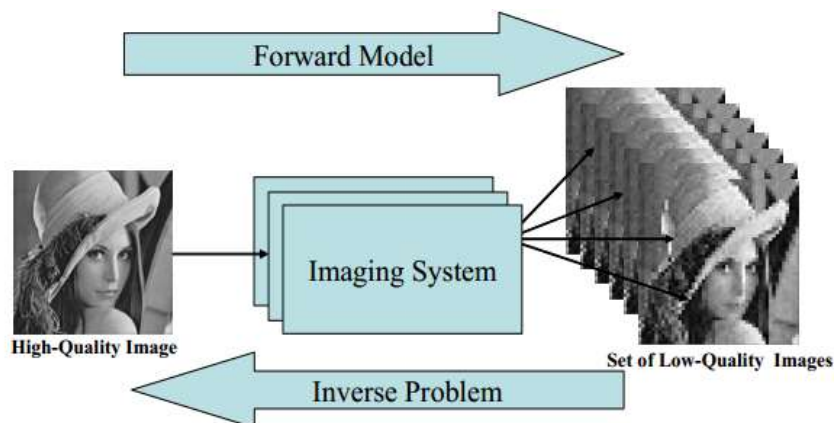
⁴ Anti-aliasing

اطلاعات مفید از تصاویر به کاربر کمک می‌کند. پردازش تصویر در چند دهه اخیر رشد قابل توجهی در زمینه‌ی سرعت و کاهش هزینه داشته است. همچنین در ساخت حس‌گرهای دوربین نیز پیشرفت‌های زیادی به وجود آمده است و دوربین‌های دیجیتال باکیفیت بالایی تولید می‌شوند. گرچه دوربین‌های وضوح‌بالا در دسترس هستند اما بسیاری از کاربردهای بینایی ماشین همانند تصویربرداری ماهواره‌ای، شناسایی هدف^۱، تصاویر پزشکی و بسیاری دیگر هنوز به داشتن تصویری با وضوح بالاتر نیازمند هستند. برآورده کردن نیازهای مطرح‌شده برای داشتن تصویری با وضوح‌بالا سبب به وجود آمدن روش‌های پردازش تصویر شدند. درونی‌یابی تصویر یک موضوع جذاب در زمینه‌ی پردازش تصویر است. در سیستم‌های دیجیتال یک سیگنال توسط تعداد محدودی از نمونه‌ها ارائه می‌شود. درونی‌یابی چنین سیگنال‌هایی توسط نمونه‌هایی که مشخص هستند، انجام می‌گیرد. چنین مسئله‌ای یک مسئله بدحالت است چون سیگنال‌های بسیاری وجود دارند که دارای نمونه‌های یکسانی هستند. در بیشتر موارد فرض بر این است که سیگنال‌ها باند محدود هستند. درونی‌یابی در صورتی امکان‌پذیر است که فرآیند نمونه‌برداری داده‌ها از شرط نایکوئیست پیروی کند. ممکن است این شرط همیشه برقرار نباشد. روش‌های فراتفکیک‌پذیری برای افزایش وضوح فضایی استفاده می‌شوند. فراتفکیک‌پذیری فرآیندی است که سبب افزایش وضوح یک تصویر می‌شود. در واقع مقادیر نمونه‌های جدید توسط نمونه‌هایی از سیگنال که در دسترس هستند تخمین زده می‌شوند، به طوری که سیگنال جدید جزئیات بیشتر و اطلاعات بهتری را فراهم می‌آورد.

این روش‌ها با ترکیب چندین عکس گرفته شده توسط تراشه‌ها و تجهیزات نوری ارزان قیمت می‌توانند یک تصویر وضوح‌بالا تولید نمایند. بلوک دیاگرام ارائه شده برای این مسئله در شکل ۱-۱ نشان داده شده است. حرکت نسبی بین دوربین و صحنه، نویز سیستم تصویربرداری (مثلاً CCD) و خصوصیات

^۱Target detection

لنز نوری در طول زمان تغییر خواهند کرد. بنابراین هرکدام از تصاویر گرفته شده مقداری از اطلاعات صحنه‌ی اصلی را ضبط خواهند کرد.



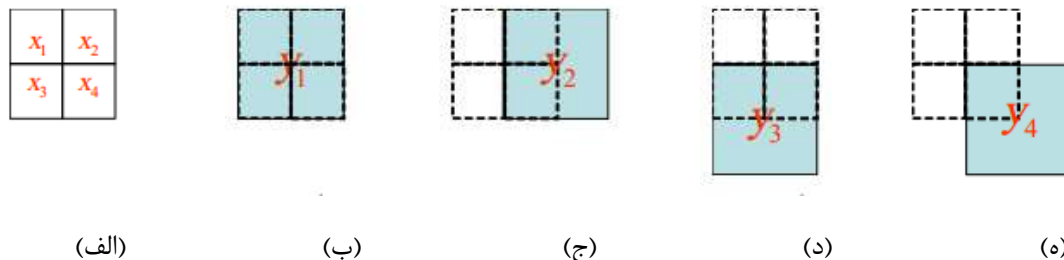
شکل ۱-۱: دیاگرام ارائه شده از تشکیل تصویر و بازسازی چند فریمی در یک سیستم تصویربرداری دیجیتال.

تفکیک بسیاری از سیستم‌های تصویربرداری به تجهیزات نوری آن‌ها محدود می‌باشد. کوچک‌ترین تفکیک قابل دسترس این سیستم‌ها از قانون ریلی [۶] تبعیت می‌کند، که وابسته به طول موج نور و قطر دریچه‌ی دوربین می‌باشد. سیستم‌های تصویربرداری دیجیتال معمولاً توسط چگالی سنسورها (مثلاً CCD) محدود می‌شوند [۷]. در این پایان‌نامه فقط مشکلات تفکیک‌پذیری که مربوط به چگالی حس‌گرها می‌باشد (نمونه‌برداری زیر فرکانس نایکوئیست) بررسی می‌شود.

ایده‌ی اصلی فراتفکیک‌پذیری، ترکیب رشته‌ای از تصاویر مات و نویزی و وضوح پایین (LR) برای ساخت تصاویر وضوح بالا (HR) می‌باشد. تصویر وضوح‌بالای ساخته شده محتویات فرکانس بالای بیشتری دارد و نویز آن کاهش یافته است.

کارهای اولیه‌ای که روی موضوع فراتفکیک‌پذیری انجام شده است نشان می‌دهد که اثر روی هم

افتادگی فرکانسی^۱ ما را قادر می‌سازد تا تصویر وضوح بالایی از تصاویر ترکیب شده^۲ به دست آوریم، مشروط بر اینکه یک حرکت زیر پیکسلی نسبی بین تصاویر زیر نمونه‌برداری شده^۳ ورودی وجود داشته باشد [۸]. یک آزمایش بسیار ساده‌ی فراتفکیک‌پذیری در شکل ۱-۲ نشان داده شده است.



شکل ۱-۲: نمایش ساده‌ای از مسئله‌ی فراتفکیک‌پذیری مبتنی بر حرکت. (الف) یک تصویر وضوح بالا که از ۴ پیکسل ساخته شده است. (ب) تا (ه) تصاویر وضوح پایین شامل یک پیکسل که دارای حرکت زیر پیکسلی می‌باشد.

شکل ۱-۲ اصول الگوریتم‌های فراتفکیک‌پذیری مبتنی بر حرکت را نشان می‌دهد. یک صحنه‌ی وضوح بالا شامل ۴ پیکسل در شکل (الف) نشان داده شده است. یک دوربین مجازی که دارای حرکت زیر پیکسلی می‌باشد چندین تصویر را از این صحنه شکار می‌کند. شکل (ب) تا (ه) تصاویر LR گرفته شده را نمایش می‌دهند. مطمئناً هیچ‌کدام از این تصاویر LR نمی‌توانند جزئیات صحنه‌ی مورد اشاره را بازگو نمایند. فرض کنید تابع نقطه گستر دوربین^۴ (PSF) مجازی یک تابع خطی شناخته شده باشد و مقدار روشنایی تمام پیکسل‌های حاشیه صفر می‌باشد. معادلات ارتباط تصاویر LR به تصویر HR به صورت زیر می‌باشد.

¹ Aliasing

² Fused

³ Under-sampled

⁴ Point spread function

$$y_1 = h_1 \cdot x_1 + h_2 \cdot x_2 + h_3 \cdot x_3 + h_4 \cdot x_4 + v_1 \quad 1-1$$

$$y_2 = 0 \cdot x_1 + h_2 \cdot x_2 + 0 \cdot x_3 + h_4 \cdot x_4 + v_2 \quad 2-1$$

$$y_3 = 0 \cdot x_1 + 0 \cdot x_2 + h_3 \cdot x_3 + h_4 \cdot x_4 + v_3 \quad 3-1$$

$$y_4 = 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + h_4 x_4 + v_4 \quad 4-1$$

که y_i ها ($i = 1,2,3,4$) تصاویر LR ضبط شده می‌باشند و x_i ها مقادیر روشنایی پیکسل‌های داخل تصویر وضوح‌بالا می‌باشند. h_i ها عناصر PSF شناخته شده‌اند و v_i ها نویز تصادفی CCD می‌باشند. در حالتی که نویز اضافه شده بسیار کوچک باشد ($v_i \cong 0$) مجموعه معادلات بالا قابل حل می‌باشند و می‌توان مقدار پیکسل وضوح‌بالا را به دست آورد. متأسفانه همان‌طور که در قسمت‌های بعدی خواهیم دید این فرض ساده کننده در شرایط واقعی مسئله را غیرقابل حل می‌کند. آزمایش موجود در شکل ۳-۱ یک مثال واقعی از فناوری فراتفکیک‌پذیری ارائه می‌نماید.



(الف)



(ب)



(ج)

شکل ۳-۱: (الف) یک تصویر واقعی، (ب) نتیجه‌ی درون‌یابی و (ج) نتیجه‌ی اجرای الگوریتم فراتفکیک‌پذیری

در این آزمایش تعداد ۲۶ تصویر با یک دوربین گرفته شده است که یکی از آن‌ها در شکل (الف)

نشان داده شده است. متأسفانه به دلیل تعداد پایین پیکسل‌های دوربین دیجیتال جزئیات این تصویر واضح نمی‌باشد. تصویر درون‌یابی شده در شکل (ب) نشان داده شده است. فراتفکیک‌پذیری به ما کمک می‌کند که جزئیات از دست رفته در فرآیند تصویربرداری را بازسازی کنیم. نتایج اعمال الگوریتم فراتفکیک‌پذیری [۹] به این تصویر در شکل (ج) نشان داده شده است. تصویر خروجی یک تصویر وضوح‌بالا می‌باشد که تعداد پیکسل‌های آن ۱۶ برابر بیشتر از تعداد پیکسل‌های هر کدام از تصاویر وضوح پایین می‌باشد (نرخ بهبود کیفیت در هر فریم ۴ می‌باشد).

فراتفکیک‌پذیری تلاش می‌کند یک تصویر وضوح‌بالا از تعدادی تصویر که به خاطر محدودیت‌های سیستم تصویربرداری افت کیفیت پیدا کرده‌اند، استخراج شود. به این نوع از مسائل، مسئله معکوس گفته می‌شود. زیرا منبع اطلاعات از داده‌های جمع‌آوری شده ساخته می‌شود. حل یک مسئله معکوس در حالت کلی نیاز به ساخت یک مدل مستقیم دارد. تا به حال معمول‌ترین مدل ارائه شده برای مسئله فراتفکیک‌پذیری به صورت معادله‌ی زیر می‌باشد.

$$Y = MX + v$$

۵-۱

که Y داده‌ی اندازه‌گیری شده است (یک یا تعدادی تصویر)، X تصویر وضوح‌بالای ناشناخته است و v نویز تصادفی سیستم تصویربرداری می‌باشد. در این فرمول‌بندی تصویر دوبعدی با اسکن تصویر به صورت یک‌بعدی تبدیل می‌شود. ماتریس M در مدل مستقیم بالا، سیستم تصویربرداری را ارائه می‌دهد که شامل فرآیندهایی می‌باشد که روی کیفیت تصویر به دست آمده اثر می‌گذارند. ساده‌ترین شکل ماتریس M ماتریس واحد می‌باشد که مسئله را به یک مسئله نویز زدایی تقلیل می‌دهد. مسائل پیچیده‌تر با در نظر گرفتن مدل پیچیده‌تری برای M به دست می‌آید. برای مثال معمولاً برای تعریف مسئله‌ی

فرا تفکیک پذیری تصاویر خاکستری یک سیستم تصویربرداری را شامل مراحل ماتی، پیچ و تاپ دادن^۱ و کاهش نمونه برداری^۲ در نظر می گیرند. با در نظر گرفتن مدل مستقیم یک راه حل ساده مسئله می تواند با تکنیک شبه معکوس به دست آید.

$$X = (M^T M)^{-1} M^T Y \quad ۶-۱$$

متأسفانه ابعاد ماتریس M آن قدر بزرگ است که ذخیره کردن $M^T M$ حتی بدون در نظر گرفتن معکوس غیر عملی می باشد. در عمل برای تخمین X یک تابع هزینه تعریف می شود و در یک فرایند بهینه سازی با روش های تکراری حل می شود. این نوع از توابع هزینه اطمینان می دهند که نتیجه نهایی به داده های ورودی وفادار می باشد. برای مثال با تعریف تابع هزینه حداقل مربعات که نرم L_2 بردار باقیمانده را حداقل می کند داریم:

$$\bar{X} = \operatorname{argmin}_X J(X) = \operatorname{argmin}_X \|Y - MX\|_2^2 \quad ۷-۱$$

برای حالتی که v نویز گوسی سفید با میانگین صفر می باشد این شیوه، تفسیری از تخمین بیشترین شباهت^۳ ML به \bar{X} می باشد [۱۰].

کاربردهای فرا تفکیک پذیری به سه دسته تقسیم می شود:

✚ کاربردهای صنعتی: ساخت دوربین دیجیتال ارزان قیمت، معاینه ی IC، ساخت

الگوریتم های فشرده سازی HDTV و وضوح بالا با نرخ بیت^۴ پایین.

^۱ Warping

^۲ Down-sampling

^۳ Maximum Likelihood

^۴ Bit rate

✚ کاربردهای علمی: ستاره‌شناسی (بهبود تصاویر گرفته شده از تلسکوپ)، زیست‌شناسی

(بهبود تصاویر گرفته شده از میکروسکوپ‌های الکترونیکی و نوری)

✚ کاربردهای امنیتی و قضایی: بهبود تصاویر گرفته شده از طریق دوربین‌های نظارتی.

در این پایان‌نامه می‌خواهیم یک چهارچوب احتمالاتی چند فریمی برای ترکیب اطلاعات چندین تصویر وضوح پایین و رسیدن به یک تصویر وضوح بالا ارائه کنیم. مسائل تئوری و پیاده‌سازی الگوریتم پیشنهادی با در نظر گرفتن کاربردهای واقعی فراتفکیک‌پذیری می‌باشد. تصاویر خروجی روش ارائه شده، دارای لبه‌های تیزتر، وضوح فضایی بالاتر و نویز کمتر نسبت به روش‌های موجود می‌باشند.

سازمان‌دهی این پایان‌نامه به شرح زیر است:

در فصل ۲ مروری بر روش‌های انجام‌گرفته در زمینه‌ی فراتفکیک‌پذیری خواهد شد و روش‌هایی که در حوزه‌ی مشابه با این کار قرار دارند معرفی می‌شوند و مزایا و معایب هرکدام از این روش‌ها شرح داده خواهد شد. در فصل ۳ مباحث نظری موردنیاز برای انجام فرآیند فراتفکیک‌پذیری احتمالاتی مطرح خواهد شد. در فصل ۴ روش پیشنهادی برای ثبت احتمالاتی تصویر مطرح می‌شود. در فصل ۵ روش پیشنهادی فراتفکیک‌پذیری احتمالاتی بیان خواهد شد. در نهایت در فصل ۶ نتایج کلی پایان‌نامه مطرح می‌شود و پیشنهادهایی برای کارهای آتی بیان خواهد شد.

۲ فصل دوم: مرور روش های گذشته

هدف اصلی بازسازی تصویر با استفاده از فراتفکیک‌پذیری ساخت یک تصویر وضوح‌بالا بر اساس یک یا مجموعه‌ای از تصاویر وضوح پایین می‌باشد. در حالت کلی روش‌های بازسازی فراتفکیک‌پذیری^۱ (SRR) به دو دسته تقسیم می‌شوند. (۱) فراتفکیک‌پذیری تک فریمی (۲) فراتفکیک‌پذیری چند فریمی. در فراتفکیک‌پذیری تک فریمی با داشتن یک تصویر وضوح پایین از صحنه سعی می‌شود تصویر وضوح‌بالایی از صحنه تخمین زده شود، در حالی که در فراتفکیک‌پذیری چند فریمی با داشتن چند تصویر وضوح پایین از یک صحنه، تصویر وضوح‌بالای نهایی ساخته می‌شود.

۱-۲ فراتفکیک‌پذیری تک فریمی

۱-۱-۲ روش‌های SR مبتنی بر پایگاه داده خارجی

این روش‌ها الگوریتم‌های مبتنی بر یادگیری یا الگوریتم‌های هالوسینه می‌باشند که ابتدا در [۱۱] معرفی شدند. در این مقاله از شبکه عصبی برای بهبود وضوح تصاویر اثرانگشت استفاده شده است. این الگوریتم‌ها دارای یک فاز آموزش هستند که در آن‌ها رابطه‌ی میان تعدادی تصاویر HR و LR از یک کلاس ویژه شامل تصاویر چهره، اثرانگشت و غیره یاد گرفته می‌شود. سپس این رابطه به‌عنوان جمله‌ی پیشین برای بازسازی تصویر وضوح‌بالا به‌کار برده می‌شود. پایگاه آموزشی الگوریتم‌های SRR مبتنی بر پایگاه داده خارجی باید قابلیت عمومیت‌گرایی خوبی داشته باشند [۱۲]. پایگاه داده خارجی بزرگ لزوماً نتایج بهتری تولید نمی‌کند. در مقابل زیاد بودن نمونه‌های نامرتبب‌نه تنها بار محاسباتی پیدا کردن بهترین نگاشت را زیاد می‌کنند بلکه باعث آشفتگی این عملیات نیز می‌شوند.

¹ Super resolution reconstruction

۲-۱-۲ استفاده از روش‌های خود یادگیری

اگر رابطه‌ی بین وصله‌های وضوح‌بالا و پایین با توجه به وصله‌های خود تصویر در مقیاس‌های متفاوتی یاد گرفته شود مشکلات مطرح‌شده در روش‌های مبتنی بر پایگاه داده خارجی برطرف می‌گردد. برای این منظور از هرم تصویر استفاده می‌شود. روش‌هایی که اخیراً در حوزه‌ی SRR تک فریمی پیشنهاد شده‌اند مبتنی بر ارائه تُنک می‌باشند. ارائه تُنک از یک سیگنال، بر پایه‌ی یک سری توابع پایه، در بسیاری از زمینه‌ها کاربرد دارد. این روش در زمینه‌هایی همچون حذف نویز^۱، اصلاح تصویر^۲، فرا تفکیک‌پذیری و طبقه‌بندی^۳ و غیره کاربرد دارد. بسیاری از داده‌های واقعی می‌توانند به صورت تُنک ارائه شوند به طوری که توابع پایه می‌توانند از طریق مجموعه‌ای از داده‌ها یاد گرفته شوند. به عنوان مثال یک تصویر می‌تواند به صورت ترکیب وزن‌دار از الگوهای پایه بیان شود. تکرار وصله‌ها در مقیاس‌های متفاوتی از یک تصویر جفت وصله‌های HR/LR را مهیا می‌کند که همان پایه و اساس روش‌های مبتنی بر یادگیری می‌باشند [۱۳]. تکرار وصله‌ها در مقیاس‌های متفاوتی از تصویر ورودی، نیاز ما را به جمع‌آوری پایگاه داده خارجی از بین می‌برد. فرض وجود وصله‌های مشابه در این مقاله بر این اساس است که اگر ابعاد وصله‌ها کوچک انتخاب شود (به عنوان مثال 5×5) تعداد وصله‌های مشابه در یک مقیاس و یا در مقیاس‌های متفاوتی از تصویر زیاد می‌شود حتی ممکن است به لحاظ بصری این ساختار تکراری قابل درک نباشد. این فرضیه از این حقیقت ناشی می‌شود که وصله‌های کوچک شامل لبه‌ها، گوشه‌ها و ... می‌باشند. ایرانی و همکاران [۱۳] این فرضیه را به صورت آماری بررسی کردند و به این نتیجه رسیدند که به طور متوسط ۹۰ درصد وصله‌ها در هر تصویر ۹ وصله مشابه یا بیشتر در همان مقیاس تصویر اصلی دارند. ۸۰ درصد از وصله‌های ورودی ۹ وصله مشابه یا بیشتر در ۰.۴۱ از ابعاد تصویر اصلی و ۷۰ درصد آن‌ها ۹ وصله مشابه

^۱ De-noising

^۲ Inpainting

^۳ Classification

در ۰.۲۶ از ابعاد تصویر ورودی دارند. همان‌طور که در این مقاله بررسی شد وصله‌های مشابه در سرتاسر تصویر پخش شده‌اند. به همین ترتیب یک هرم تصویر ساخته شد و با ترکیب SRR تک فریمی و چند فریمی محدودیت SRR مبتنی بر مثال_داشتن یک پایگاه داده خارجی_ برطرف گردید. اما این روش به لحاظ محاسباتی بسیار هزینه‌بر است.

۲-۲ فراتفکیک پذیری چند فریمی

در حالت کلی روش‌های بازسازی فراتفکیک‌پذیری چند فریمی به چهار دسته تقسیم می‌شوند. I) روش‌های حوزه‌ی فرکانس [۸, ۱۴-۱۶] II) روش‌های درون‌یابی [۱۷-۱۹] III) روش‌های برپایه‌ی تنظیم‌کننده [۲۰-۲۲] IV) روش‌های مبتنی بر یادگیری [۲۳-۲۵].

در زیر تعدادی از روش‌های موجود در هر کدام از دسته‌های بالا را مرور می‌نماییم.

۲-۲-۱ روش‌های بر پایه‌ی حوزه‌ی فرکانس

اولین روش ارائه شده در حوزه‌ی فرکانس کار Tsai و Huang [۸] می‌باشد. آن‌ها مسئله‌ی SRR را برای تصاویر بدون نویز در نظر گرفتند. در این تحقیق پیشنهاد شد که تصاویر وضوح پایین را به حوزه‌ی تبدیل فوریه گسسته (DFT) منتقل نموده و با توجه به رابطه‌ی بین ضرایب DFT تصاویر LR (که دارای روی هم افتادگی فرکانس می‌باشند) و تصویر وضوح‌بالای مجهول اطلاعات مؤلفه‌های بالای فرکانسی به دست آید. بعد از به دست آمدن تصویر وضوح‌بالا در حوزه‌ی فرکانس تصویر حوزه‌ی فضایی به دست می‌آید. Rhee و Kang [۲۶] از تبدیل گسسته کسینوس DCT برای حل مسئله‌ی SRR استفاده کردند. روش‌های SRR حوزه‌ی فرکانس علی‌رغم اینکه کاملاً قابل درک می‌باشند و محاسبات کمی نیز دارند اما نمی‌توان از آن‌ها در کاربردهای عملی استفاده کرد، زیرا آن‌ها نیازمند وجود جابجایی سراسری بین تصاویر LR و ماتریکس خطی یکسان در فضا می‌باشند [۲۷].

۲-۲-۲ روش‌های برپایه‌ی درون‌یابی

روش‌های فراتفکیک‌پذیری بر پایه‌ی درون‌یابی برای به دست آوردن تصویر وضوح‌بالا، تصاویر وضوح پایین را روی تصویر مرجع طرح‌ریزی^۱ می‌کنند، سپس اطلاعات موجود در تصاویر طرح‌ریزی شده را ترکیب می‌کنند و در نهایت عملیات ماتی‌زدایی را برای بهبود کیفیت تصویر خروجی انجام می‌دهند. روش‌های بر پایه‌ی درون‌یابی از این حقیقت که هر کدام از تصاویر LR ورودی مقداری اطلاعات اضافی از صحنه مهیا می‌کنند استفاده می‌نمایند. روش‌های SRR بر پایه‌ی درون‌یابی از مراحل زیر تشکیل می‌شوند (۱) مرحله‌ی ثبت تصویر برای روی هم قرار دادن تصاویر^۲ وضوح پایین ورودی (۲) مرحله‌ی درون‌یابی که یک تصویر وضوح‌بالا می‌سازد. (۳) مرحله‌ی مات‌زدایی برای بهبود تصویر ساخته شده در مرحله‌ی دوم

مرحله‌ی درون‌یابی نقشی کلیدی در این چهارچوب دارد. در طول دو دهه‌ی اخیر روش‌های زیادی برای اعمال درون‌یابی پیشنهاد شده‌اند، ساده‌ترین روش پیشنهاد شده الگوریتم نزدیک‌ترین همسایه می‌باشد. در این روش هر کدام از پیکسل‌های ناشناخته از مقدار روشنایی نزدیک‌ترین همسایه قرض گرفته می‌شود، اما این روش باعث به وجود آمدن تصویری با اثرات بلوکی می‌شود [۲۷]. Irani [۲۸] و همکارانش پیشنهاد کردند که سه مرحله‌ی گفته شده برای بازسازی تصویر به صورت تکراری انجام شود. روش آن‌ها مرسوم به IBP^۳ می‌باشد. در روش IBP تصویر وضوح‌بالا با طرح‌ریزی اختلاف بین تصاویر وضوح پایین و تصویر وضوح پایین شبیه‌سازی شده به دست می‌آید. به دلیل ماهیت بد حالت مسئله‌ی SRR این روش ممکن است به جواب واحدی ختم نشود. روش طرح‌ریزی روی مجموعه‌های محدب Pocs^۴ توسط Patti و Tekalp [۱۹] پیشنهاد شد. در روش آن‌ها تصویر وضوح‌بالا به طور هم‌زمان سازگار

^۱ Projecting

^۲ Aligning

^۳ Iterative back projection

^۴ Projection onto convex set

با اطلاعات تصاویر LR ورودی و مدل پیشین^۱ می‌باشد. روش‌های SRR بر پایه‌ی درون‌یابی به‌سادگی قابل پیاده‌سازی می‌باشند اما لزوماً جواب یکتایی ندارند.

۳-۲-۲ روش‌های SRR بر پایه‌ی تنظیم‌کننده

با توجه به اینکه مسئله‌ی SRR یک مسئله بد حالت می‌باشد الگوریتم‌های SRR زیادی بر پایه‌ی عبارت تنظیم‌کننده برای رسیدن به جواب مطلوب ارائه شده‌اند [۲۹-۳۱]. روش‌های SRR بر پایه‌ی عبارت تنظیم‌کننده از تنظیم‌کننده‌ها به عنوان دانش پیشین در حل مسئله‌ی SRR استفاده می‌نمایند. معروف‌ترین و محبوب‌ترین روش‌های SRR بر پایه‌ی بیزین، تخمین بیشترین شباهت [۲۰] ML و تخمین MAP^2 [۲۱, ۲۲] می‌باشد. اولین روش SRR بر پایه‌ی تخمین ML در [۲۰] پیشنهاد شد. آن‌ها تصویر وضوح‌بالای نهایی \hat{X} را با استفاده از تخمین ML به صورت زیر به دست آوردند.

$$\hat{X}^{ML} = \operatorname{argmax} p(Y|X) \quad ۱-۲$$

که Y مجموعه تصاویر LR ورودی می‌باشد.

کلید اصلی حل مسئله‌ی بالا نحوه‌ی بنیان نهادن تابع چگالی احتمال^۳ شرطی $p(Y|X)$ می‌باشد. از آنجاکه تصاویر LR به صورت جداگانه از تصویر اصلی به دست آمده‌اند تابع چگالی احتمال شرطی $p(Y|X)$ را می‌توان به صورت زیر نوشت:

$$p(Y|X) = \prod_{t=1}^T p(y^{(t)}|X) \quad ۲-۲$$

^۱ Prior model

^۲ Maximum a posterior

^۳ Probability density function

به صورتی که $y^{(t)}$ t-امین فریم از تصاویر LR می‌باشد. علاوه بر این طبق مدل تصویر گفته شده

در فصل قبل (۵-۱) تابع چگالی احتمال می‌تواند به صورت زیر بیان شود:

$$p(y^{(t)}|X) \propto \exp\left(-\frac{1}{2\sigma_t^2} \|y^{(t)} - M^{(t)}X\|_2^2\right) \quad ۳-۲$$

با قرار دادن ۳-۲ در ۲-۲ تابع چگالی احتمال به صورت زیر به دست می‌آید:

$$p(Y|X) \propto \prod_{t=1}^T \exp\left(\frac{1}{2\sigma_t^2} \|y^{(t)} - M^{(t)}X\|_2^2\right) = \quad ۴-۲$$

$$\exp\left(-\sum_{t=1}^T \frac{1}{2\sigma_t^2} \|y^{(t)} - M^{(t)}X\|_2^2\right)$$

با جایگذاری ۴-۲ در ۱-۲ تخمین تصویر وضوح بالا به صورت زیر می‌آید.

$$\hat{X}^{ML} = \operatorname{argmax}_X \exp\left(-\sum_{k=1}^{\rho} \frac{1}{2\sigma_k^2} \|y^{(k)} - M^{(k)}X\|_2^2\right) \quad ۵-۲$$

بر خلاف روش‌های SRR بر پایه‌ی ML که تنها رابطه‌ی بین تصویر وضوح بالا و تصاویر LR را در

نظر می‌گیرند، روش‌های SRR بر پایه‌ی MAP از یک دانش پیشین برای ساخت تصویر وضوح بالا استفاده

می‌نمایند. این روش‌ها یک چهارچوب انعطاف‌پذیر برای دربرداشتن هم‌زمان ارتباط تصاویر LR و تصویر

وضوح بالا و همچنین یک دانش پیشین ارائه می‌نمایند.

هدف اصلی روش‌های بر پایه‌ی تخمین MAP تخمین تصویر وضوح بالا به صورت زیر است.

$$\hat{X}^{MAP} = \operatorname{argmax} p(X|Y) \quad ۶-۲$$

با استفاده از قانون بیز تابع چگالی احتمال را به صورت زیر می‌نویسیم

$$p(X|Y) = p(Y|X)p(X)$$

۷-۲

تعیین مدل پیشین^۱ نقش اساسی در بازسازی تصویر وضوح بالا ایفا می‌کند. تابه‌حال در مقالات SRR مدل‌های پیشین بسیاری مانند MRF [۳۲] و Gaussian MRF [۳۳] ، HUBBER MRF [۳۴] و غیره معرفی شده است.

۴-۲-۲ روش‌های بر پایه‌ی یادگیری

اخیرا روش‌هایی برای حل مسئله‌ی SR بر پایه‌ی یادگیری پیشنهاد شده‌اند. در این روش‌ها اطلاعات فرکانس بالای تصویر مرجع به‌وسیله‌ی بازیابی اطلاعات فرکانس بالا از نمونه‌های آموزشی، بهبود داده می‌شود. Hertzmann [۳۵] یک روش برای به دست آوردن اطلاعات فرکانس بالای تصویر از طریق پایگاه داده آموزشی ارائه نمود. روش آن‌ها دو مرحله دارد (۱) مرحله‌ی آموزش (۲) مرحله‌ی بازسازی تصویر وضوح بالا.

مقالات اخیر در حوزه‌ی آماری تصویر پیشنهاد داد که می‌توان وصله‌های تصویر را به صورت ترکیب خطی عناصر موجود در یک دیکشنری از وصله‌ها ارائه داد [۳۶].

chen و همکارانش [۳۷] از روش خود یادگیری استفاده کردند. روش کار به این صورت است که ابتدا تصویر ورودی وضوح پایین، با روش bicubic با فاکتور مقیاس s بزرگ می‌شود و آن تصویر شبه HR (X_m) نامیده می‌شود. تصویر ورودی به تعدادی وصله‌های بدون هم‌پوشانی^۲ با سایز $B \times B$ تقسیم می‌شود. برای هر وصله در تصویر LR وصله‌ی متناظر آن در تصویر X_m با سایز $sB \times sB$ در همان مکان استخراج می‌شود. برای هر وصله در تصویر X_m در یک همسایگی مشخص، تمام وصله‌های اطراف آن استخراج می‌-

^۱ prior model

^۲ Overlap

شوند و هر وصله استخراج شده به صورت برداری ستونی در نظر گرفته شده و با کنار هم قرار دادن بردار وصله‌ها ماتریس hypothesis ساخته شود. هم‌زمان با تولید ماتریس hypothesis تمام وصله‌ها به ابعاد $sB \times sB$ با bicubic کوچک می‌شوند و یک ماتریس hypothesis دیگری به نام H^l ساخته می‌شود. سپس برای هر وصله وضوح پایین در تصویر ورودی با استفاده از H^l و تنظیم‌کننده تیخونوف [۳۸] ضرایب مورد نیاز برای ساخت وصله‌ی وضوح‌بالای آن به دست می‌آید و با ضرب این ضرایب در ماتریس H وصله وضوح‌بالای آن ساخته می‌شود.

Zhu و همکارانش [۳۹] یک روش سریع برای حل مساله SRR ارائه دادند. آن‌ها یک پیاده‌سازی کارا بر پایه الگوریتم K-SVD پیشنهاد کردند. این روش بدون نیاز به پایگاه داده خارجی بزرگ و فقط با استفاده از تصویر ورودی می‌باشد و از سرعت خوبی در فراتفکیک‌پذیری برخوردار است. در این مقاله وصله‌هایی که برای یادگیری دیکشنری استفاده می‌شوند با یک روش مناسب از تصویر ورودی وضوح پایین انتخاب می‌شود. روش پیشنهادی برای انتخاب وصله‌ها، تکنیک $SMSE^1$ می‌باشد. از معیار بهینه‌سازی مبتنی بر L_0 که بسیار سریع‌تر از L_1 می‌باشد برای هر دو فاز آموزش و بازسازی استفاده شده است. روش کار به این صورت است که در ابتدا به صورت تصادفی وصله‌هایی از تصویر استخراج می‌شوند. تعداد وصله‌هایی که به صورت تصادفی استخراج می‌شوند برابر با ۳۰۰۰ وصله است. با استفاده از معیار SMSE وصله‌هایی که اطلاعات بالایی دارند برای ساخت دیکشنری استفاده می‌شود. در مرحله بعد با استفاده از الگوریتم بهبودیافته K-SVD یک دیکشنری وضوح‌بالا ساخته می‌شود. سپس دیکشنری وضوح

¹ Sample mean square error

^۲ نرم L_0 که به صورت $\| \cdot \|_0$ نشان داده می‌شود تعداد عناصر غیر صفر یک بردار را نشان می‌دهد.
^۳ نرم L_1 به عنوان حداقل انحراف مطلق شناخته شده است. به طور اساسی جمع تفاوت مطلق (S) میان مقدار هدف (Y_l) و مقادیر تخمین زده شده $f(X_l)$ را مینیمم می‌کند:

$$s = \sum_{i=1}^n |y_i - f(x_i)|$$

پایین با کاهش نمونه‌ها از دیکشنری وضوح‌بالا ساخته می‌شود. برای هر وصله در تصویر ورودی با استفاده از دیکشنری وضوح پایین ضرایب موردنیاز برای ساخت وصله‌ی SR به دست می‌آید. سرانجام این ضرایب در دیکشنری وضوح‌بالا ضرب می‌شوند تا تصویر SR ساخته شود.

۵-۲-۲ روش‌های مناسب برای بازسازی ویدیوهای واقعی

هدف از فراتفکیک‌پذیری بازبازی تصویر وضوح‌بالای X با استفاده از تصاویر وضوح پایین Y می‌باشد. این الگوریتم برای بازبازی جزئیات صحنه، وابسته به تخمین حرکت بین فریم‌های ویدیو می‌باشد. اولین مرحله در روش‌هایی که تاکنون در مورد آن‌ها بحث شد (موسوم به SRR کلاسیک) به تخمین حرکت بین فریم‌ها می‌پردازد. در الگوریتم‌های SRR کلاسیک فقط تخمین سراسری حرکت (برای مثال انتقال^۱) و تبدیل‌های سراسری آفاین^۲ به‌اندازه‌ای دقیق می‌باشند که بتوان بازسازی فراتفکیک‌پذیری موفق‌تری داشت. بیشتر ویدیوهای واقعی شامل صحنه‌هایی هستند که حرکت‌های محلی به صورت ذاتی در آن‌ها وجود دارد (مثل شخصی که در حال صحبت است) بنابراین نمی‌توان ویدیوهای واقعی را تنها به حرکت‌های سراسری محدود کرد.

به دست آوردن تخمین دقیقی از حرکت محلی (که به آن جریان نوری^۳ نیز می‌گویند) کار بسیار دشواری می‌باشد مخصوصاً وقتی که نویز و روی هم افتادگی فرکانسی وجود داشته باشد. هنگامی که یکی از روش‌های SRR کلاسیک با تخمین غیر دقیق حرکت به کار برده شود، نتایج بسیار ضعیفی (حتی نسبت به روش‌های درون‌یابی) همراه با مصنوعات زیاد حاصل می‌گردد. بنابراین نمی‌توان ویدیوهای واقعی با الگوی حرکتی عمومی را با استفاده از روش‌های SRR کلاسیک بازسازی نمود. لذا در بین محققان نیاز

¹ Translation

² Affine global transform

³ Optical flow

به روشی متفاوت از SRR کلاسیک، برای ویدیوهای واقعی احساس شد. تاکنون چندین الگوریتم برای بازسازی تصویر وضوح بالا در ویدیوهای واقعی پیشنهاد شده است که در این بخش برخی از آنها را مرور می‌کنیم. نخستین روش SRR مناسب برای بازسازی ویدیو توسط نویسندگان در مرجع [۴۰] پیشنهاد شد. نقطه‌ی شروع آن‌ها برای توسعه‌ی SRR بدون تخمین دقیق حرکت، الگوریتم‌های نوین زدایی بود. اخیراً روش‌هایی برای نوین زدایی ویدیو پیشنهاد شده است که در عین حال که از تخمین حرکت اجتناب می‌نمایند، نتایج خوبی نیز ارائه می‌دهند [۴۱، ۴۲]. از میان روش‌های ارائه شده برای نوین زدایی تصویر آن‌ها فیلتر^۱ NLM را برای تعمیم به فراتفکیک‌پذیری انتخاب کردند. در مقاله‌ی مذکور مسئله‌ی نوین زدایی ویدیو به وسیله‌ی فیلتر NLM ابتدا به مسئله‌ی فراتفکیک‌پذیری تک فریمی و سپس با در نظر گرفتن اطلاعات زمانی به SRR ویدیو بدون تخمین دقیق حرکت تعمیم داده شد. آن‌ها در مقاله‌ی [۴۳] با فرض اینکه می‌توان حرکت‌های محلی را به صورت جمع وزن دار حرکت‌های سراسری مدل کرد، موفق شدند یک چهارچوب احتمالاتی برای حل مسئله‌ی SRR ویدیو پیشنهاد دهند. تابع جریمه‌ی پیشنهادی آن‌ها مستقیماً از مدل تصویر SRR کلاسیک استخراج می‌شود به طوری که عملگر waring در فرمولاسیون SRR کلاسیک با معادل آن (یعنی جمع وزن دار حرکت‌های سراسری) جایگزین می‌شود. به منظور سادگی حل مسئله پیشنهاد شد به جای معادله‌ی ماتریسی-برداری به دست آمده از معادل پیکسلی آن استفاده نماییم. آن‌ها نشان دادند که معادل پیکسلی به دست آمده با چهارچوب پیشنهادی قبلی آن‌ها (یعنی [۴۰]) یکسان می‌باشد. در فصل بعد درباره‌ی این مقاله مفصل توضیح خواهیم داد.

Takeda و میلان فر [۴۴] چهارچوب جدیدی برای حل مسئله‌ی SRR ویدیو بدون تخمین دقیق حرکت پیشنهاد دادند. روش آن‌ها علاوه بر بزرگ‌نمایی تصویر در فضا، قادر بود در زمان هم فریم‌های

^۱ Nonlocal means

جدیدی به رشته‌ی ویدیویی اضافه کند. این روش بر پایه‌ی رگرسیون کرنلی^۱ چند بعدی می‌باشد و هر پیکسل در داخل رشته‌ی ویدیویی با یک سری سه بعدی (تیلور) که از اطلاعات پیکسل‌های همسایه ساخته شده است، تخمین زده می‌شود. علی‌رغم اینکه روش آن‌ها بسیار کارآمد است اما به لحاظ محاسباتی بسیار پیچیده می‌باشد. در [۴۵] چهارچوب جدیدی برای سرعت بخشیدن به روش ارائه شده در [۴۴] پیشنهاد شده است.

در [۴۶] یک روش جدید بر پایه‌ی [۴۰] توسعه داده شد. در این روش ابتدا برای هر پیکسل تصویر وضوح‌بالای اولیه تعدادی کاندید به‌وسیله‌ی کدگذاری تنک و انتشار باور^۲ در فریم‌های وضوح پایین پیدا کردند. سپس با استفاده از رابطه‌ی فیلتر NLM پیکسل هدف را بازسازی کردند. در [۴۷] یک روش جدید SRR ویدیو بر پایه‌ی کدگذاری تنک پیشنهاد شد. در این روش ابتدا به‌وسیله‌ی الگوریتم block matching حرکت بین فریم‌ها را با دقت زیر پیکسلی تخمین زدند. سپس از کدگذاری تنک برای بازسازی فریم وضوح‌بالا استفاده کردند. در [۴۸] یک روش جدید برای SRR ویدیو با استفاده از یک پایگاه داده‌ی خارجی پیشنهاد شده است. در این روش به جای تخمین دقیق حرکت بین فریم‌ها پیشنهاد کردند جزئیات فرکانس بالا به‌وسیله‌ی مثال‌های خارجی به دست آید. آن‌ها از یک دیکشنری که از مثال‌های خارجی ساخته می‌شود برای بازسازی تصویر استفاده کردند. همچنین در این مقاله از مفهوم دگرشکل-پذیری که در کار قبلی نویسندگان [۴۹] توسعه داده شده بود برای تغییر شکل وصله‌ها و بهبود لبه‌ها استفاده شد.

¹ kernel regression

² belief propagation

۳ فصل سوم: توری

بازسازی فراتفکیک‌پذیری ترکیب چندین تصویر وضوح پایین $\{y_t\}_{t=1}^T$ برای ساخت یک تصویر وضوح بالا X می‌باشد به طوری که تصویر وضوح بالای X دارای تفکیک نوری بالاتر می‌باشد.

تعداد الگوریتم‌های زیادی در طی دو دهه‌ی اخیر برای حل مسئله‌ی SRR توسعه داده شده‌اند. در مدل مشهوری که اکثر این مقالات از آن استفاده می‌نمایند فرض شده است که تصویر وضوح بالای X به وسیله‌ی رشته‌ای از عملگرها به دست می‌آید. این عملگرها عبارتند از:

(۱) تغییر شکل هندسی (F_t) ^۱

(۲) ماتری خطی و ثابت در فضا (H) ^۲

(۳) کاهش ابعاد (D)

(۴) اضافه کردن نویز سفید گوسی با میانگین صفر (n_t) که هم نویز اندازه‌گیری و هم عدم انطباق مدل تصویر را مدل می‌کند [۹].

عملگرهای بالا تماماً خطی می‌باشند و به صورت ضرب ماتریسی در تصویر عمل می‌نمایند. در تمام این پایان‌نامه فرض می‌کنیم عملگرهای D و H برای تمام فریم‌های رشته‌ی ویدیویی مورد مطالعه یکسان می‌باشند. مدل گفته شده منجر به معادله‌ی زیر می‌شود:

$$y_t = DHF_t X + n_t \quad \text{for } t=1,2,\dots, T \quad ۱-۳$$

بنابراین بازیابی X از $\{y_t\}_{t=1}^T$ یک مسئله‌ی معکوس می‌باشد که از نویز زدایی، ماتری زدایی، افزایش مقیاس و ترکیب چندین تصویر در یک مسئله‌ی واحد تشکیل شده است. با قرار دادن $F_1 = I$ به y_1 به عنوان تصویر مرجع می‌رسیم و می‌خواهیم X را به عنوان نسخه‌ی وضوح بالای آن بسازیم.

^۱ Geometrical warp

^۲ Linear space invariant blur

SRR کلاسیک فرض می‌کند که D ، H و F_t شناخته شده می‌باشند یا اینکه می‌توان آن‌ها را از طریق داده‌های ورودی به صورت مطلوبی تخمین زد.

در عمل این نوع بازسازی تصویر وابسته به توانایی تخمین حرکت در صحنه (با دقت زیر پیکسلی) می‌باشد زیرا لازم است مقادیر مختلف با شبکه‌ی نمونه‌برداری^۱ متفاوت را باهم ادغام نماییم. اکثر الگوریتم‌های *SRR* با تخمین حرکت رشته‌ی ویدیویی با دقت زیر پیکسلی شروع می‌شوند [۲۸، ۵۰].

تخمین حرکت محلی بسیار دقیق (که به عنوان جریان نوری شناخته می‌شود) یک مسئله‌ی به شدت نامعین^۲ می‌باشد. وقتی در الگوریتم‌های *SRR* کلاسیک حرکت بین فریم‌ها را به صورت غیر دقیق تخمین بزنیم خروجی غالباً به مصنوعات مزاحم^۳ ختم می‌شود. به همین دلیل در اکثر الگوریتم‌های *SRR* کلاسیک تعدادی فرضیات ساده شونده در ساختار حرکت الگوریتم در نظر گرفته می‌شود (مانند حرکت-های کلی^۴ یا اجسام صلب^۵). فقط تحت این شرایط الگوریتم‌های *SRR* کلاسیک می‌توانند به تخمین قابل قبولی از تصویر وضوح‌بالا برسند. به عنوان نتیجه می‌توان گفت الگوریتم‌های *SRR* کلاسیک توانایی بازسازی فیلم‌های واقعی با محتوای عمومی را ندارند.

با استفاده از ۱-۳ تخمین بیشترین شباهت با مینیمم کردن تابع جریمه‌ی زیر نسبت به X به دست می‌آید.

$$\epsilon_{ML}^2(X) = \frac{1}{2} \sum_{t=1}^T \|DHF_t X - y_t\|_2^2 \quad 2-3$$

¹ Sampling grid

² Severely under-determined

³ Disturbing artifacts

⁴ Global warping

⁵ Rigid bodies

این توصیف پیشنهاد می‌کند که تصویر X باید به گونه‌ای به دست آید که به درستی تصاویر LR را توصیف نماید. مشتق گرفتن از رابطه‌ی ۲-۳ به معادله‌ی زیر ختم می‌شود.

$$\frac{\partial \epsilon_{ML}^2(X)}{\partial X} = \sum_{t=1}^T F_t^T H^T D^T (D H F_t X - y_t) = 0 \quad 3-3$$

اگر $A = \sum_{t=1}^T F_t^T H^T D^T D H F_t$ و $b = \sum_{t=1}^T F_t^T H^T D^T y_t$ در نظر گرفته شود، معادله‌ی سیستم خطی $A \hat{X}_{ML} = b$ حاصل می‌گردد.

در خیلی از مواقع مقادیر اندازه‌گیری شده $\{y_t\}_{t=1}^T$ به اندازه کافی وجود ندارد که باعث می‌شود ماتریس A منفرد^۱ یا بدحالت شود و نتوانیم X را بازیابی کنیم. در این حالت‌ها ما نیاز به عبارت تنظیم‌کننده^۲ داریم. تخمین ماکزیمم احتمال پسین MAP تابع جریمه‌ی زیر را پیشنهاد می‌دهد.

$$\epsilon_{MAP}^2(X) = \epsilon_{ML}^2(X) + \lambda R(X) \quad 4-3$$

که تابع R عبارت تنظیم‌کننده می‌باشد و به معکوس ماتریس A پایداری جبری می‌دهد.

علاوه بر پایداری، از R به عنوان یک دانش پیشین درباره‌ی X نیز استفاده می‌شود (برای مثال صافی فضایی^۳، تنگی ارائه‌ی موجک، آنتروپی کمینه و ...). ما در این پایان‌نامه از $total\ variation$ استفاده می‌نماییم. به طوری که $R(X) = Tv(X)$ جمع نرم L_1 گرادیان تصویر را مینیمم می‌کند و باعث صافی تصویر می‌شود [۵۱]. بنابراین تخمین MAP مورد استفاده‌ی ما به صورت زیر می‌باشد.

¹ Singular

² Regularization

³ Spatial smoothness

$$\epsilon_{MAP}^2(X) = \frac{1}{2} \sum_{t=1}^T \|DHF_t X - y_t\|_2^2 + \lambda TV(X)$$

که عموماً توسط الگوریتم‌های تکراری [۹، ۵۰] قابل حل می‌باشد. این تکنیک هسته‌ی اصلی روش پیشنهادی این پایان‌نامه می‌باشد.

در تمام بحث‌های انجام شده تا به حال فرض بر این بود که عملگر D ، H و F_t شناخته شده می‌باشند. عملگر D وابسته به ضریب بزرگنمایی تفکیک تصویر می‌باشد و بنابراین همیشه ثابت است. فرض می‌کنیم که فاکتور وضوح یک عدد صحیح $S > 1$ در هر دو محور مختصات می‌باشد. در بیشتر موارد عملگر ماتی H ، به PSF دوربین ارتباط پیدا می‌کند و بنابراین در اغلب اوقات قابل دسترس می‌باشد. برخلاف این عملگرها ماتریس‌های F_t به سختی به دست می‌آیند. آن‌ها وابسته به صحنه می‌باشند و برای ساخت آن‌ها به تخمین دقیق حرکت در صحنه نیازمندیم. از آنجایی که در حالت کلی رسیدن به چنین دقتی بسیار مشکل می‌باشد روش‌هایی همانند [۴۰] تکنیک‌هایی را برای کنار گذاشتن تخمین دقیق حرکت پیشنهاد داده‌اند.

در این پایان‌نامه ما بر اساس کار ارائه شده در [۴۳] به دنبال روشی برای تخمین حرکت در SRR احتمالاتی می‌باشیم. بنابراین تمرکز خود را بر روی نظریه‌ی تخمین احتمالاتی حرکت می‌گذاریم.

فرض شود عملگر F_t به صورت ارتباط یک‌به‌یک بین پیکسل‌های داخل فریم مرجع و t -امین فریم در نظر گرفته شود. اگر اندازه بیشترین حرکت بین دو فریم D پیکسل باشد بنابراین تعداد تمام جابه‌جایی‌های ممکن $M = (2D + 1)^2$ می‌باشد. اگر $\{F_m\}_{m=1}^M$ مجموعه‌ی تمام جابه‌جایی‌های سراسری ممکن در تصویر باشد می‌توان معادله‌ی زیر را نوشت:

$$F_t X = \sum_{m=1}^M Q_{m,t} F_m X \quad ۶-۳$$

که این عبارت عملیات تغییر شکل تصویر X را بر پایه‌ی عملگر F_t بیان می‌کند. ماتریس $\{Q_{m,t}\}_{m=1}^M$ ماتریس قطری وزن دهی می‌باشد که شامل تعدادی ۱ در راستای قطر اصلی برای پیکسل-هایی که حرکت آن‌ها F_m است و صفر برای سایر پیکسل‌ها می‌باشد. در این حالت می‌توان حرکت‌های پیچیده محلی را به صورت ترکیب خطی حرکت‌های سراسری بیان کرد. در این فرمول نویسی عملگر F_t با ترکیب خطی حرکت‌های سراسری عوض شده است. فرمول ارائه شده در ۶-۳ یک رابطه یک‌به‌یک بین پیکسل‌های هر دو تصویر ارائه می‌دهد. برای معرفی میدان حرکت احتمالی لازم است $Q_{m,t}$ را به‌گونه‌ای تغییر داد که بتوان به‌وسیله‌ی مقادیر پیوسته، مقدار اطمینان هر پیکسل و هر مسیر حرکت را بیان نمود. این مفهوم منجر به نوع جدیدی از فراتفکیک‌پذیری خواهد شد که F_t را با تجزیه‌ای از آن جابه‌جا می‌کند.

فرمول ML معرفی شده در ۲-۳ را به صورت ماکزیمم شباهت احتمالاتی PML می‌نویسیم.

$$\epsilon_{PML}^2(X) = \frac{1}{2} \sum_{m=1}^M \sum_{t=1}^T \|DHF_m X - y_t\|_{w_{m,t}}^2 \quad ۷-۳$$

این فرمول همان توضیحات گفته شده در بالا را نشان می‌دهد به‌جز اینکه به صورت متفاوتی بیان شده است^۱. در این فرمول علاوه بر جمع کردن حرکت‌های سراسری مختلف (برای اعمال اثر F_t) مربعات خطایی که از این حرکت‌های سراسری مختلف به دست می‌آید نیز جمع شده است و به هر کدام از آن‌ها یک وزن $w_{m,t}$ اختصاص داده شده است.

قابل ذکر است که وزن‌های استفاده شده در ۷-۳ با آن چیزی که در ۶-۳ معرفی شده متفاوت

^۱ $\|Q\|_w^2 = a^T w a$

می‌باشد به طوری که $Q_{m,t}$ برای پیکسل‌های تصویر وضوح بالا تعریف می‌شود اما $w_{m,t}$ ماتریس قطری می‌باشد که روی شبکه‌ی وضوح پایین تعریف می‌شود.

فعلا فرض می‌شود $w_{m,t}$ شناخته شده است، اما در فصل ۵ مفصل درباره‌ی آن بحث خواهیم کرد. اگرچه این فرمولاسیون فقط شامل حرکت‌های سراسری می‌باشد اما قادر است هر میدان حرکتی پیچیده-ای را پردازش نماید. اگر میدان حرکتی شناخته شده باشد می‌توان با قرار دادن ۱ برای پیکسل‌هایی که حرکت آن‌ها F_m است و قرار دادن صفر برای سایر پیکسل‌ها محاسبات را انجام داد. استفاده از PML مذکور یک ویژگی متمایزکننده دیگر هم دارد و آن مقاوم کردن الگوریتم نسبت به نقاط پرت می‌باشد. برای مثال اگر یکی از تصاویر موجود در مجموعه پرت باشد، وزن اختصاص داده شده به پیکسل‌های داخل این تصویر صفر می‌باشد، زیرا این تصویر با تصویر مرجع انطباق پیدا نمی‌کند بنابراین این پیکسل‌ها در بهینه‌سازی حذف می‌شوند.

تابعی که قرار است بهینه شود شامل دو جمله می‌باشد: $\mathcal{E}_{PML}^2(X)$ و جمله تنظیم‌کننده‌ی $TV(X)$. می‌توان به جای حل مستقیم مسئله آن را به پیروی از روش‌های توسعه داده شده در [۹, ۵۲] تجزیه نمود. از آنجایی که هم H و هم F_m عملگرهای ثابتی در فضا می‌باشند، می‌توان جای آن‌ها را با هم عوض کرد. بنابراین با تعریف $Z = HX$ ابتدا با مینیمم سازی معادله‌ی زیر نسخه‌ی مات تصویر وضوح بالای Z به دست می‌آید

$$\mathcal{E}_{PML}^2(X) = \frac{1}{2} \sum_{m=1}^M \sum_{t=1}^T \|DF_t Z - y_t\|_{w_{m,t}}^2 \quad ۸-۳$$

که به آن مرحله‌ی ترکیب گفته می‌شود. سپس عملیات ماتی زدایی با مینیمم کردن تابع هزینه به دست می‌آید.

$$\epsilon_{DB}^2(X) = \|HX - Z\|_2^2 + \lambda TV(X) \quad 9-3$$

از آنجا که مرحله‌ی دوم شناخته شده و معروف می‌باشد، از اینجا به بعد روی قسمت نخست الگوریتم متمرکز می‌شویم. نکته اینک ماتریس زدایی معرفی شده در این قسمت نسبتاً ساده است، بنابراین در صورتی که از تکنیک‌های پیشرفته‌تر استفاده شود جواب بهتری به دست می‌آید. با مشتق‌گیری از ۸-۳ داریم:

$$\frac{\partial \epsilon_{PML}^2(X)}{\partial Z} = \sum_{m=1}^M \sum_{t=1}^T F_m^T D^T w_{m,t} (DF_m Z - y_t) = 0 \quad 10-3$$

که به معادلات یک سیستم خطی منجر می‌شود. برای سادگی در بیان مفهوم، نماد نویسی جدید زیر را معرفی می‌نماییم:

$$\hat{w}_m = \sum_{t=1}^T w_{m,t} \quad 11-3$$

$$\hat{y}_m = \sum_{t=1}^T w_{m,t} y_t \quad 12-3$$

ماتریس \hat{w}_m یک ماتریس قطری می‌باشد زیرا از جمع ماتریس‌های قطری به دست آمده است. داریم:

$$\left[\sum_{m=1}^M F_m^T D^T \hat{w}_m DF_m \right] Z = \sum_{m=1}^M F_m^T D^T \hat{y}_m \quad \text{معادله } 13-3$$

حل معادله‌ی فوق به نظر پیچیده می‌آید، بنابراین سعی می‌کنیم این معادله را برای هر کدام از پیکسل‌های تصویر Z در یک فرم بسته بنویسیم.

سمت راست معادله ۳-۱۳ یک تصویر هم بعد Z می‌باشد. علاوه بر این همانطور که نشان خواهیم داد ماتریسی که در سمت چپ ماتریس Z ضرب شده است یک ماتریس مثبت معین قطری می‌باشد. بنابراین می‌توانیم فرمولاسیون ماتریس برداری بالا را به نسخه‌ی پیکسلی آن تبدیل نماییم. اجازه دهید یک پیکسل در محل $[i,j]$ در تصویر Z در نظر بگیریم و بازسازی آن را مشاهده کنیم. همانطور که خواهید دید این پیکسل در سمت راست نیز فقط به $[i,j]$ امین پیکسل وابسته می‌باشد.

برای یک F_m خاص که با $(dx(m), dy(m))$ جابجا شده است عبارت $F_m^T V$ مکان $[i+dx(m), j+dy(m)]$ $dy(m)$ امین المان از تصویر V را نشان می‌دهد (ترانهاده تاثیر جابجایی معکوس دارد). تصویر $u = D^T \hat{y}_m$ نسخه‌ی افزایش مقیاس یافته‌ی تصویر وضوح پایین \hat{y}_m به وسیله‌ی پر کردن با صفر^۱ می‌باشد. این بیان می‌کند که اگر مکان $[i+dx(m), j+dy(m)]$ مضرب صحیحی از s (نرخ تفکیک) نباشد این مکان مقداری برابر صفر خواهد داشت. در غیر این صورت این مکان مقداری برابر $\hat{y}_m[k, L]$ دارد به طوری که:

$$[k, L] = [i + dx(m), j + dy(m)] \quad ۱۴-۳$$

بنابراین در مکان $[i, j]$ داریم:

$$RHS[i, j] = \sum_{[k, L] \in N(i, j)} \hat{y}_m[k, L] \quad ۱۵-۳$$

همسایگی $N(i, j)$ به صورت زیر تعریف می‌شود

$$N(i, j) = \{[k, L] | \forall m \in [1, M], S.K = i + dx(m), S.L = j + dy(m)\} \quad ۱۶-۳$$

^۱ Zero-filling

اگر از ۱۲-۳ معادل \hat{w}_m را جایگزین نماییم، داریم:

$$RHS[i, j] = \sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m,t}[K, L] y_t[K, L] \quad ۱۷-۳$$

در عبارت بالا $w_{m,t}[K, L]$ به عنصری روی قطر اصلی $w_{m,t}$ اشاره می‌کند که در $[K, L]$ امین درایه y_t ضرب می‌شود.

حال درباره‌ی عبارت سمت چپ معادله ۱۳-۳ بحث می‌کنیم. عملگر $D^T \hat{w}_m D$ یک ماتریس قطری می‌باشد که در هر محور مختصات تصویر را با فاکتور وضوح S کوچک می‌کند، سپس به هر کدام از پیکسل‌ها به وسیله‌ی ماتریس قطری \hat{w}_m یک وزن اختصاص می‌دهد و سپس تصویر را با فاکتور S به وسیله‌ی قرار دادن صفر بزرگنمایی می‌کند. بدین معنی که وقتی این عملگر روی تصویر V در مکان $[i, j]$ اثر می‌کند، اگر $[i, j]/S$ یک عدد صحیح نباشد صفر (NULL) می‌شود در غیر این صورت به سادگی وزن می‌گیرد و تبدیل به $\hat{w}_m[i, j]V[i, j]$ می‌شود. وقتی که عملگر $F_m^T D^T \hat{w}_m D F_m$ به $[i, j]$ امین پیکسل اعمال شود این عملگر آن را به مکان $[i + dx(m), j + dy(m)]$ جابجا می‌کند، سپس بسته به اینکه آیا $[i + dx(m), j + dy(m)]/S$ یک عدد صحیح است یا نه آن را صفر می‌کند و یا به آن وزن می‌دهد، در نهایت آن را به اندازه‌ی $[-dx(m), -dy(m)]$ جابجا می‌کند و به مکان اصلی $[i, j]$ بازمی‌گرداند. در حقیقت عملگر $F_m^T D^T \hat{w}_m D F_m$ هر پیکسل را به مکان اصلی خودش بازمی‌گرداند. از آنجایی که هر پیکسل خروجی فقط وابسته به مقدار پیکسل ورودی از همان نقطه می‌باشد پس این ماتریس یک ماتریس قطری است. بنابراین مقدار عددی که در $[i, j]$ امین پیکسل Z ضرب می‌شود برابر است با:

$${}^1LHS[i, j] = \sum_{[K, L] \in N(i, j)} \sum_{t=1}^T \hat{w}_m[K, L] Z[i, j] \quad 18-3$$

با استفاده از تعریف w_m در ۱۱-۳ داریم

$$LHS[i, j] = \sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m,t}[K, L] Z[i, j] \quad 19-3$$

فرض می‌کنیم که جمع وزن‌ها در ۱۹-۳ مقداری مثبت باشد (یعنی حداقل یک مقدار غیر صفر

در آن وجود داشته باشد) با ترکیب ۱۷-۳ و ۱۹-۳ به یک معادله بافرم بسته برای پیکسل $[i, j]$ در تخمین

Z خواهیم رسید:

$$\hat{Z}[i, j] = \frac{\sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m,t}[K, L] y_t[k, L]}{\sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m,t}[K, L]} \quad 20-3$$

¹ Left hand side

۴ فصل چہارم: مثبت احتمالاتی تصویر

در سالیان اخیر ثبت تصویر^۱ در شاخه‌ی پردازش تصویر به صورت گسترده‌ای در مقالات مورد بررسی قرار گرفته است. ثبت تصویر یک مرحله پیش‌پردازش مهم در کاربردهای متنوعی مثل تصاویر پزشکی، فراتفکیک‌پذیری و تصویربرداری از راه دور می‌باشد. هنگامی که به دنبال تطبیق چگال^۲ دو تصویر باشیم تبدیل‌ها و دگر شکل‌پذیری‌های^۳ بسیاری شامل تغییرات رنگ، تغییر مقیاس، چرخش و جابجایی باید در نظر گرفته شود، در حالی که اکثر روش‌هایی که تاکنون ارائه شده‌اند فقط قسمتی از این تبدیل‌ها را در نظر می‌گیرند که باعث نتیجه نادرست در خروجی می‌شود. در این فصل یک روش جدید ثبت چگال^۴ تصویر بر پایه‌ی Relaxation labelling ارائه شده است. در روش پیشنهادی، از بلوک‌های تصویر به عنوان ویژگی استفاده شده است، سپس برای انتخاب بهینه‌ترین نظیر، Relaxation labelling به نقاط کاندیدا اعمال می‌شود. روش پیشنهادی از لحاظ کیفیت بصری و همچنین معیار ارزیابی PSNR و Bad pixel به نتایج رضایت بخشی رسیده است و قادر است به صورت مقاوم صحنه‌ها را ثبت نماید و نسبت به سایر روش‌های موجود نتایج بهتری را ارائه دهد.

ثبت تصویر فرآیند روی هم قرار دادن چند تصویر گرفته شده از یک صحنه در زمان‌های مختلف و از زوایای دید^۵ مختلف می‌باشد [۵۳]. ثبت تصویر در کاربردهای زیادی شامل تصویربرداری از راه دور^۶، فراتفکیک‌پذیری، ردیابی تغییرات بصری^۷ [۵۴]، بخش‌بندی تصویر^۸ [۵۵] و دسته‌بندی تصویر^۹ [۵۶] کاربرد دارد. کاربرد برجسته‌ی دیگر ثبت تصویر بازسازی صحنه با مقیاس بزرگ با استفاده از تعدادی عکس می‌باشد. ثبت تصویر ممکن است برای تطبیق تصاویری که از یک صحنه و یا از صحنه‌های متفاوتی گرفته

¹ Image Registration

² Dense Matching

³ Deformation

⁴ Dense Registration

⁵ view point

⁶ Remote sensing

⁷ Visual tracking

⁸ Image segmentation

⁹ Image classification

شده است استفاده شود. اگر تصاویر ورودی از یک صحنه در زمان‌های یکسان اما از زاویه‌ی دید متفاوتی گرفته شده باشند به آن تطبیق استریو^۱ نیز گفته می‌شود.

روش‌های ثبت تصویر را می‌توان به دو گروه ثبت چگال تصویر^۲ و ثبت تُنک^۳ تقسیم کرد. روش‌های ثبت چگال تصویر سعی در پیدا کردن نقاط مشابه از تصویر مرجع برای تمام نقاط تصویر تست را دارند. در حالی که روش‌های ثبت تُنک تصویر سعی در تطبیق تعداد اندکی نقطه کلیدی درون تصویر تست و مرجع را دارند. تطبیق تُنک در مقالات بسیاری [۵۷-۵۹] به دلیل کارایی بالایی که دارد استفاده شده است. الگوریتم‌های تطبیق تُنک عموماً مبتنی بر ویژگی می‌باشند. در الگوریتم‌های مبتنی بر ویژگی تمام شرایط چرخش، تغییر روشنایی، تغییر مقیاس و حتی دگرشکل‌پذیری در ویژگی آن‌ها (برای مثال SIFT [۶۰]) در نظر گرفته شده است. اگرچه بسیاری از کاربردها به صورت ذاتی به تطبیق چگال نیاز دارند [۶۱، ۶۲].

الگوریتم‌های ثبت تصویر را می‌توان به دو گروه محلی یا سراسری به لحاظ نحوه‌ی جستجو تقسیم‌بندی کرد، که به ترتیب به معنی جست و جو در منطقه‌ی محدودی از تصویر و یا جست و جو در سراسر تصویر می‌باشد. جست و جو در پنجره محلی می‌تواند بسیار سریع‌تر باشد اما دقت پایین‌تری دارد، در حالی که جست و جو در سراسر تصویر می‌تواند دقیق‌تر اما با سرعت پایین‌تری انجام شود.

یکی از کارهای کلاسیک ثبت چگال تصویر روش جریان نوری^۴ [۶۳] می‌باشد. این روش به صورت عمده برای تخمین نقشه‌ی اختلاف^۵ بین دو تصویر مشابه استفاده می‌شود. در طی دهه‌ی گذشته

¹ Stereo matching

² Dense matching

³ Sparse matching

⁴ Optical flow

⁵ Disparity map

پیشرفت‌های عمده‌ای در انطباق استریو به وجود آمده است [۶۴-۶۷]. برای مثال نویسندگان در [۶۷] فرضیات وفاداری روشنایی و وفاداری گرادیان در جریان نوری را باهم ترکیب کردند. همچنین بسیاری از محققان شرط صافی^۱ یا ارتباط فضایی^۲ را با استفاده از بهینه‌سازی در فرمول نویسی تطبیق تصاویر وارد می‌کنند [۶۸-۷۱]. برای مثال [۷۰] تطبیق چگال بین دو تصویر را با استفاده از ویژگی‌های تنک در چارچوب بهینه‌سازی انرژی پیشنهاد می‌دهد. نویسندگان در [۶۹] یک روش ثبت چگال با استفاده از Sparse coding و Belief propagation پیشنهاد می‌دهند.

در این فصل یک روش جدید برای ثبت چگال تصویر با استفاده از Relaxation labelling [۷۲] پیشنهاد شده است. در این روش ابتدا از تمام نقاط هر دو تصویر مورد انطباق ویژگی استخراج می‌شود. در مرحله بعد تعدادی پیکسل کاندید در تصویر مرجع برای هر کدام از پیکسل‌های تصویر تست انتخاب می‌شود. سپس با استفاده از اطلاعات محلی پیکسل‌ها و همچنین اطلاعات نقاط همسایه در الگوریتم Relaxation labelling احتمال صحت هر کدام از نقاط کاندید محاسبه می‌شود. در نهایت کاندید بالاترین احتمال به عنوان بهترین انطباق انتخاب می‌شود. این فصل به صورت زیر سازماندهی شده است. در قسمت ۴-۱ روش پیشنهادی توضیح داده شده است. در قسمت ۴-۲ نتایج مقایسه‌ای روش پیشنهادی با سایر روش‌ها آورده شده است.

۴-۱ روش پیشنهادی

این قسمت جزئیات روش پیشنهادی را توضیح می‌دهد. عملیات ثبت چگال تصویر به ۵ مرحله تقسیم شده است که با جزئیات کامل در ادامه توضیح داده می‌شود.

¹ smoothness

² Spatial coherence

۴-۱-۱ استخراج چگال ویژگی با استفاده از بلاک‌های تصویر

در مرحله نخست نیاز به استخراج ویژگی از تصویر مرجع A و تصویر تست B داریم. برای استخراج چگال ویژگی در اطراف پیکسل‌های هرکدام از تصاویر A و B یک وصله با ابعاد $(2k + 1) \times (2k + 1)$ در نظر گرفته می‌شود که k یک عدد صحیح مثبت می‌باشد. برای پیکسل i -ام تصویر مرجع (a_i) در مکان $[x_i, y_i]$ وصله‌ی p_{x_i, y_i}^A به صورت برداری $f_i^A \in R^{s \times 1}$ تبدیل می‌شود به طوری که

$$s = (2k + 1)^2$$

$$p_{x_i, y_i}^A = G_{x_i, y_i} A \quad ۱-۴$$

$$f_i^A = \text{vec}(p_{x_i, y_i}^A) \quad ۲-۴$$

به طوری که $p_{x_i, y_i}^A \in R^{\sqrt{s} \times \sqrt{s}}$ وصله‌ی استخراج شده از مکان $[x_i, y_i]$ از تصویر A می‌باشد. عملگر G_{x_i, y_i} عملگر استخراج وصله با ابعاد $\sqrt{s} \times \sqrt{s}$ از مکان $[x_i, y_i]$ می‌باشد. همچنین $\text{vec}(\cdot)$ تابع تبدیل‌کننده‌ی یک وصله به بردار می‌باشد. f_i^A ویژگی استخراج شده از پیکسل i -ام از تصویر A می‌باشد. برای تطبیق تصویر تست با تصویر مرجع لازم است همانند مراحل گفته شده در بالا از تصویر تست B نیز ویژگی استخراج شود.

$$p_{x_i, y_i}^B = G_{x_i, y_i} B \quad ۳-۴$$

$$f_i^B = \text{vec}(p_{x_i, y_i}^B) \quad ۴-۴$$

به طوری که $p_{x_i, y_i}^B \in R^{\sqrt{s} \times \sqrt{s}}$ وصله‌ی استخراج شده از مکان $[x_i, y_i]$ از تصویر B می‌باشد. $f_i^B \in R^{s \times 1}$ برداری شده‌ی وصله‌ی استخراج شده از پیکسل i -ام (B_i) در مکان $[x_i, y_i]$ از تصویر تست B می‌باشد.

۴-۱-۲ پیدا کردن نقاط کاندید

هدف اصلی این مرحله پیدا کردن پیکسل‌هایی از تصویر A (a_i هایی) می‌باشد که بیشترین شباهت را به پیکسل j -ام تصویر تست (b_j) از لحاظ روشنایی دارند. در این مرحله لازم است مقدار ماکزیمم بازه‌ی اختلاف^۱ دو تصویر تخمین زده شود. این مقدار به وسیله‌ی بازرسی بصری مشخص خواهد شد. بنابراین برای هر پیکسل b_j یک ناحیه به مرکز $[x_j, y_j]$ با در نظر گرفتن بیشترین بازه‌ی اختلاف در تصویر A مشخص می‌شود. اگر a مجموعه‌ی m پیکسل موجود در تصویر A باشد:

$$a = \{a_1, a_2, \dots, a_m\} \quad ۵-۴$$

اگر C_j مجموعه l پیکسل موجود در همسایگی مشخص شده برای b_j در تصویر A تعریف شود:

$$C_j = \{c_1, c_2, \dots, c_l\} \quad ۶-۴$$

به طوری که $C_j \subset a$ می‌باشد.

حال لازم است در داخل مجموعه‌ی C_j به دنبال تعداد n تا از نزدیک ترین پیکسل‌ها به پیکسل

b_j بگردیم. اگر فاصله‌ی اقلیدسی بین f_j^B را با تمام توصیفگرهای اعضای C_j به دست آورده شود:

$$\varepsilon_j = \{\varepsilon_j(c_1), \varepsilon_j(c_2), \dots, \varepsilon_j(c_l)\} \quad ۷-۴$$

حال تعداد n تا از پیکسل‌های C_j که در مجموعه‌ی ε_j مقادیر کوچک‌تری دارند به عنوان نقاط

کاندید پیکسل b_j انتخاب می‌شود.

^۱ Disparity range

$$\Omega_j = \{d_1, d_2, \dots, d_n\}$$

۸-۴

به طوری که $\Omega_j \subset C_j$ مجموعه‌ی کاندیدهای پیکسل b_j می‌باشد و $d_i \in C_j$ و $i=1,2,\dots,n$ است.

۴-۱-۳ لبه یابی

دشواری ثبت چگال تصویر عموماً بستگی زیادی به محتویات تصویر ورودی دارد. برای مثال در نواحی که از اجزای فرکانس بالا تشکیل شده باشند مانند لبه‌ها و یا بافت‌ها^۱ فرآیند ثبت تصویر آسان‌تر و دقیق‌تر خواهد بود اما در نواحی صاف^۲ که به طور طبیعی ویژگی مشخص و متمایزی ندارند فرآیند ثبت تصویر دشوارتر و همراه با خطا خواهد بود.

در روش ارائه‌شده، با توجه به اینکه هرکدام از همسایه‌های یک پیکسل با چه احتمالی و در چه مکانی ثبت شوند، سعی می‌شود مکان نقطه‌ی مرکزی این همسایگی تخمین زده شود. یعنی در روش ارائه شده موقعیت مکانی همسایه‌ها کاملاً به یکدیگر وابسته می‌باشند. ثبت نقاط در نواحی هموار به دلیل عدم وجود تمایز بین دو نقطه مجاور، همراه با خطا خواهد بود. حال در صورتی که تمام پیکسل‌های تصویر B هم‌زمان ثبت شوند، خطای موجود در ثبت نقاط هموار روی ثبت نقاط لبه و بافت نیز اثر می‌گذارد و خروجی همراه با خطا خواهد بود. با توجه به مقدمه گفته شده، ابتدا تمام نقاط لبه را ثبت می‌نماییم. سپس مکان سایر نقاط را با توجه به مکان لبه‌های همسایه (که با دقت بالایی به دست آمده است) تخمین می‌زنیم.

در این مرحله پیکسل‌های موجود در تصویر B را به دو دسته تقسیم می‌شوند (۱) پیکسل‌های

روی لبه (۲) پیکسل‌های نقاط صاف و هموار

^۱ Texture

^۲ Smooth

$$V = \{b_i \mid E_{x_i, y_i} = 1\} \quad 9-4$$

$$U = \{b_i \mid E_{x_i, y_i} = 0\} \quad 10-4$$

به طوری که V مجموعه‌ی نقاط لبه و U مجموعه‌ی نقاط هموار تصویر B می‌باشند. E_{x_i, y_i} به صورت زیر تعریف می‌شود.

$$E_{x_i, y_i} = \begin{cases} 1 & \text{if } b_i \text{ is an edge} \\ 0 & \text{otherwise} \end{cases} \quad \text{معادله ۱۱-۴}$$

برای به دست آوردن لبه‌های تصویر B از لبه یاب Prewitt استفاده شده است.

۴-۱-۴ اعمال Relaxation labelling

همانطور که در بخش قبلی توضیح داده شد برای هر نقطه از تصویر B تعداد n کاندید از تصویر A استخراج شده است. حال از کاندیدهای استخراج شده برای پیکسل b_j موجود در مجموعه‌ی Ω_j به عنوان دانش پیشین استفاده می‌شود تا بهترین انطباق برای داده‌های تست به دست آید. باید ذکر شود که در بخش قبلی نقاط کاندید انطباق، بر اساس خصوصیات محلی پیکسل‌های ورودی انتخاب شده‌اند و خصوصیات هندسی انطباق در نظر نگرفته شده است. برای مثال به جز نقاط محدودی اطراف مرز اشیاء، می‌توان انتظار داشت که پیکسل‌های همسایه در تصویر B ، باید به پیکسل‌های همسایه در تصویر A انطباق پیدا کنند. برای اضافه کردن این خصوصیات هندسی به حل مسئله، ما مسئله ثبت تصویر را به صورت یک مسئله احتمالاتی مدل می‌کنیم و با اعمال Relaxation labelling بهترین انطباق را می‌یابیم.

مسئله‌ی انطباق گراف^۱ در مقالات مختلف حوزه‌ی بینایی ماشین در سالیان اخیر در کاربردهای

^۱ Graph matching

مختلفی به کار گرفته شده است. در روش‌های پیشنهاد شده برای انطباق گراف، Relaxation labelling به عنوان یکی از موثرترین روش‌ها شناخته شده است [۷۲]. این روش از اطلاعات بافت برای به‌روزرسانی احتمال توزیع برچسب هر کدام از گره‌های گراف، در یک روش تکراری استفاده می‌کند. این روش نخستین بار توسط [۷۳] پیشنهاد شده است. در روش پیشنهادی، هر کدام از تصاویر به صورت یک گراف در نظر گرفته می‌شوند، به صورتی که پیکسل‌های تصویر نودهای گراف را تشکیل می‌دهند. نودهای گراف توسط بردار ویژگی استخراج شده از آن پیکسل توصیف می‌شود. گراف پیشنهادی از نوع تمام متصل می‌باشد و هر نود آن، با تمام نودهای گراف در یک همسایگی مشخص ارتباط دارد.

در روش پیشنهادی، با استفاده از احتمال نقاط همسایه‌ی پیکسل b_j ، احتمال کاندیدهای موجود در مجموعه‌ی Ω_j به‌روزرسانی می‌شود. به‌روزرسانی احتمال نقاط کاندید به دو مرحله تقسیم می‌شود. ابتدا احتمال کاندیدهای موجود در مجموعه‌ی V از تصویر به‌روزرسانی می‌شود، سپس با استفاده از احتمال به دست آمده، مکان نقاط موجود در مجموعه U تخمین زده می‌شود.

۱-۴-۱-۴ به‌روزرسانی احتمال کاندیدهای نقاط لبه

در این مرحله مجموعه‌ی نقاط همسایه‌ی پیکسل $b_j \in V$ به صورت زیر تعریف می‌شود.

$$N_j = \{b_i | x_j - r < x_i < x_j + r \ \& \ y_j - r < y_i < y_j + r \ \& \ E_{x_i, y_i} = 1\} \quad ۱۲-۴$$

به‌طوری‌که r پارامتر تعیین ابعاد همسایگی اطراف پیکسل b_j می‌باشد.

حال با استفاده از فرمول به‌روزرسانی احتمال Relaxation labelling [۷۲, ۷۴] احتمال هر کدام

از کاندیدهای $b_j \in V$ را به‌روزرسانی می‌کنیم.

$$P^{(n+1)}(b_j = a_{b_j}) = \frac{P^{(n)}(b_j = a_{b_j})Q^{(n)}(b_j = a_{b_j})}{\sum_{a_\lambda \in \Omega_j} P^{(n)}(b_j = a_\lambda)Q^{(n)}(b_j = a_\lambda)} \quad ۱۳-۴$$

به طوری که $P^{(n)}(b_j = a_{b_j})$ احتمال انتساب کاندید $a_{b_j} \in \Omega_j$ به b_j در تکرار n -ام می باشد. در هر تکرار تمام انتقال های ممکن پیکسل b_j در نظر گرفته می شود و احتمال مربوط به هر کدام با استفاده از احتمال تکرار قبلی و پشتیبانی صورت گرفته از طرف سایر پیکسل ها به روزرسانی می شود. تابع $Q(b_j = a_{b_j})$ مقدار پشتیبانی صورت گرفته از انتساب b_j به $a_{b_j} \in \Omega_j$ از طرف سایر نودهای موجود در گراف می باشد که به صورت زیر تعریف می شود.

$$Q^{(n)}(b_j = a_\alpha) = \sum_{b_i \in N_j} \sum_{a_\beta \in \Omega_i} P(b_i = a_\beta) P(A_{ij} | b_j = a_\alpha, b_i = a_\beta) \quad ۱۴-۴$$

به طوری که Ω_i مجموعه نقاط کاندید پیکسل $b_i \in N_j$ می باشد. $P(A_{ij} | b_j = a_\alpha, b_i = a_\beta)$ احتمال انتساب همزمان b_j به a_α و b_i به a_β می باشد. در مقالات مختلف بسته به کاربرد مورد نظر، تعاریف متفاوتی از $P(A_{ij})$ در نظر گرفته شده است. در این پایان نامه فرض بر این است که پیکسل های همسایه در تصویر تست، در تصویر مرجع نیز همسایه باشند. بنابراین اگر بردار جابه جایی پیکسل b_j به صورت زیر تعریف شود:

$$u_j = \begin{bmatrix} x_j \\ y_j \end{bmatrix} - \begin{bmatrix} x_\alpha \\ y_\alpha \end{bmatrix} \quad ۱۵-۴$$

به طوری که $[x_j, y_j]$ مکان پیکسل b_j و $[x_\alpha, y_\alpha]$ مکان پیکسل متناظر با b_j در تصویر مرجع می باشد. بردار جابه جایی نقطه $b_i \in N_j$ به صورت زیر تعریف می شود.

$$u_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} - \begin{bmatrix} x_\beta \\ y_\beta \end{bmatrix} \quad ۱۶-۴$$

به طوری که $[x_i, y_i]$ مکان پیکسل $b_i \in N_j$ می باشد و $[x_\beta, y_\beta]$ مکان نقطه متناظر در تصویر

مرجع می‌باشد. با توجه به فرض صافی^۱ (یعنی اندازه‌ی بردار جابه‌جایی نقاط همسایه تقریباً یکسان می‌باشد) لازم است اختلاف جابه‌جایی پیکسل b_j و $b_i \in N_j$ ناچیز باشد. بنابراین احتمال هم‌زمانی انتساب به نحوی تعریف می‌شود که اگر اختلاف جابه‌جایی b_j و $b_i \in N_j$ کوچک باشد احتمال بزرگی به آن نسبت دهد و بالعکس.

$$P(A_{ij} | b_j = a_\alpha, b_i = a_\beta) = \exp\left(-\frac{\|u_i - u_j\|_2^2}{2\sigma^2}\right) \quad ۱۷-۴$$

لازم به ذکر است برای استفاده از معادله‌ی ۴-۱۳ لازم است به هر کدام از کاندیدها یک احتمال اولیه نسبت داده شود. احتمال اولیه‌ی نسبت داده شده به کاندید i -ام پیکسل b_j موجود در مجموعه Ω_j با فاصله‌ی اقلیدسی f_j^B (ویژگی پیکسل b_j) و ویژگی کاندید i -ام آن رابطه‌ی عکس دارد. اگر فاصله‌ی اقلیدسی بیشتر باشد نشانگر این است که دو نقطه نامنطبق می‌باشند در نتیجه احتمال کمتری به آن‌ها نسبت داده می‌شود و بالعکس. برای هر عضو مجموعه Ω_j احتمال اولیه به صورت زیر محاسبه می‌شود.

$$P^{(0)}(d_i) = \frac{1}{\varepsilon_j(d_i)+1} \frac{1}{\sum_{d_\lambda \in \Omega_j} \frac{1}{\varepsilon_j(d_\lambda)+1}} \quad ۱۸-۴$$

به طوری که $d_i \in \Omega_j$ و $\varepsilon_j(d_i)$ فاصله اقلیدسی f_j^B از ویژگی مربوط به d_i می‌باشد. مراحل بالا را برای پیکسل‌های لبه تکرار می‌کنیم تا پیکسل‌های لبه با احتمال بالا ثبت شوند.

۴-۱-۴-۲ به روزرسانی احتمال کاندیدهای نقاط هموار

حال که مکان پیکسل‌های لبه‌ی تصویر B با احتمال بالا پیدا شد لازم است تا مکان سایر نقاط تصویر نیز تخمین زده شود. ابتدا برای هر پیکسل $b_j \in U$ یک همسایگی به شکل زیر تعریف می‌شود:

^۱ Smoothness

$$N_j = \{b_i | x_j - r < x_i < x_j + r \ \& \ y_j - r < y_i < y_j + r \} \quad ۱۹-۴$$

حال با استفاده از ۱۳-۴، ۱۴-۴ و ۱۷-۴ که در بخش قبل معرفی شدند می‌توانیم احتمال کاندیدهای نقاط $b_j \in U$ را به دست آوریم.

تفاوت عملیات انجام شده در این قسمت با مرحله‌ی قبل در این است که در این مرحله فقط احتمال کاندیدهای مربوط به $b_j \in U$ به دست آورده شده در حالی که در مرحله قبل فقط احتمال کاندیدهای $b_j \in V$ به دست آورده می‌شود. کاملاً واضح است که با توجه به همسایگی تعریف شده‌ی N_j در این مرحله که در به‌روزرسانی احتمال کاندیدهای $b_j \in U$ از احتمال به دست آمده در مرحله قبل برای نقاط $b_j \in V$ نیز استفاده می‌کنیم. در صورتی که در مرحله قبل برای محاسبه احتمال کاندیدهای $b_j \in V$ فقط از احتمال نقاط همسایه‌ی b_j که عضو مجموعه‌ی V باشد استفاده کردیم.

لازم به ذکر است احتمال اولیه‌ی نقاط عضو مجموعه $b_j \in U$ طبق ۱۸-۴ محاسبه می‌شود و در هر تکرار بروز می‌شود، اما احتمال اولیه‌ی نقاط عضو مجموعه V در این مرحله همان مقادیر به دست آمده در مرحله قبل باقی خواهد ماند و دیگر به‌روز نمی‌شوند.

۴-۱-۵ تفسیر نتایج Relaxation labelling

بعد از چندین مرحله تکرار در الگوریتم Relaxation labelling احتمال بروز رسانی شده هر کدام از کاندیدها به دست آورده می‌شود. این احتمال می‌تواند با عنوان امتیاز انطباق^۱ در نظر گرفته شود. در روش ارائه شده ما کاندیدهای با بالاترین احتمال را به عنوان بهترین انطباق انتخاب می‌نماییم. فرض می‌کنیم که روش ارائه شده در صورتی موفق به پیدا کردن انطباق یک نقطه می‌شود که کاندید با بالاترین

^۱ Matching score

احتمال، احتمالی بالاتر از آستانه (T) داشته باشد، در غیر این صورت فرض می‌کنیم انطباق خوبی برای آن یافت نشده است. بر اساس نتایج تجربی به دست آمده می‌توان گفت عدم انطباق عموماً زمانی رخ می‌دهد که به دلیل پوشیدگی^۱ یا مشکلات مرزها^۲ انطباقی وجود نداشته باشد و یا تعداد تکرار الگوریتم Relaxation labelling به اندازه کافی نباشد.

۲-۴ نتایج تجربی

در این قسمت آزمایشات متفاوتی را برای ارزیابی روش پیشنهادی ارائه شده است. مسئله ثبت دو تصویر که از دو نقطه دید متفاوت گرفته شده اند را در نظر می‌گیریم. برای ارزیابی روش ارائه شده، تست-ها روی مجموعه داده‌هایی شامل تصاویر با مقیاس‌های متفاوت، صحنه‌هایی همراه با چرخش، صحنه‌هایی شامل اشیاء دگردیسی شده^۳ انجام می‌دهیم. تصاویر تست در شرایط نرمال محیطی گرفته شده‌اند. برای بازسازی^۴ تصاویر تست، هر کدام از پیکسل‌های تصویر تست با کاندیدهای انتخاب شده از تصویر مرجع عوض شده‌اند. تصاویر بازسازی شده خروجی در شکل ۴-۵ تا شکل ۴-۸ نشان داده شده‌اند. همچنین نقشه‌ی احتمال هر کدام از تصاویر در قسمت (ز) این شکل‌ها نشان داده شده است. در این نقشه احتمال یک نقطه روشن‌تر، احتمال بالاتر را نشان می‌دهد. اگرچه معیار^۵ PSNR نمی‌تواند به صورت صحیح دقت ثبت را در این آزمایش ارزیابی کند، اما می‌تواند تخمینی از شباهت تصاویر بازسازی شده و تصویر تست بدهد. به منظور مقایسه عددی مقدار PSNR خروجی هر کدام از روش‌ها در جدول ۴-۱ آورده شده است.

¹ Occlusion

² Boundary issues

³ Deform

⁴ synthesize

⁵ Peak signal to noise ratio

در آزمایش دوم روش ارائه شده روی مجموعه دادگان Middlebury database¹ ارزیابی شده

است و نتایج برای ۴ تصویر تست در شکل ۱-۴ تا شکل ۴-۴ نمایش داده شده است. در جدول ۲-۴ مقادیر عددی برای مقایسه با سایر روش‌ها آورده شده است.

همانطور که از نتایج تجربی مشخص است روش ارائه شده به لحاظ تغییرات در تضاد رنگ‌ها، مقیاس، چرخش و از شکل افتادن نسبت به سایر روش‌ها مقاوم‌تر است.

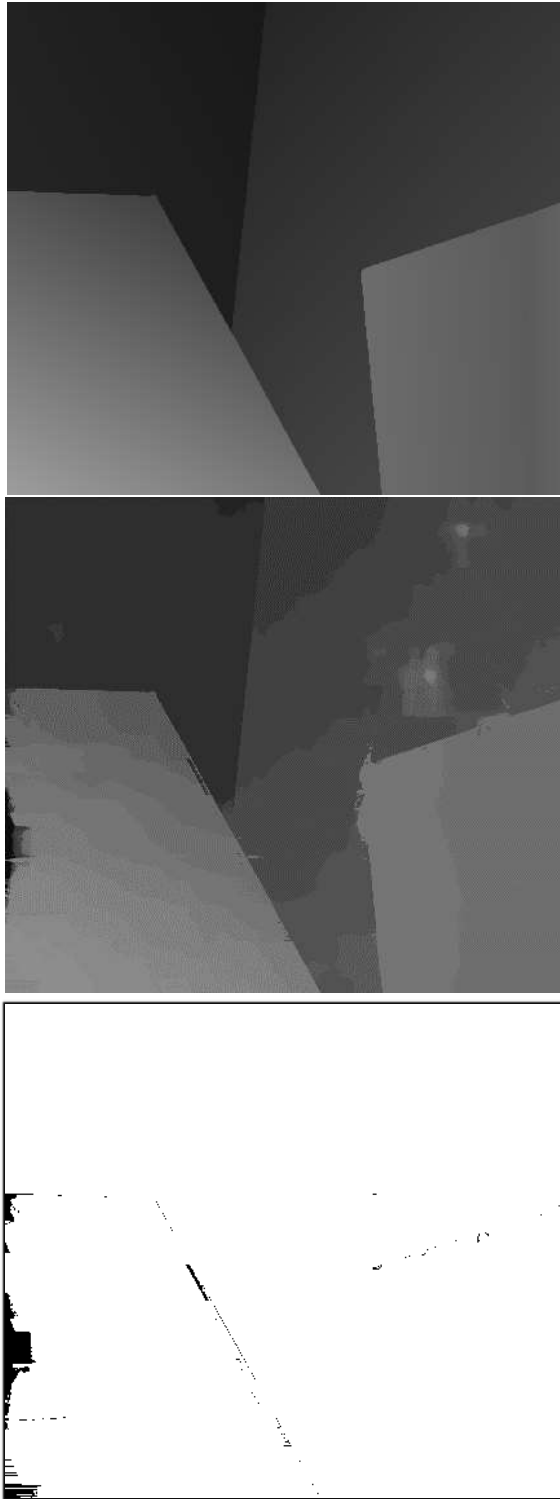
جدول ۱-۴: مقایسه PSNR روش پیشنهادی با سایر روش‌ها

روش	جابجایی (شکل ۶-۴)	چرخش (شکل ۷-۴)	از شکل انداختن (شکل ۸-۴)	بزرگنمایی (شکل ۵-۴)
DSP[71]	21.22	14.89	15.01	19.22
Optical flow[58]	17.21	16.22	11.25	12.95
Sift flow[68]	17.79	14.73	12.51	14.33
روش پیشنهادی	31.94	22.59	26.85	26.49

جدول ۲-۴: مقایسه خطای نقشه‌ی اختلاف به دست آمده توسط روش پیشنهادی با سایر روش‌ها

روش	Cones			Teddy			Venus			Tsukuba		
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc
ARAP [11]	2.08	6.73	6.17	3.01	6.47	9.51	0.38	0.53	4.86	3.07	3.55	11.8
ObjectStereo [12]	2.20	6.99	6.36	4.13	7.59	11.2	0.59	0.69	4.61	1.22	1.62	6.36
PMBP [13]	2.22	6.64	6.48	2.88	8.57	8.99	0.30	0.49	3.57	1.96	2.21	9.22
روش پیشنهادی	1.90	4.90	3.25	5.03	8.76	10.89	0.24	0.34	2.93	1.37	1.34	6.43

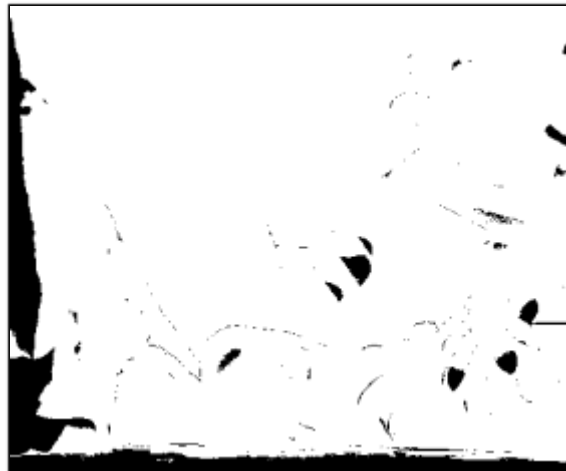
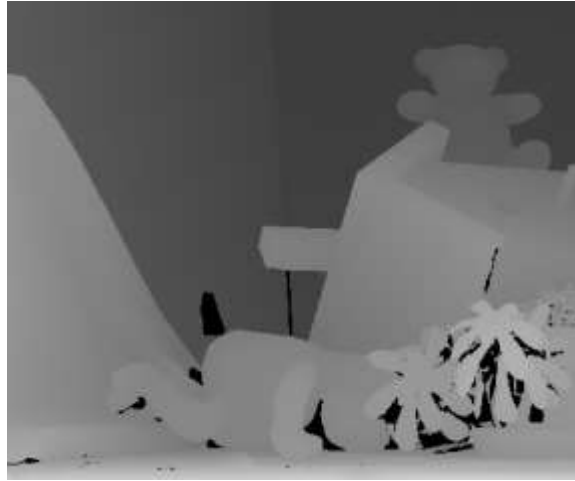
¹ <http://vision.middlebury.edu/stereo/>



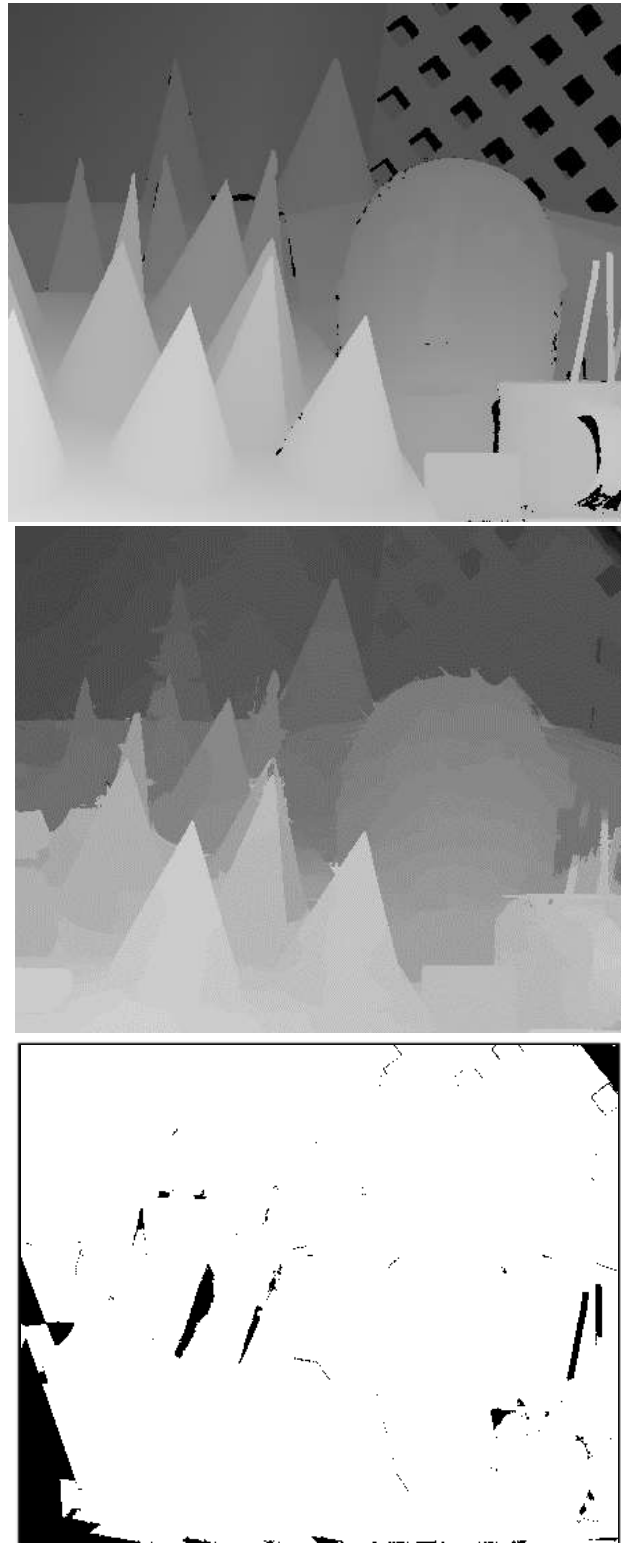
شکل ۴-۱: نقشه‌ی اختلاف تصویر Venus. (بالا) نقشه‌ی صحیح، (وسط) خروجی روش پیشنهادی، (پایین) خطا



شکل ۴-۲: نقشه‌ی اختلاف تصویر Tsukuba. (بالا) نقشه‌ی صحیح، (وسط) خروجی روش پیشنهادی، (پایین) خطا



شکل ۴-۳: نقشه‌ی اختلاف تصویر Teddy. (بالا) نقشه‌ی صحیح، (وسط) خروجی روش پیشنهادی، (پایین) خطا



شکل ۴-۴: نقشه‌ی اختلاف تصویر Cones. (بالا) نقشه‌ی صحیح، (وسط) خروجی روش پیشنهادی، (پایین) خطا



الف



ب



ج



د



ه



و



ز

شکل ۴-۵: نتایج بازسازی رشته‌ی بزرگنمایی الف) تصویر مرجع ب) تصویر تست ج) [71] DSP (د) Optical flow [58] (ه) Sift flow [68] (و) روش پیشنهادی ز) نقشه‌ی احتمال روش پیشنهادی



الف



ب



ج



د



ه



و



ز

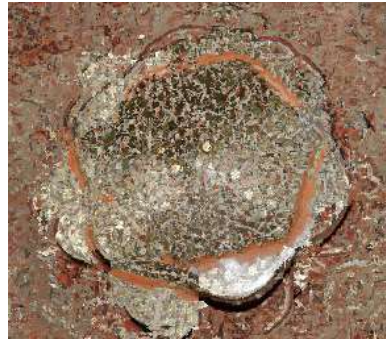
شکل ۴-۶: نتایج بازسازی رشته‌ی جابجایی الف) تصویر مرجع ب) تصویر تست ج) [71] DSP د) Optical [58] flow ه) [68] Sift flow و) روش پیشنهادی ز) نقشه‌ی احتمال روش پیشنهادی



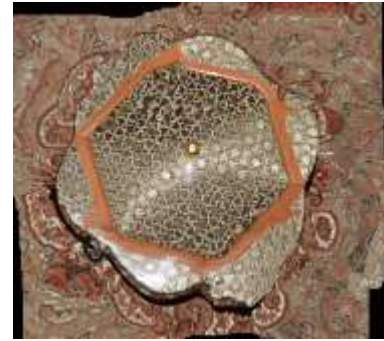
الف



ب



ج



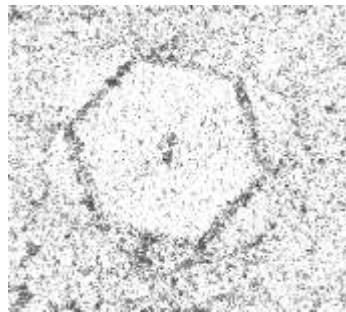
د



ه



و



ز

شکل ۴-۷: نتایج بازسازی رشته‌ی چرخش الف) تصویر مرجع ب) تصویر تست ج) DSP[71] د) Optical Sift flow[68] ه) flow[58] و) روش پیشنهادی ز) نقشه‌ی احتمال روش پیشنهادی



الف



ب



ج



د



ه



و



ز

شکل ۴-۸: نتایج بازسازی رشته‌ی از شکل انداختن الف) تصویر مرجع ب) تصویر تست ج) [71] DSP د) Optical flow [58] ه) Sift flow [68] و) روش پیشنهادی ز) نقشه‌ی احتمال روش پیشنهادی

۵ فصل پنجم: فرا تکیک پذیری احتمالاتی

همانطور که در فصول قبلی بیان شد، بازسازی تصویر وابسته به توانایی تخمین حرکت در صحنه با دقت زیر پیکسلی می‌باشد. اما تخمین حرکت بسیار دقیق یک مسئله‌ی به شدت نامعین می‌باشد. در نتیجه می‌توان گفت الگوریتم‌های SRR^1 کلاسیک توانایی بازسازی فیلم‌های واقعی با محتوای عمومی را ندارند. از آنجایی که تخمین دقیق حرکت با دقت بالا بسیار مشکل می‌باشد در مقالات مختلف تکنیک‌هایی برای کنار گذاشتن تخمین دقیق حرکت پیشنهاد شده است. در این پایان‌نامه سعی شده است نظریه‌ی جابجایی فازی ارائه شده در [۴۳] به گونه‌ای توسعه داده شود که بتواند با تخمین دقیق‌تری از حرکت پیکسل‌ها، SRR احتمالاتی را بهبود ببخشد. در این بخش ابتدا مشکلات و معایب SRR احتمالاتی را بررسی می‌نماییم و سپس با توسعه‌ی سیستم وزن دهی SRR احتمالاتی چهارچوب جدیدی برای فراتفکیک‌پذیری احتمالاتی پیشنهاد خواهیم داد که بتواند بر مشکلات روش‌های پیشین غلبه کند.

در فصل سوم تئوری SRR احتمالاتی به طور مفصل بیان شد. همانطور که مشاهده شد، نویسندگان [۴۳] موفق شدند با تعویض عملگر حرکت با جمع وزن‌دار حرکت‌های سراسری، به معادله‌ای بافرم بسته به شکل زیر دست پیدا کنند:

$$\hat{Z}[i, j] = \frac{\sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m, t}[K, L] y_t[k, L]}{\sum_{[K, L] \in N(i, j)} \sum_{t=1}^T w_{m, t}[K, L]} \quad 1-5$$

در تمام فصل سوم فرض بر این بود که $w_{m, t}[K, L]$ یک مقدار شناخته شده می‌باشد. با توجه به مشکلات تخمین دقیق حرکت، تعیین مقدار دقیق $w_{m, t}[K, L]$ با مشکلات زیادی همراه بود. در مقاله [۴۳] پیشنهاد شد که به جای تعیین دقیق حرکت پیکسل‌ها بین دو فریم متوالی، از احتمال حرکت پیکسل‌ها در داخل یک همسایگی مشخص در فریم‌های دیگر استفاده نماییم. در این مقاله پیشنهاد شد

¹ Super resolution reconstruction

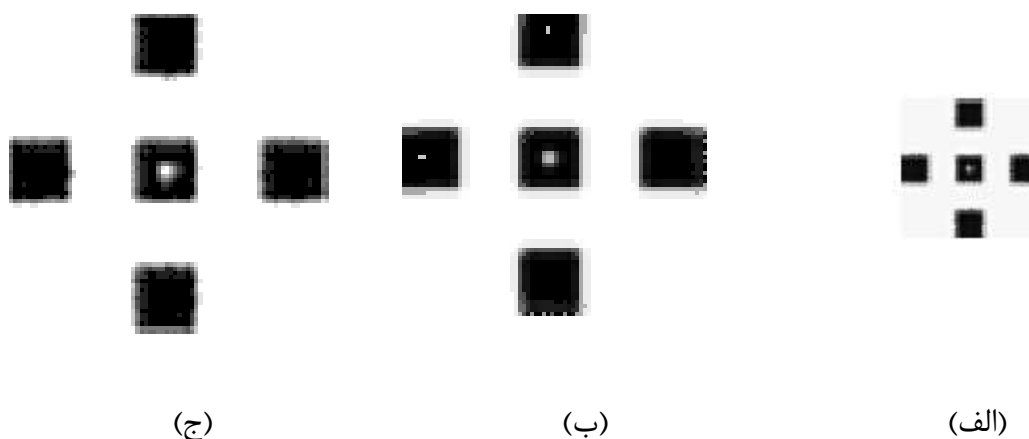
وزن‌ها به نحوی به دست آید که توجیه‌کننده‌ی مدل تصویر باشد. با در نظر گرفتن رابطه‌ی ۲-۳، پیکسل‌های ماتریس Z بعد از جابجایی به وسیله‌ی عملگر F_a و کاهش نمونه به وسیله عملگر D باید حداقل خطا را نسبت به g_t داشته باشند. همچنین پیشنهاد شد برای تعیین دقیق‌تر وزن‌ها، از پشتیبانی فضایی پیکسل‌های همسایه نیز بهره برده شود. بدین منظور به مرکزیت هر پیکسل یک وصله کوچک در نظر گرفته می‌شود و تعیین وزن‌ها به وسیله آن انجام می‌شود. اگر G_{x_j, y_j} عملگر استخراج وصله در مکان $[x_j, y_j]$ با اندازه از پیش تعیین‌شده‌ی $((2k + 1) \times (2k + 1))$ باشد آن‌ها پیشنهاد دادند محاسبه‌ی وزن‌ها به صورت زیر باشد:

$$w_{a,t}[x_j, y_j] = \exp\left\{-\frac{\|G_{x_j, y_j}(DF_a Z - g_t)\|_2^2}{2\sigma^2}\right\} \quad ۲-۵$$

همانطور که مشخص می‌باشد، برای تعیین وزن‌ها فقط از روش‌های وصله‌های تصویر استفاده شده است و هیچ گونه قید هندسی برای حرکت پیکسل‌ها، در این معادله در نظر گرفته نشده است.

فرااتفکیک‌پذیری احتمالاتی وابستگی زیادی به نحوه‌ی تعیین مجموعه‌ی همسایگی $N(i, j)$ برای هر پیکسل و نحوه‌ی تعریف تابع وزن دهی دارد. در عمل همسایگی تاثیر بسیار زیادی روی عملکرد فرااتفکیک‌پذیری احتمالاتی دارد. همسایگی باید آنقدر بزرگ باشد که بتوان از خصوصیات غیر محلی بودن الگوریتم بهره برد. از طرفی افزایش اندازه پنجره‌ی همسایگی محاسبات را به شدت افزایش می‌دهد. تابع وزن دهی یکی دیگر از پارامترهای مهم فرااتفکیک‌پذیری احتمالاتی می‌باشد. با بهبود تابع وزن دهی می‌توان عملکرد فرااتفکیک‌پذیری احتمالاتی را به شدت افزایش داد. اطلاعات تمام پیکسل‌های موجود در همسایگی برای بازسازی یک پیکسل ناشناخته مورد نیاز نمی‌باشد. تابع وزن‌دهی میزان مشارکت هر پیکسل درون همسایگی برای بازسازی پیکسل هدف را تعیین می‌کند. توانایی تابع وزن دهی در مشارکت دادن پیکسل‌های مرتبط با پیکسل هدف و حذف پیکسل‌های نامربوط، تعیین‌کننده کیفیت خروجی

فرا تفکیک پذیری احتمالاتی خواهد بود. این عملیات ارتباط زیادی با ثبت تصویر دارد. زیرا عملیات کلیدی در ثبت تصویر پیدا کردن ارتباط بین نقاط دو تصویر می باشد. از آنجایی که ثبت تصویر با دقت بالا بسیار مشکل می باشد، در این فصل از ثبت احتمالاتی تصویر که در فصل قبل توسعه داده شد استفاده می شود.



شکل ۵-۱: مقایسه تابع وزن دهی روش پیشنهادی و روش فرا تفکیک پذیری احتمالاتی [۴۳].

برای اثبات این ادعا که تابع وزن دهی چقدر می تواند در خروجی الگوریتم فرا تفکیک پذیری موثر باشد در این قسمت آزمایش بسیار ساده ای مطرح شده است. شکل ۵-۱ (الف) یک فریم از ویدیوی وضوح پایین تست را نشان می دهد. این رشته ی ویدیویی شامل چهار فریم وضوح پایین می باشد که نسبت به یکدیگر دارای حرکات انتقالی زیر پیکسلی می باشند. همانطور که پیداست این تصویر شامل پنج مربع سیاه می باشد که در یک زمینه ی سفید قرار گرفته اند. در وسط یکی از مربع ها یک نقطه ی سفید رنگ قرار داده شده است. شکل ۵-۱ (ب) نتیجه ی اعمال الگوریتم ارائه شده در [43] به این ویدیو را نشان می دهد. همانطور که در تصویر خروجی این الگوریتم مشخص است در وسط سه مربع نقطه ی سفید رنگی ایجاد شده است، در حالی که در تصویر ورودی فقط در یکی از مربع ها نقطه ی سفید رنگ قرار داشت. دلیل

اضافه شدن نقاط سفید رنگ به تصویر، نحوه کارکرد تابع وزن دهی الگوریتم [۴۳] می‌باشد. در این الگوریتم از روشنایی پیکسل‌ها برای وزن دهی استفاده می‌شود. از آنجاکه مربع‌ها بسیار مشابه می‌باشند و فقط در یک نقطه متفاوت می‌باشند، پیکسل مرکزی هر کدام از مربع‌ها با وزن زیادی در بازسازی پیکسل مرکزی سایر مربع‌ها نقش آفرینی می‌کند. بنابراین یک نقطه‌ی سفید در مرکز یکی از مربع‌ها باعث بوجود آمدن مصنوعات مزاحم در مرکز سایر مربع‌ها شده است. استفاده از پشتیبانی مکانی سایر پیکسل‌ها در فرایند وزن دهی باعث رفع این مشکل خواهد شد. خروجی روش پیشنهادی در شکل ۵-۱ (ج) نشان داده شده است. همانطور که در شکل پیداست دیگر اثری از مصنوعات مزاحم در شکل دیده نمی‌شود.

ایراد دیگر الگوریتم [۴۳] صافی بیش از حد تصویر خروجی است. نویسندگان این مقاله برای محاسبه‌ی احتمال هر مسیر^۱ از پشتیبانی روشنایی سایر پیکسل‌ها استفاده کرده اند. بنابراین اگر یک پیکسل دارای روشنایی متفاوتی با سایر پیکسل‌ها باشد مقدار آن تغییر خواهد کرد. به عبارتی این الگوریتم پیکسل‌های با مقدار متفاوت در همسایگی را به عنوان نویز در نظر می‌گیرد. بنابراین خروجی بسیار صاف خواهد بود. در این پایان نامه از پشتیبانی روشنایی و پشتیبانی مکانی سایر پیکسل‌های همسایه برای تعیین وزن هر پیکسل مشارکت کننده در فرایند بازسازی استفاده شده است.

اگر فرض کنیم که پیکسل‌های همسایه در یک فریم، در فریم‌های بعد نیز همسایه باقی بمانند، مفهوم پشتیبانی مکانی واضح‌تر خواهد شد. بدیهی است که این فرض در تمام نقاط تصویر بجز نقاط اندکی اطراف مرزهای اشیا و نقاط انسداد برقرار می‌باشد. پشتیبانی مکانی، احتمال حرکت یک پیکسل در یک مسیر، برحسب احتمال حرکت پیکسل‌های همسایه می‌باشد. با اضافه کردن این قید هندسی به فرایند وزن دهی یک پیکسل داریم:

^۱ طبق آنچه در فصل سوم گفته شد، منظور از مسیر مکان‌هایی در فریم‌های دیگر رشته‌ی ویدیویی می‌باشد که احتمال می‌رود پیکسل هدف به آن مکان جایجا شده باشد.

$$w_{a,t}[x_j, y_j] = P_{a,t}^{intensity}[x_j, y_j] \times P_{a,t}^{geometry}[x_j, y_j] \quad ۳-۵$$

به طوری که $P_{a,t}^{intensity}[x_j, y_j]$ و $P_{a,t}^{geometry}[x_j, y_j]$ به ترتیب احتمال مسیر a-ام در t-امین فریم بر پایه ی شباهت روشنایی و پشتیبانی مکانی سایر پیکسل های همسایه می باشد. $P_{a,t}^{intensity}[x_j, y_j]$ با کمی تفاوت در فرایند نرمال سازی، مشابه تابع وزندهی فراتفکیک پذیری احتمالاتی [۴۳] محاسبه می شود، نحوه ی محاسبه ی این احتمال در بخش ۵-۱-۳ و رابطه ی ۵-۱۶ توضیح داده خواهد شد. $P_{a,t}^{geometry}[x_j, y_j]$ پشتیبانی مکانی سایر پیکسل ها از یک جابجایی می باشد و نحوه ی محاسبه ی آن در بخش ۵-۱-۲ و در رابطه ی ۵-۱۰ شرح داده شده است. در حقیقت روش ارائه شده در این قسمت ترکیبی از ثبت احتمالاتی تصویر ارائه شده در فصل پیشین و فراتفکیک پذیری احتمالاتی [۴۳] می باشد (برای کسب اطلاعات بیشتر راجع به این روش به فصل سوم مراجعه نمایید).

۵-۱ روش پیشنهادی

در این بخش روش پیشنهادی برای فراتفکیک پذیری ویدیو با در نظر گرفتن حرکت های محلی بیان می شود. روش پیشنهادی به دو بخش تقسیم شده است. در قسمت اول روش ارائه شده برای ثبت تصاویر وضوح پایین و نسخه ی وضوح بالای اولیه بیان می شود و در قسمت دوم فرآیند بازسازی تصویر وضوح بالا بیان خواهد شد. قسمت نخست این بخش شباهت بسیار زیادی به مطالب گفته شده در فصل چهارم دارد. با این تفاوت که در فصل چهارم هر دو تصویری که نسبت به هم ثبت می شدند هم بعد بودند، اما در این بخش ابعاد دو تصویر با هم متفاوت است. به همین دلیل، تعدادی از روابط بیان شده در فصل قبل برای ثبت احتمالاتی تصویر، در این حالت قابل استفاده نیستند و دچار تغییر می شوند. به منظور جلوگیری از اطاله کلام و خستگی خواننده، فقط تغییرات ایجاد شده در ثبت احتمالاتی تصویر با جزییات بیان شده است و مطالبی از ثبت احتمالاتی تصویر که بدون تغییر باقی می ماند به صورت خلاصه بیان

می‌شود.

ابتدا لازم است جزئیات مشابه در تصاویر وضوح پایین پیدا شود و با توجه به مدل تصویر و فرایند تصویربرداری تصویر وضوح بالا ساخته شود. برای شروع، یک تخمین اولیه از تصویر وضوح بالا در نظر می‌گیریم و به وسیله‌ی جزئیات موجود در تصاویر وضوح پایین آن را به روزرسانی می‌نماییم. تصویر وضوح بالای اولیه از طریق درون‌یابی تصویر وضوح پایین مرجع (فریم وضوح پایینی که قرار است فرایند فراتفکیک پذیری روی آن اعمال شود) به دست می‌آید.

فرض کنید می‌خواهیم تصویر وضوح بالای اولیه‌ی Z را با تصویر وضوح پایین g_t ثبت احتمالاتی نماییم. طبق روندی که در فصل قبل بیان شد، ابتدا لازم است که از g_t و Z ویژگی استخراج شود. در فصل پیشین برای استخراج چگال ویژگی در اطراف هر پیکسل یک وصله با ابعاد $(2k + 1) \times (2k + 1)$ استخراج می‌شد، به طوری که k یک عدد صحیح مثبت می‌باشد. اما تصاویر Z و g_t هم‌بعد نمی‌باشند و مقایسه و پیدا کردن کاندیدا را با مشکل مواجه می‌کند. یک راه این می‌باشد که با درون‌یابی تصویر g_t ، این تصویر با تصویر Z هم‌بعد شوند و سپس آن‌ها را ثبت نماییم. اما در این حالت اندازه وصله‌های تصویر و پنجره‌ی جستجو بزرگ‌تر می‌شود و پیچیدگی محاسباتی الگوریتم بسیار زیاد می‌شود. راه حل دیگر این است که دو وصله با ابعاد متفاوت از این دو تصویر استخراج شود، سپس به وسیله‌ی مدل تصویر ارائه شده در ۲-۳ وصله‌های تصویر Z و g_t به منظور مقایسه به فضای یکسانی برده شوند.

۵-۱-۱ ثبت احتمالاتی تصویر

۵-۱-۱-۱ پیدا کردن نقاط کاندید

در مرحله نخست نیاز به استخراج ویژگی از تصویر وضوح بالای اولیه‌ی Z و تصویر ورودی وضوح

پایین g_t داریم. برای i -امین پیکسل در تصویر وضوح بالای اولیه $(Z[x_i, y_i])$ ، وصله‌ی $p_{x_i, y_i}^Z \in$

$R^{s.(q-1)+1}$ به فرم برداری تبدیل می‌شود. این وصله بعد از عملیات ماتی و کاهش نمونه به بردار ویژگی $f_i^Z \in R^{(2k+1)^2 \times 1}$ تبدیل می‌شود. به صورتی که $q = (2k + 1) \times (2k + 1)$ است و s پارامتر بزرگنمایی می‌باشد.

$$p_{x_i, y_i}^Z = G_{x_i, y_i} Z \quad ۴-۵$$

$$f_i^Z = DHvec(p_{x_i, y_i}^Z) \quad ۵-۵$$

p_{x_i, y_i}^Z وصله‌ی استخراج شده در مکان $[x_i, y_i]$ از تصویر Z می‌باشد. عملگر استخراج وصله با ابعاد $s.(q - 1) + 1$ به مرکز $[x_i, y_i]$ می‌باشد. همچنین $vec(\cdot)$ تابع تبدیل یک وصله به بردار می‌باشد. f_i^Z بردار ویژگی استخراج شده از i -امین پیکسل تصویر Z است. برای هر پیکسل از تصویر g_t نیز ویژگی مناسب استخراج می‌شود.

$$p_{x_j, y_j}^{g_t} = G_{x_j, y_j} g_t \quad ۶-۵$$

$$f_j^{g_t} = vec(p_{x_j, y_j}^{g_t}) \quad ۷-۵$$

به صورتی که $p_{x_j, y_j}^{g_t} \in R^q$ وصله‌ی استخراج شده از تصویر g_t در مکان $[x_j, y_j]$ می‌باشد. $f_j^{g_t} \in R^{(2k+1)^2 \times 1}$ بردار ویژگی استخراج شده از j -امین پیکسل g_t در مکان $[x_j, y_j]$ می‌باشد.

در این مرحله شبیه ترین پیکسل‌ها به i -امین پیکسل تصویر Z ، درون تصویر g_t پیدا می‌شوند. بنابراین برای هر پیکسل $Z[x_i, y_i]$ یک ناحیه جستجو به مرکز $[x_i', y_i']$ در تصویر g_t در نظر گرفته می‌شود. به صورتی که $[x_i, y_i] = ([x_i', y_i'] - 1) \times s + 1$.

اگر C_i مجموعه پیکسل‌های موجود در همسایگی مشخص شده برای $Z[x_i, y_i]$ در تصویر g_t

باشد، با در نظر گرفتن فاصله‌ی اقلیدسی ویژگی f_i^z و ویژگی سایر پیکسل‌های موجود در مجموعه‌ی C_i ، تعداد n تا از نزدیک ترین پیکسل‌ها به پیکسل $Z[x_i, y_i]$ را درون این مجموعه می‌یابیم.

$$\Omega_i = \{d_1, d_2, \dots, d_n\} \quad ۸-۵$$

به طوری که $\Omega_i \subset C_i$ مجموعه‌ی کاندیدهای پیکسل $Z[x_i, y_i]$ می‌باشد.

۲-۱-۵ اعمال Relaxation labelling

تا این مرحله برای هر کدام از پیکسل‌های تصویر Z تعدادی کاندیدا با توجه به شباهت روشنایی انتخاب شده است. در این پایان‌نامه برای اعمال قیود هندسی به پیکسل‌ها و انتخاب بهترین کاندید از Relaxation Labelling استفاده می‌شود. همانند آنچه در فصل قبل بیان شد، ابتدا تمام نقاط لبه را ثبت می‌نماییم. سپس مکان سایر نقاط را با توجه به مکان لبه‌های همسایه (که با دقت بالایی به دست آمده است) تخمین می‌زنیم.

پیکسل‌های موجود در تصویر Z را به دو مجموعه‌ی نقاط لبه V و نقاط هموار U تقسیم می‌نماییم. حال از کاندیدهای استخراج شده برای پیکسل $Z[x_i, y_i]$ موجود در مجموعه‌ی Ω_i به عنوان دانش پیشین استفاده می‌شود تا بهترین انطباق برای داده‌های تست به دست آید.

اگر مجموعه‌ی نقاط همسایه‌ی پیکسل $Z[x_i, y_i] \in V$ به صورت زیر تعریف شود، داریم:

$$N_i = \{Z[x_j, y_j] \mid x_i - r < x_j < x_i + r \ \& \ y_i - r < y_j < y_i + r \ \& \ Z[x_j, y_j] \in V\} \quad ۹-۵$$

به طوری که r پارامتر تعیین سائز همسایگی اطراف پیکسل $Z[x_i, y_i]$ می‌باشد.

حال با استفاده از فرمول به‌روزرسانی احتمال Relaxation labelling احتمال هر کدام از

کاندیدهای $Z[x_i, y_i] \in V$ را به روزرسانی می‌کنیم. برای سادگی بیشتر در فرمول نویسی زیر از Z_i ، به جای $Z[x_i, y_i]$ و از g_α به جای پیکسل α -ام فریم وضوح پایین g_t استفاده شده است.

$$P_{\text{geometry}}^{(n+1)}(Z_i = g_{z_i}) = \frac{P^{(n)}(Z_i = g_{z_i})Q^{(n)}(Z_i = g_{z_i})}{\sum_{g_\lambda \in \Omega_i} P^{(n)}(Z_i = g_\lambda)Q^{(n)}(Z_i = g_\lambda)} \quad 10-5$$

به طوری که $P_{\text{geometry}}^{(n)}(Z_i = g_{z_i})$ احتمال انتساب کاندید $g_{z_i} \in \Omega_i$ به Z_i در تکرار n -ام می‌باشد. تابع $Q(Z_i = g_{z_i})$ مقدار پشتیبانی صورت گرفته از انتساب Z_i به $g_{z_i} \in \Omega_i$ از طرف سایر نودهای موجود در گراف می‌باشد که به صورت زیر تعریف می‌شود.

$$Q^{(n)}(Z_i = g_\alpha) = \sum_{Z_j \in N_i} \sum_{g_\beta \in \Omega_j} P(Z_j = g_\beta) P(A_{ij} | Z_i = g_\alpha, Z_j = g_\beta) \quad 11-5$$

به طوری که Ω_j مجموعه نقاط کاندید پیکسل $Z_j \in N_i$ می‌باشد. $P(A_{ij} | Z_i = g_\alpha, Z_j = g_\beta)$ احتمال انتساب همزمان Z_i به g_α و Z_j به g_β می‌باشد. اگر بردار جابه‌جایی پیکسل Z_i به صورت زیر تعریف شود:

$$u_i = \begin{bmatrix} x_i' \\ y_i' \end{bmatrix} - \begin{bmatrix} x_\alpha \\ y_\alpha \end{bmatrix} \quad 12-5$$

به طوری که اگر $[x_i, y_i]$ مکان پیکسل Z_i و $[x_\alpha, y_\alpha]$ مکان پیکسل متناظر با Z_i در تصویر g_t می‌باشد $[x_i, y_i] = ([x_i', y_i'] - 1) \times s + 1$. بردار جابه‌جایی نقطه‌ی $Z_j \in N_i$ به صورت زیر تعریف می‌شود.

$$u_j = \begin{bmatrix} x_j' \\ y_j' \end{bmatrix} - \begin{bmatrix} x_\beta \\ y_\beta \end{bmatrix} \quad 13-5$$

به طوری که اگر $[x_j, y_j]$ مکان پیکسل $Z_j \in N_i$ می‌باشد و $[x_\beta, y_\beta]$ مکان نقطه‌ی متناظر در تصویر g_t می‌باشد $[x_j, y_j] = ([x_j', y_j'] - 1) \times s + 1$. احتمال همزمانی انتساب به صورت زیر تعریف می‌شود:

$$P(A_{ij}|Z_i = g_\alpha, Z_j = g_\beta) = \exp\left(-\frac{\|u_i - u_j\|_2^2}{2\sigma^2}\right) \quad ۱۴-۵$$

حال که مکان پیکسل‌های لبه‌ی تصویر Z با احتمال بالا پیدا شد لازم است تا مکان سایر نقاط تصویر نیز تخمین زده شود. ابتدا برای هر پیکسل $Z_i \in U$ یک همسایگی به شکل زیر تعریف می‌شود:

$$N_i = \{Z[x_j, y_j] \mid x_i - r < x_j < x_i + r \ \& \ y_i - r < y_j < y_i + r \} \quad ۱۵-۵$$

حال با استفاده از ۱۰-۵ و ۱۱-۵ و ۱۴-۵ می‌توانیم احتمال کاندیدهای نقاط $Z_i \in U$ را به دست آوریم.

۳-۱-۵ بازسازی تصویر وضوح بالا

بعد از چندین تکرار الگوریتم Relaxation Labelling، احتمال نهایی هر کدام از کاندیدها به دست می‌آید. بعد از اینکه تمام فریم‌های وضوح پایین با فریم وضوح بالای اولیه ثبت شدند، تعداد $n \times T$ کاندید برای هر پیکسل موجود در فریم وضوح بالای اولیه به دست می‌آید. به طوری که n تعداد کاندیدهای پیدا شده در هر فریم وضوح پایین می‌باشد و T تعداد فریم‌های وضوح پایین است. نکته قابل توجه این است که در روش پیشنهادی تعداد مسیرهای حرکت هر پیکسل در هر فریم به n مسیر کاهش یافته است، بنابراین در فرایند وزن دهی هر پیکسل فقط $n \times T$ پیکسل شرکت خواهند کرد.

احتمال شباهت روشنایی را به صورت زیر تعریف می‌نماییم:

$$P_{Intensity}^{d,j,t}[x_i, y_i] = \frac{\exp\left\{-\frac{\|f_i^Z - f_j^{gt}\|_2^2}{2\sigma^2}\right\}}{\sum_{d_\lambda \in \Omega_i} \exp\left\{-\frac{\|f_i^Z - f_\lambda^{gt}\|_2^2}{2\sigma^2}\right\}} \quad ۱۶-۵$$

به طوری که f_i^Z ویژگی استخراج شده از پیکسل i -ام تصویر Z می‌باشد و f_j^{gt} ویژگی استخراج

شده از کاندیدای t -ام در تصویر g_t می‌باشد.

حال ۳-۲۰ را طوری بازنویسی می‌نماییم تا با چهارچوب پیشنهادی وفق داده شود:

$$Z[x_i, y_i] = \frac{\sum_{t=1}^T \sum_{d_{\lambda,t} \in \Omega_i^t} w_{d_{\lambda,t}} \times g_{d_{\lambda,t}}}{\sum_{t=1}^T \sum_{d_{\lambda,t} \in \Omega_i^t} w_{d_{\lambda,t}}} \quad ۱۷-۵$$

به طوری که $g_{d_{\lambda,t}}$ مقدار روشنایی کاندید $d_{\lambda} \in \Omega_i^t$ در فریم t -ام می‌باشد، $w_{d_{\lambda,t}}$ به صورت زیر

تعریف می‌شود:

$$w_{d_{\lambda,t}} = P_{Intensity}^{d_{\lambda,t}}[x_i, y_i] \times P_{Geometry, (L)}^{d_{\lambda,t}}[x_i, y_i] \quad ۱۸-۵$$

$P_{Geometry, (L)}^{d_{\lambda,t}}[x_i, y_i]$ پشتیبانی فضایی به دست آمده برای کاندید $d_{\lambda} \in \Omega_i^t$ در فریم t -ام

بعد از L بار تکرار الگوریتم Relaxation Labelling می‌باشد.

۵-۲ نتایج تجربی

در این بخش عملکرد الگوریتم پیشنهادی با سایر روش‌های ارتقای وضوح موجود مقایسه می‌شود. بازدهی روش پیشنهادی بر روی رشته‌های ویدیویی Foreman، MissAmerica و Suzie بررسی می‌شود. هر کدام از این رشته‌های ویدیویی شامل ۳۰ فریم می‌باشند. تصویر ورودی به‌عنوان نسخه‌ی وضوح‌بالا در نظر گرفته‌شده و برای ساخت نسخه‌ی وضوح پایین، ابتدا تصاویر وضوح‌بالا با تابع نقطه گستر 3×3 مات می‌شوند و سپس با فاکتور وضوح ۳:۱ در هر دو محور افقی و عمودی کاهش نمونه می‌یابند. سپس نویز گوسی با انحراف معیار ۲ به هر کدام از فریم‌ها اضافه می‌شود. رشته‌های ویدیویی تست فقط شامل حرکت اشیا می‌باشند و دوربین در تمام رشته‌ها ثابت می‌باشد. برای بازسازی فریم t -ام تمام فریم‌های ۱:۳۰ در فرایند بازسازی شرکت داده می‌شوند. در تمام تست‌ها پارامتر $\sigma = 2.2$ در نظر گرفته شده

است. بیشترین تکرار الگوریتم فراتفکیک پذیری دو مرتبه می‌باشد و بیشترین تکرار الگوریتم برچسب گذاری آزاد ۲۰ بار می‌باشد. در این پایان نامه از معیار PSNR^۱ به‌عنوان معیار سنجش کیفیت تصاویر تخمین زده‌شده استفاده شده است.

معیار PSNR به‌صورت زیر محاسبه می‌شوند:

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2 m}{\|\hat{X} - X\|^2} \right) \quad ۱-۵$$

که m تعداد کل پیکسل‌های تصویر HR است. \hat{X} و X به ترتیب تصویر HR تخمین زده‌شده و تصویر اصلی می‌باشند.

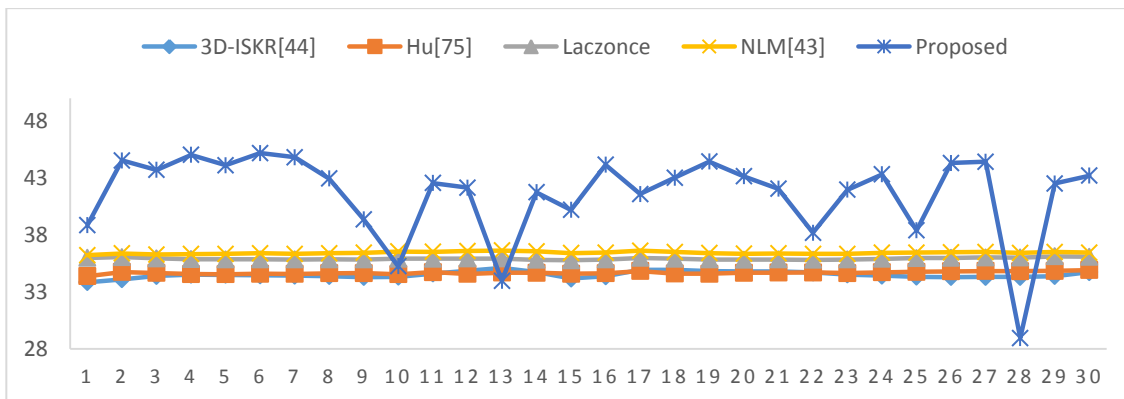
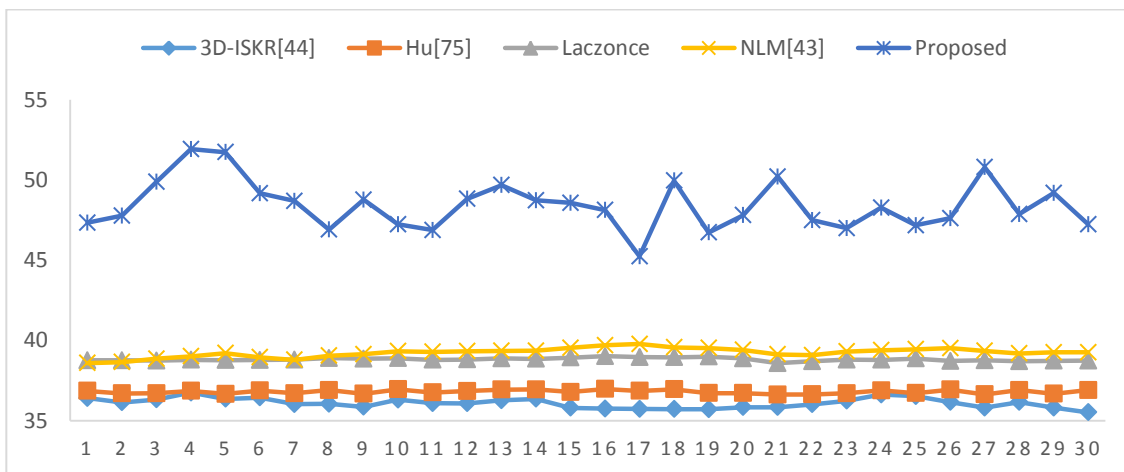
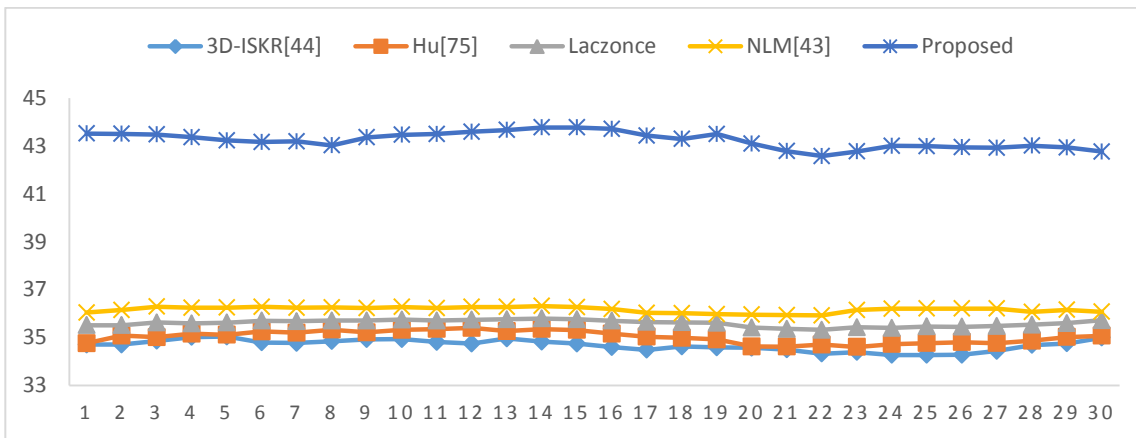
پارامتر PSNR به‌تنهایی نمی‌تواند معیار مناسبی برای بررسی عملکرد الگوریتم‌های SRR باشد و مقایسه بصری نیز باید انجام شود. برای مشاهده بهتر نتایج در هر بخش تنها قسمتی از تصاویر نشان داده شده است. روش پیشنهادی با روش درون‌یابی Laczonce و همچنین چندین روش SRR همانند روش‌های [۴۳]، [۴۴] و [۷۵] مقایسه شده است. تصویر ورودی با فاکتور وضوح ۳ بزرگ شده است. در تصویر وضوح‌پایین وصله‌ها با ابعاد ۵×۵ می‌باشند و در تصویر وضوح‌بالا ابعاد وصله‌ها ۱۳×۱۳ استخراج شده‌اند. شکل ۲-۵ مقدار PSNR فریم به فریم روش پیشنهادی و سایر روش‌ها را برای رشته‌های ویدیویی تست نشان می‌دهد. همانطور که مشخص است روش پیشنهادی در حالت کلی از سایر روش‌ها بهتر عمل می‌نماید. در جدول ۱-۵ نتایج مقایسه متوسط PSNR تمام ۳۰ فریم تولید شده توسط روش پیشنهادی و سایر روش‌ها نشان داده شده است. همان‌طور که در این جدول مشاهده می‌شود از نظر PSNR نسبت به سایر روش‌ها نتایج بهتری کسب شده است. در شکل ۳-۵ تا شکل ۹-۵ خروجی روش

^۱Peak signal to noise ratio

پیشنهادی و سایر روش ها به منظور مقایسه‌ی بصری آورده شده است. همانطور که مشخص است درونیایی Laczonce تعداد پیکسل‌ها را افزایش می‌دهد اما جزییاتی به تصویر اضافه نمی‌نماید. اگرچه روش [۷۵] نسبت به نویز مقاوم می‌باشد اما در تصاویر خروجی این روش، لبه‌ها بسیار صاف می‌باشند. نتایج [۴۴] در نواحی اطراف لبه کمی صاف می‌باشد. [۴۳] در نواحی از تصویر که وصله‌های مشابه کمی وجود دارد اثرات بلوکی تولید نموده است. از لحاظ بصری کاملاً مشخص است که روش پیشنهادی لبه‌های تیزتری تولید می‌کند و اثرات مصنوعی کمتری دارد.

جدول ۵-۱: مقایسه PSNR روش پیشنهادی با سایر روش ها

میانگین	ویدیوی Suzie	ویدیوی Miss-America	ویدیوی Foreman	الگوریتم
۳۶.۷۷	۳۵.۹۲	۳۸.۸	۳۵.۶	Lanczos
۳۵.۵۰	۳۴.۶۹	۳۶.۸۱	۳۵.۰۲	الگوریتم [۷۵]
۳۵.۱	۳۴.۵۲	۳۶.۱۰	۳۴.۶۸	الگوریتم [۴۴]
۳۷.۲۸	۳۶.۴۴	۳۹.۲۵	۳۶.۱۶	الگوریتم [۴۳]
۴۴.۴۳	۴۱.۶۱	۴۸.۴۳	۴۳.۲۵	روش پیشنهادی



شکل ۵-۲: مقایسه PSNR فریم به فریم رشته های ویدیویی تست بالا (Foreman) و وسط (Miss-America) پایین (Suzie)



الف

ب

ج

د



ه

و

ز

شکل ۳-۵: مقایسه بصری فریم اول رشته‌ی ویدیویی Foreman. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونبایی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی



الف

ب

ج

د



ه

و

ز

شکل ۴-۵: مقایسه بصری فریم هشتم رشته‌ی ویدیویی Foreman. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونیابی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی



الف

ب

ج

د



ه

و

ز

شکل ۵-۵: مقایسه بصری فریم سیزدهم رشته‌ی ویدیویی Foreman. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونیابی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی

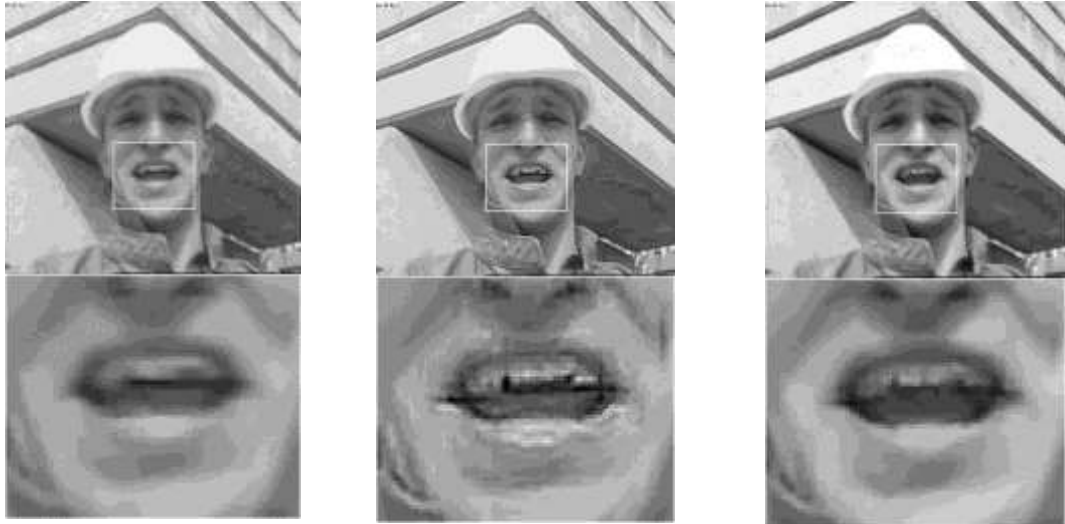


الف

ب

ج

د



ه

و

ز

شکل ۵-۶: مقایسه بصری فریم بیست و دوم رشته‌ی ویدیویی Foreman. (الف) تصویر وضوح پایین (ب) تصویر اصلی وضوح بالا (ج) درونیایی Lanczos (د) [۷۵] (ه) [۴۴] (و) [۴۳] (ز) روش پیشنهادی

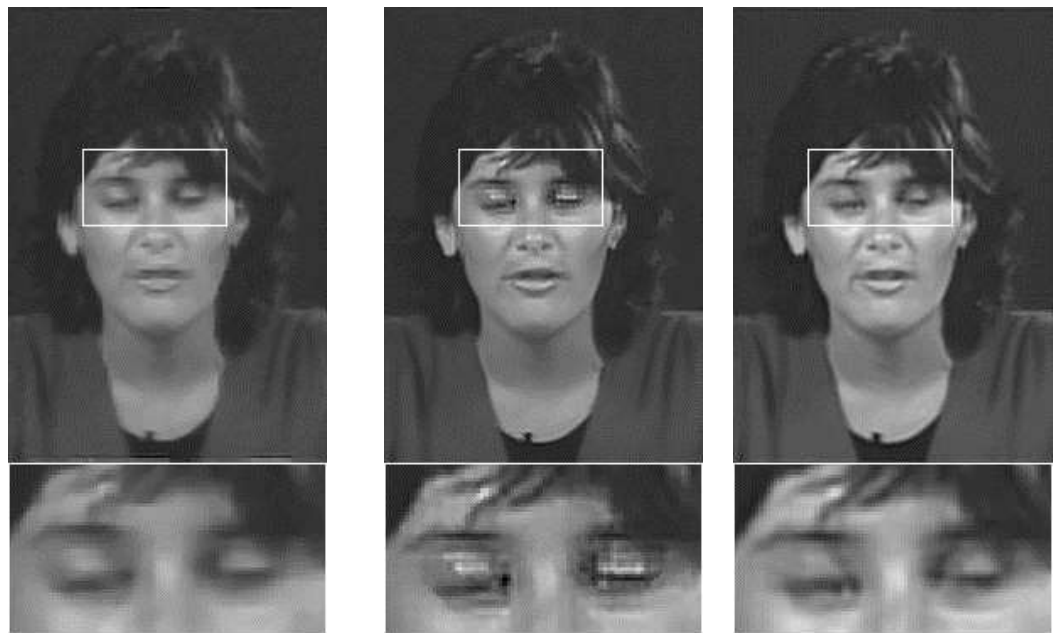


الف

ب

ج

د

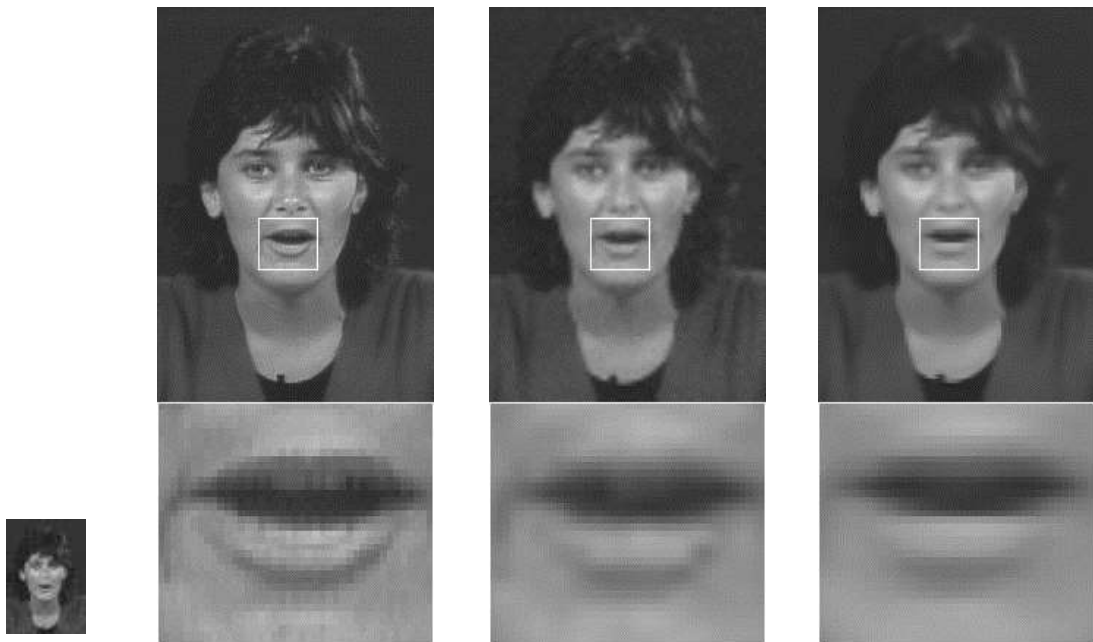


ه

و

ز

شکل ۵-۷: مقایسه بصری فریم هشتم رشته‌ی ویدیویی Miss-America. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونیایی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی

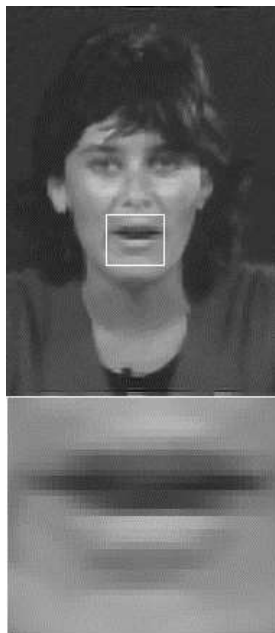


الف

ب

ج

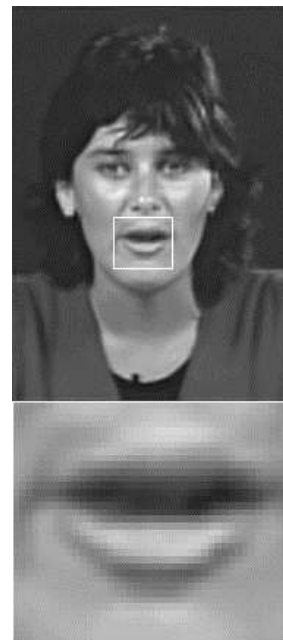
د



ه

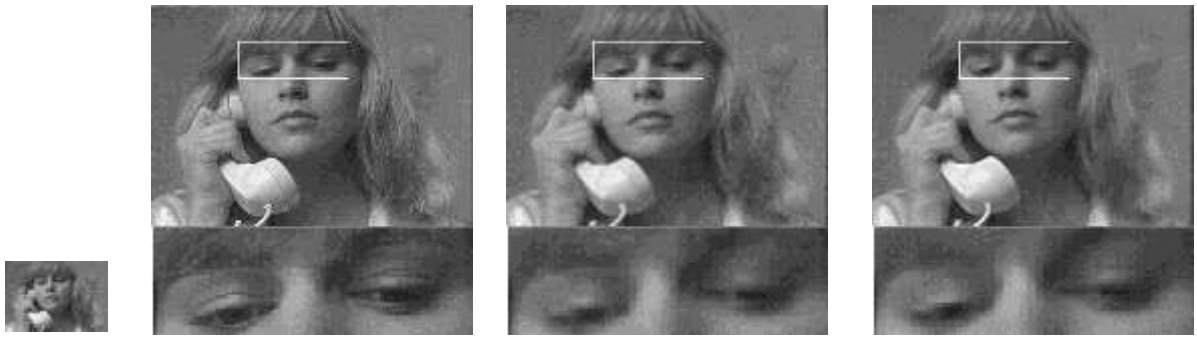


و



ز

شکل ۵-۸: مقایسه بصری فریم بیستم رشته‌ی ویدیویی Miss-America. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونیایی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی



الف

ب

ج

د



ه

و

ز

شکل ۵-۹: مقایسه بصری فریم اول رشته‌ی ویدیویی Suzie. الف) تصویر وضوح پایین ب) تصویر اصلی وضوح بالا ج) درونیایی Lanczos د) [۷۵] ه) [۴۴] و) [۴۳] ز) روش پیشنهادی

۶ نتیجه گیری

ما یک روش جدید برای افزایش وضوح تصویر بر پایه‌ی Relaxation labelling ارائه دادیم. در روش ارائه شده تعدادی نقطه‌ی کاندید برای هر کدام از پیکسل‌های تصویر وضوح‌بالای اولیه در تصاویر وضوح پایین به دست می‌آوردیم. سپس با اعمال Relaxation labelling نقاط کاندید بهینه را به دست آوردیم. با ترکیب وزن‌دار نقاط کاندیدا تصویر وضوح‌بالای نهایی به دست خواهد آمد. نتایج تجربی نشان می‌دهد که روش پیشنهادی در مقایسه با سایر روش‌ها دقت تخمین بالاتری دارد. ما باور داریم که روش ارائه شده می‌تواند در کاربردهای مختلف فراتفکیک‌پذیری ویدیو مانند ویدیوهای نظارتی به کاربرده شود.

پیشنهادهایی برای کارهای آتی:

در طی پیاده‌سازی روش‌های ارتقای وضوح تصویر، با چالش‌هایی روبرو شدیم که با برطرف کردن آن می‌توانیم روش پیشنهادی را بهبود دهیم. در ادامه به توضیح این مسائل می‌پردازیم.

- با توجه به چالش‌های موجود در ثبت تصویر با افزایش دقت ثبت تصویر، پیچیدگی الگوریتم افزایش خواهد یافت. بنابراین ارائه‌ی الگوریتمی سریع در حوزه‌ی ثبت تصویر برای سرعت بخشیدن به فرآیند فراتفکیک‌پذیری ضروری به نظر می‌رسد.
- اکثر الگوریتم‌های فراتفکیک‌پذیری ویدیو خود را به فاکتور وضوح ۳ محدود کرده‌اند. بزرگنمایی تصویر با فاکتوری بزرگتر از ۳ می‌تواند استفاده‌های متعددی در حوزه‌های قضایی و امنیتی داشته باشد.
- در حوزه فراتفکیک‌پذیری تاکنون مقالات زیادی منتشر شده است. با این حال هنوز معیار کاملاً مناسبی جهت ارزیابی عملکرد الگوریتم‌ها ارائه نشده است. مقدار PSNR یکی از رایج‌ترین معیارهای سنجش

عملکرد الگوریتم‌های SR است، که به‌طور قطعی نمی‌تواند بیانگر کیفیت مناسب تصویر بازسازی باشد. بنابراین ارائه یک معیار سنجش مناسب در این حوزه مورد نیاز است.

- [١] S. Chaudhuri and M. V. Joshi, "Research on image super-resolution," *Motion-Free Super-Resolution*, pp. 15-31, 2005.
- [٢] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE signal processing magazine*, vol. 20, pp. 21-36, 2003.
- [٣] J. Cui, Y. Wang, J. Huang, T. Tan, and Z. Sun, "An iris image synthesis method based on PCA and super-resolution," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, 2004, pp. 47.٤٧٤-١
- [٤] T. Gotoh and M. Okutomi, "Direct super-resolution and registration using raw CFA images," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, pp. II-600-II-607 Vol. 2.
- [٥] H. Nyquist, "Certain topics in telegraph transmission theory," *Proceedings of the IEEE*, vol. 90, pp. 280-305, 2002.
- [٦] H. H. Barrett and K. J. Myers, *Foundations of image science*: John Wiley & Sons, 2013.
- [٧] S. Borman, "Topics in multiframe superresolution restoration," 2004.
- [٨] R. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in computer vision and Image Processing*, vol. 1, pp. 317-339, 1984.
- [٩] S. Farsiu, "A fast and robust framework for image fusion and enhancement," Citeseer, 2005.
- [١٠] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE transactions on image processing*, vol. 6, pp. 1646-1658, 1997.
- [١١] E. Mjolsness, "Fingerprint Hallucination," California Institute of Technology, 1985.
- [١٢] D. Kong, M. Han, W. Xu, H. Tao, and Y. Gong, "A conditional random field model for video super-resolution," in *18th International Conference on Pattern Recognition (ICPR'06)*, 2006, pp. 6.٦٢٢-١٩
- [١٣] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 349-356.
- [١٤] N. Nguyen and P. Milanfar, "A wavelet-based interpolation-restoration method for superresolution (wavelet superresolution)," *Circuits, Systems and Signal Processing*, vol. 19, pp. 321-338, 2000.
- [١٥] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouky, B. M. Salam, and F. A. El-Samie, "Regularized super-resolution reconstruction of images using wavelet fusion," *Optical Engineering*, vol. 44, pp. 097001-097001-10, 2005.
- [١٦] H. Ji and C. Fermüller, "Wavelet-based super-resolution reconstruction: theory and

- algorithm," in *European Conference on Computer Vision*, 2006, pp. 295-307.
- [17] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 181-186, 1992.
- [18] N. K. Bose and N. A. Ahuja, "Superresolution and noise filtering using moving least squares," *IEEE Transactions on Image Processing*, vol. 15, pp. 2239-2248, 2006.
- [19] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing*, vol. 6, pp. 1099-1106, 1997.
- [20] B. C. Tom and A. K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images," in *Image Processing, 1995. Proceedings.*, pp. 539-542.
- [21] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, vol. 6, pp. 1621-1633, 1997.
- [22] J. Tian and K.-K. Ma, "Stochastic super-resolution image reconstruction," *Journal of Visual Communication and Image Representation*, vol. 21, pp. 232-244, 2010.
- [23] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, pp. I-I.
- [24] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International journal of computer vision*, vol. 40, pp. 25-47, 2000.
- [25] L. C. Pickup, S. J. Roberts, and A. Zisserman, "A sampled texture prior for image super-resolution," in *Advances in neural information processing systems*, 2003.
- [26] S. Rhee and M. G. Kang, "Discrete cosine transform based regularized high-resolution image reconstruction algorithm," *Optical Engineering*, vol. 38, pp. 1348-1356, 1999.
- [27] J. Tian and K.-K. Ma, "A survey on super-resolution imaging," *Signal, Image and Video Processing*, vol. 5, pp. 329-342, 2011.
- [28] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical models and image processing*, vol. 53, pp. 231-239, 1991.
- [29] K. V. Suresh and A. N. Rajagopalan, "Robust and computationally efficient superresolution algorithm," *JOSA A*, vol. 24, pp. 984-992, 2007.
- [30] S. P. Belekos, N. P. Galatsanos, and A. K. Katsaggelos, "Maximum a posteriori video super-resolution using a new multichannel image prior," *IEEE Transactions on Image Processing*, vol. 19, pp. 1451-1464, 2010.
- [31] H. Shen, L. Zhang, B. Huang, and P. Li, "A MAP approach for joint motion estimation, segmentation, and super resolution," *IEEE Transactions on Image processing*, vol. 16, pp. 479-490, 2007.
- [32] S. Z. Li, *Markov random field modeling in computer vision*: Springer Science &

- Business Media, 2012.
- [۳۳] H. Rue and L. Held, *Gaussian Markov random fields: theory and applications*: CRC Press, 2005.
 - [۳۴] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE transactions on image processing*, vol. 5, pp. 996-1011, 1996.
 - [۳۵] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin, "Image analogies," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 327-340.
 - [۳۶] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, pp. 1838-1857, 2011.
 - [۳۷] C. Chen and J. E. Fowler, "Single-image super-resolution using multihypothesis prediction," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, 2012, pp. 608-612.
 - [۳۸] A. N. Tikhonov and V. Y. Arsenin, "Solutions of ill-posed problems. ۱۹۷۷ ",
 - [۳۹] Z. Zhu, F. Guo, H. Yu, and C. Chen, "Fast single image super-resolution via self-example learning and sparse representation," *IEEE Transactions on Multimedia*, vol. 16, pp. 2178-2190, 2014.
 - [۴۰] M. Protter, M. Elad, H. Takeda, and P. Milanfar " ,Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on image processing*, vol. 18, pp. 36-51, 2009.
 - [۴۱] A. Buades, B. Coll, and J.-M. Morel, "Denoising image sequences does not require motion estimation," in *IEEE Conference on Advanced Video and Signal Based Surveillance, 2005.*, 2005, pp. 70-74.
 - [۴۲] J. Boulanger, C. Kervrann, and P. Bouthemy, "Space-time adaptation for patch-based image sequence restoration," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, pp. 1096-1102, 2007.
 - [۴۳] M. Protter and M. Elad, "Super resolution with probabilistic motion estimation," *IEEE Transactions on Image Processing*, vol. 18, pp. 1899-1904, 2009.
 - [۴۴] H. Takeda, P. Milanfar, M. Protter, and M. Elad, "Super-resolution without explicit subpixel motion estimation," *IEEE Transactions on Image Processing*, vol. 18, pp. 1958-1975, 2009.
 - [۴۵] W.-s. Yu, M.-h. Wang, H.-w. Chang, and S.-q. Chen, "A Fast Kernel Regression Framework for Video Super-Resolution," *TIIS*, vol. ۱, pp. 232-248, 2014.
 - [۴۶] N. Barzigar, A. Roozgard, P. Verma, and S. Cheng, "A Video Super-Resolution Framework Using SCoBeP," *IEEE Transactions on circuits and systems for video technology*, vol. 26, pp. 264-277, 2016.
 - [۴۷] T. Kato, H. Hino, and N. Murata, "Multi-frame image super resolution based on sparse coding," *Neural Networks*, vol. 66, pp. 64-78, 2015.
 - [۴۸] Y. Zhu, Y. Zhang, and J. Sun, "Super-Resolution of Video Using Deformable

- Patches," in *International Conference on Intelligent Science and Big Data Engineering*, 2015, pp. 647-656.
- [٤٩] Y. Zhu, Y. Zhang, and A. L. Yuille, "Single image super-resolution using deformable patches," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2917-2924.
- [٥٠] S. Farsiu ,D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, pp. 47-57, 2004.
- [٥١] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259-268, 1992.
- [٥٢] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing*, vol. 10, pp. 1187-1193, 2001.
- [٥٣] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, pp. 977-1000, 2003.
- [٥٤] S. Avidan, "Ensemble tracking," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, pp. 261-271, 2007.
- [٥٥] S. Bagon, O. Boiman, and M. Irani, "What is a good image segment? a unified approach to segment extraction," in *European Conference on Computer Vision*, 2008, pp. 30-44.
- [٥٦] J. Kim and K. Grauman, "Asymmetric region-to-image matching for comparing images with generic object categories," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010, pp. 2344-2351.
- [٥٧] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, pp. 1615-1630, 2005.
- [٥٨] T. Brox and J. Malik, "Large displacement optical flow: descriptor matching in variational motion estimation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, pp. 500-513, 2011.
- [٥٩] I. Simon and S. M. Seitz, "A probabilistic model for object recognition, segmentation, and non-rigid correspondence," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-7.
- [٦٠] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, pp. 91-110, 2004.
- [٦١] O. Boiman and M. Irani, "Detecting irregularities in images and in video," *International Journal of Computer Vision*, vol. 74, pp. 17-31, 2007.
- [٦٢] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1-8.
- [٦٣] B. D .Lucas and T. Kanade, "An iterative image registration technique with an

- application to stereo vision," in *IJCAI*, 1981, pp. 674-679.
- [१ॢ] C. Liu, "Beyond pixels: exploring new representations and applications for motion analysis," Citeseer, 2009.
 - [१ॣ] D. Tell, "Wide baseline matching with applications to visual servoing," in *PHD THESIS, COMPUTATIONAL VISION AND ACTIVE PERCEPTION LABORATORY (CVAP), ROYAL INSTITUTE OF TECHNOLOGY*, 2002.
 - [१।] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *European conference on computer vision*, 2004, pp. 25-36.
 - [१॥] O. Enqvist, K. Josephson, and F. Kahl, "Optimal correspondences from pairwise constraints," in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 1295-1302.
 - [१०] C. Liu, J. Yuen, and A. Torralba, "Sift flow: Dense correspondence across scenes and its applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, pp. 978-994, 2011.
 - [११] A. Roozgard, N. Barzigar, S. Cheng, and P. Verma, "Dense image registration using sparse coding and belief propagation," in *Signal Processing and Communication Systems (ICSPCS), 2011 5th International Conference on*, 2011, pp. 1-5.
 - [१ॡ] X. Qing, J. Shen, X. Li, and Y. Jia, "A new sparse feature-based patch for dense correspondence," in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, 2014, pp. 1-6.
 - [१Ӏ] J. Kim, C. Liu, F. Sha, and K. Grauman, "Deformable spatial pyramid matching for fast dense correspondences," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 2307-2314.
 - [१ӑ] A. Ahmadyfard and J. Kittler, "Region-based representation for object recognition by relaxation labelling," in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, 2000, pp. 297-307.
 - [१Ӓ] A. Rosenfeld, R. A. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 420-433, 1976.
 - [१ӓ] J. Kittler and J. Illingworth, "Relaxation labelling algorithms—a review," *Image and Vision Computing*, vol. 3, pp. 206-216, 1985.
 - [१Ӕ] J. Hu and Y. Luo, "Noise-robust video super-resolution using an adaptive spatial-temporal filter," *Multimedia Tools and Applications*, vol. 74, pp. 9259-9278, 2015.

Abstract:

Super-resolution (SR) is a promising technique of digital imaging which attempts to generate a raster image with a high resolution from low resolution input image(s). This thesis focuses on multi-frames super resolution. In interpolation methods increasing resolution is performed by up sampling the image. These methods only increase number of pixels of image and missing details are not reconstructed. Super-resolution methods are proposed for solving the problem of interpolation methods. Super-resolution technique not only improves the quality of under sampled, low resolution images by increasing their spatial resolution but also attempts to filter out distortion. Classic SR methods do not produce a good result on actual videos, because these methods require accurate motion estimation. New methods are introduced in video super resolution that can produce effective results without accurate motion estimation, recently. SR method based on NLM filter use fuzzy motion instead of exact motion estimation. In this thesis we propose a new method based on probability. In proposed method, by introducing the concept of probabilistic registration, fuzzy motion that has been introduced in earlier work is improve. In the proposed method using other video frames we can restore the missing high resolution details. Based on proposed method user can overcome problems in imaging systems using moderate cost hardware. Experimental results show that the proposed method can remove artificial effects and produce sharp edges. The proposed method in terms of PSNR criteria is better than the other methods introduced in this thesis.

Key words: super resolution, relaxation labelling, probabilistic registration



Shahrood University of Technology

Faculty of Electrical Engineering and Robotic

M.Sc. Thesis in Digital Electronic Systems Engineering

Video Super Resolution without Explicit Motion Estimation

By:

Mahmood Amiri

Supervisor:

Dr. Alireza Ahmadyfard

Advisor:

Dr. Vahid Abolghasemi

January 2017