دانشکده مهندسی برق و رباتیک

گروه الکترونیک

پایان‌نامه کارشناسی ارشد

افزایش وضوح تصویر به کمک یادگیری جزئیات از سطوح مختلف هرم تصویر

آزاده مکاری

استاد راهنما:

دکتر علیرضا احمدی فرد

بهمن ۹۴

تقدیم به:

خانواده‌ی عزیزم

تشکر و قدردانی

نخستین سپاس و ستایش از آن خداوندی است که بنده کوچکش را در دریای بیکران اندیشه، قطره‌ای ساخت تا وسعت آن را از دریچه اندیشه‌های ناب آموزگارانی بزرگ به تماشا نشیند. لذا اکنون که در سایه سار بنده‌نوازی‌هایش پایان‌نامه حاضر به انجام رسیده است، بر خود لازم می‌دانم مراتب سپاس و قدردانی را از بزرگوارانی به جا آورم که اگر دست یاریگرشان نبود، هرگز این پایان‌نامه به انجام نمی‌رسید.

در ابتدا از استاد عزیزم، جناب آقای دکتر علیرضا احمدی‌فرد، که دلسوزانه مرا در انجام این پروژه یاری کردند و از راهنمایی‌ها و تجربیاتشان بهره‌های فراوان بردم تشکر ویژه دارم.

سپاس آخر را به مهربانترین همراهان زندگیم، به پدر و مادر و برادر عزیزم تقدیم می‌کنم که حضورشان در فضای زندگیم مصداق بی ریای سخاوت بوده است.

تعهد نامه

اینجانب آزاده مکاری دانشجوی دوره کارشناسی ارشد رشته برق دانشکده برق دانشگاه صنعتی شاهرود نویسنده پایان نامه افزایش وضوح تصویر به کمک یادگیری جزئیات از سطوح مختلف هرم تصویر تحت راهنمائی دکتر علیرضا احمدی فرد متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای 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 تفکیک‌پذیری فنی در پردازش تصویر است که سعی در بازسازی تصویری با وضوح‌بالا (HR[1]) از روی تصویر یا تصاویر وضوح پایین (LR[2]) ورودی دارد. این پایان‌نامه بر روی فرا تفکیک‌پذیری تک فریم تمرکز دارد . در روش‌های مبتنی بر درون‌یابی افزایش ابعاد تصویر به‌وسیله‌ی نمونه‌برداری مجدد انجام می شود. در این روش‌ها جزئیاتی که در نتیجه‌ی فرایند کاهش ابعاد ازدست‌رفته است، بازسازی نمی شود. برای اینکه خرابی‌های به وجود آمده در فرایند افزایش ابعاد تصویر را از بین ببریم یک روش جدید فرا تفکیک‌پذیری مبتنی بر خود یادگیری ارائه می‌گردد. درروش پیشنهادی، برای یادگیری رابطه‌ی میان وصله‌های وضوح‌بالا و پایین بدون نیاز به داشتن پایگاه داده خارجی و فقط با استفاده از هرم تصویر می‌توانیم جزئیات فرکانس بالای ازدست‌رفته‌ی تصویر را بازسازی کنیم. بر پایه روش ارائه‌شده در این پایان‌نامه به کاربر این امکان داده می‌شود بدون پرداخت هزینه برای خرید سخت‌افزار بهتر، بر مشکلات موجود در سیستم‌های تصویربرداری غلبه کند و از طرفی نیازمند جمع‌آوری پایگاه داده خارجی که شامل تصاویر وضوح‌بالا و پایین است نمی‌باشد. روش پیشنهادی یک روش سریع است به‌طوری‌که به متوسط زمان ۵ ثانیه برای تولید خروجی وضوح‌بالا نیاز دارد. همچنین روش پیشنهادی ازلحاظ PSNR به نتایج رضایت بخشی رسیده است. 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 می دهیم که روش پیشنهادی ازنقطه‌نظر سرعت و معیار PSNR و SSIM بهتر از سایر روش‌های معرفی‌شده در این پایان‌نامه عمل می کند.

واژه‌های کلیدی: فرا تفکیک‌پذیری، تنظیم‌کننده‌ی تیخونوف، دیکشنری، هرم تصویر، تنظیم کننده BTV

---

[1] High-Resolution

[2] Low-Rseloution

لیست مقالات مستخرج از پایان‌نامه:

1. Azade Mokari, Alireza Ahmadyfard, “A fast method for single image super resolution using dictionary learning”, the first international conference on signal processing and intelligent systems, Amirkabir University of technology, December 2015.
2. Azade Mokari, Alireza Ahmadyfard, “An adaptive single image method for super resolution”, the first international conference on signal processing and intelligent systems, Amirkabir University of technology, December 2015.
3. Azade Mokari, Alireza Ahmadyfard, “Fast single image super resolution via dictionary learning”, Submitted to Image processing, IET Jornal.

فهرست مطالب

فهرست شکل‌ها

فهرست جداول

۱ فصل اول: مقدمه

جهان پیشرفت‌های شگرفی در زمینه‌ی سخت‌افزار و نرم‌افزار به خود دیده است. تصاویر دیجیتال امروزه در همه جا حضور دارند در وب، بر روی سیستم‌های ماهواره‌ای و غیره. 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 اطلاعات مفید از تصاویر به کاربر کمک می‌کند. پردازش تصویر در چند دهه اخیر رشد قابل‌توجهی در زمینه‌ی سرعت و کاهش هزینه داشته است. همچنین در ساخت حس‌گرهای دوربین نیز پیشرفت‌های زیادی به وجود آمده است تا دوربین‌های دیجیتال باکیفیت بالایی تولید شود. 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 وضوح‌بالا نیازمند هستند. برآورده کردن نیازهای مطرح‌شده برای داشتن تصویری با وضوح‌بالا سبب به وجود آمدن فن‌های پردازش تصویر شدند. موضوع فرا تفکیک‌پذیری اولین بار در سال ۱۹۸۴ توسط Tsai و همکارانش[۱] مطرح شد. درون‌یابی تصویر یک موضوع قابل تحقیق و جذاب درزمینه‌ی پردازش تصویر است. در سیستم‌های دیجیتال یک سیگنال توسط تعداد محدودی از نمونه‌ها ارائه می‌شود. درون‌یابی چنین سیگنال‌هایی توسط نمونه‌هایی که مشخص هستند، انجام می‌گیرد. چنین مسئله‌ای یک مسئله بدحالت است چون سیگنال‌های بسیاری وجود دارند که ممکن است نمونه‌های یکسانی داشته باشند. در بیشتر موارد فرض بر این است که سیگنال‌ها باند محدود هستند. درون‌یابی در صورتی امکان‌پذیر است که فرآیند نمونه‌برداری داده‌ها از شرط نایکوئیست پیروی کند. این شرط ممکن است که همیشه برقرار نباشد. فن‌های فرا تفکیک‌پذیری برای افزایش وضوح فضایی استفاده می‌شوند. فرا تفکیک‌پذیری فرآیندی است که سبب افزایش وضوح یک تصویر می‌شود. درواقع مقادیر نمونه‌های جدید توسط برخی مقادیر نمونه‌های سیگنال که در دسترس هستند تخمین زده می‌شود، به‌طوری‌که سیگنال جدید جزئیات بیشتر و اطلاعات بهتری را فراهم می‌آورد.

---

[۱] Target detection

۱-۱ وضوح تصویر

وضوح بصری[۱] معیاری برای توانایی سیستم تصویربرداری یا جزئی از سیستم تصویربرداری (دوربین) است که جزئیات تصویر را نشان می‌دهد. به‌عبارت‌دیگر وضوح تصویر کوچک‌ترین جزئیاتی است که در یک تصویر قابل‌مشاهده است. وضوح یک موضوع بنیادی در قضاوت کیفیت سیستم پردازش یا وسیله به دست آوردن تصویر می‌باشد. در ساده‌ترین حالت وضوح تصویر به‌عنوان کوچک‌ترین جزئیات قابل‌اندازه‌گیری در نمایش بصری تعریف می‌شود. در حوزه پردازش تصاویر دیجیتال و ویدیو مفهوم وضوح به یکی از سه تعریف ذیل مربوط می‌شود [۲]:

وضوح فضایی: هر تصویر از کوچک‌ترین عنصر به نام پیکسل ساخته‌شده است. وضوح فضایی به تعداد پیکسل‌ها در یک تصویر مربوط می‌شود و با معیار پیکسل بر طول اندازه‌گیری می‌شود. هرچه وضوح فضایی یک پیکسل 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. این قابل‌بیان است که تعداد سطوح کوانتیزاسیون به نرخ نمونه‌برداری فضایی مربوط می‌شود. اگر حس‌گر دوربین سطوح کوانتیزاسیون کمتری داشته باشد باید نمونه‌برداری با نرخ بالاتری انجام شود تا بتواند روشنایی صحنه را تسخیر کند.

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

---

[۱] Optical -Resolution

نرخ فریم‌های معمولی برای یک ویدیو در حدود ۲۵ فریم در ثانیه می‌باشد[۲].

در این پایان‌نامه واژه افزایش وضوح به‌وضوح فضایی مربوط می‌شود. همان‌طور که قبلاً نیز بیان شد وضوح فضایی به‌طور ویژه تعداد کل پیکسل‌ها در یک تصویر را نشان می‌دهد. به‌عنوان‌مثال برای یک تصویر دیجیتال ۳۰۰×۳۰۰ 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 اندازه پیکسل‌ها ۲) افزایش اندازه چیپ[۳, ۴]. 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 فن‌های پردازش تصویر می‌باشد. این روش‌ها به‌طور ویژه به روش‌های فرا تفکیک‌پذیری مربوط می‌شوند. فرا تفکیک‌پذیری به فرآیندی گفته می‌شود که سعی در بازسازی تصویری وضوح‌بالا از یک یا چند تصویر وضوح‌پایین دارد. به عبارتی در فرا تفکیک‌پذیری در بازیابی اطلاعات فرکانس بالایی که در طول فرآیند تصویربرداری از بین رفته‌اند، تلاش می‌شود. روش‌های نرم‌افزاری ازلحاظ اقتصادی مقرون‌به‌صرفه می‌باشند و امکان تولید تصویری با وضوح بالاتر را فراهم می‌آورند. به عبارتی در این روش‌ها می‌توانیم از محدوده‌ی توانایی سیستم تصویربرداری بالاتر رویم. این فن 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[۵, ۶] و غیره کاربرد دارد. مسئله فرا

تفکیک‌پذیری را نباید با فن‌های مشابه همانند درون‌یابی[۱] و بازیابی[۲] اشتباه گرفت. در مقایسه با روش‌های بهبود وضوح فن فرا تفکیک‌پذیری نه‌تنها کیفیت تصویر را افزایش می‌دهد بلکه باعث افزایش وضوح فضایی یک تصویر نیز می‌شود. در روش‌های درون‌یابی جزئیات فرکانس بالا برخلاف روش‌های فرا تفکیک‌پذیری بازیابی نمی‌شوند[۷]. در روش‌های بازیابی ابعاد تصویر ورودی و خروجی یکسان است و تنها کیفیت تصویر خروجی افزایش پیداکرده است. در فرا تفکیک‌پذیری علاوه بر اینکه کیفیت تصویر خروجی افزایش پیدا می‌کند ابعاد آن نیز نسبت به تصویر ورودی افزایش پیدا می‌کند.

۱-۲ فاکتورهای خرابی تصویر

تصویری که به دست کاربر می‌رسد معمولاً حالت رضایت بخشی از جزئیات صحنه موردنظر را (به دلیل کامل نبودن سیستم تصویربرداری و شرایط تصویربرداری) نشان نمی‌دهد. این تصویر یک نسخه تنزل یافته (خراب‌شده) از صحنه طبیعی است. این خرابی‌ها به‌واسطه‌ی فاکتورهایی مثل ماتی[۳]، نویز و روی‌هم افتادگی فرکانسی[۴] اتفاق می‌افتد. شکل ۱-۱ مثالی از یک صحنه‌ی طبیعی و تصویری که درنهایت توسط سیستم تصویربرداری تولید می‌شود، نشان می‌دهد. 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

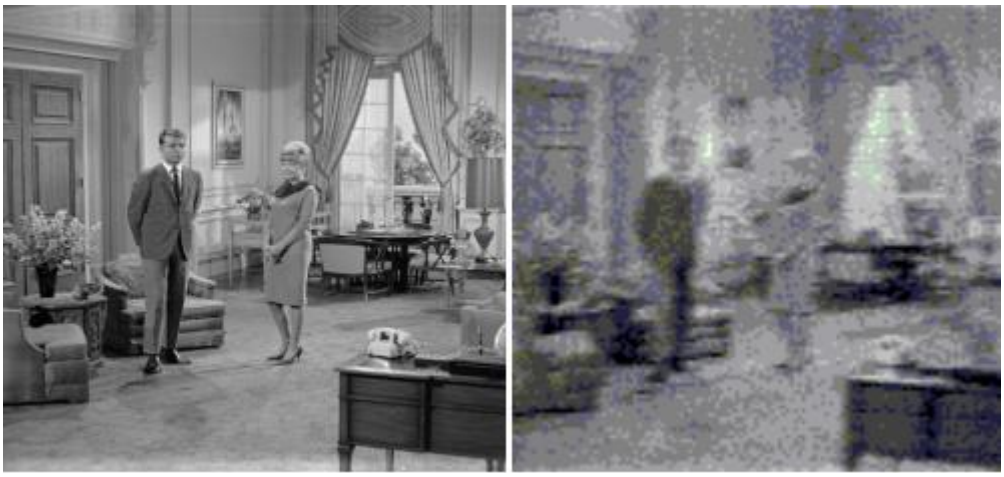
۴- کم بودن نرخ نمونه‌برداری

---

[۱] Interpolation method
[۲] Restoration
[۳] Blur
[۴] Aliasing
[۵] Motion

فاکتوری شبیه حرکت نسبی دوربین و صحنه، فوکوس نادقیق، توربولنس اتمسفر، تابع نقطه 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 در زمان گرفتن یک تصویر شناخته‌شده باشد فن حذف ماتی آسان‌تر و با دقت بالاتری انجام می‌گیرد. شکل ۱-۲ مثالی از تصویر مات شده و صحنه اصلی را نشان می‌دهد.

اگرچه تاکنون برای نویز مدل‌های متفاوتی در این موضوع پیشنهادشده است اما فقط نویز سفید گوسی را در نظر می‌گیرند زیرا مدل خوبی برای بیشتر سیستم‌های تصویربرداری می‌باشد. همچنین فرض بر این است که نویز به تصویر هیچ شباهتی ندارد به این معنی که بین مقادیر پیکسل‌های تصویر و مؤلفه‌های نویز همبستگی[۳] وجود ندارد. شکل ۱-۳ مثالی از تصویر نویزی و صحنه اصلی را نشان می‌دهد.



شکل ۱-۱ تصویر اصلی (سمت چپ) و نسخه نویزی، مات شده و کاهش وضوح یافته‌ی آن (سمت راست).

[۱] Point spread function

[۲] De-bluring

[۳] correlation

شکل ۱-۲. مثالی از تصویر اصلی (سمت چپ) و تصویر مات شده‌ی آن (سمت راست)

شکل ۱-۳. مثالی از تصویر نویزی (سمت چپ) و صحنه‌ی اصلی (سمت راست).

فاکتور دیگری که در افزایش وضوح تصویر تأثیر دارد نرخ نمونه‌برداری می‌باشد. بر طبق تئوری نمونه‌برداری نایکوئیست[۸, ۹] 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 زیر نمونه‌برداری[۱] شناخته می‌شود که در بسیاری از حس‌گرهای تصویربرداری اتفاق

[۱] Under-sampling

می‌افتد. به دلیل پدیده زیر نمونه‌برداری، مؤلفه‌ی فرکانس بالا با مؤلفه‌های فرکانس پایین همپوشانی خواهد داشت و به‌عنوان سیگنال یا تصویر بازسازی‌شده‌ای معرفی می‌شود که وضوح آن تنزل یافته است. چنین تنزل‌هایی به‌عنوان پدیده‌ی روی‌هم افتادگی فرکانسی در نظر گرفته می‌شوند که باعث از بین رفتن اطلاعات جزئی تصویر می‌شوند. روی‌هم افتادگی فرکانسی سبب تولید اثرات مصنوعی در بازسازی سیگنال می‌شود. برای کاهش این اثرات مصنوعی باید از فن‌های حذف اثر روی‌هم افتادگی فرکانسی[1] استفاده شود. شکل ۱-۴ مثالی از صحنه اصلی و تصویر زیر نمونه‌برداری شده را نشان می‌دهد.

شکل ۱-۴. مثالی از تصویر زیر نمونه‌برداری شده (سمت چپ) و صحنه طبیعی (سمت راست)

۱-۳ معنی فرا تفکیک‌پذیری

همان‌طور که قبلاً هم بیان شد وضوح فضایی یک تصویر به تعداد پیکسل‌ها در یک تصویر اشاره دارد، بنابراین هرچه تعداد پیکسل‌ها بیشتر باشد جزئیات بیشتر و اطلاعات بیشتری در تصویر مشاهده خواهد شد؛ بنابراین یک سؤال اساسی به وجود می‌آید که چرا به فرا تفکیک‌پذیری برای افزایش وضوح نیاز است. برای پیدا کردن جواب اساسی این سؤال ابتدا باید با حس‌گر تصویر آشنا شد. یک حس‌گر تصویر یا دوربین، وسیله‌ای است که انرژی نوری را به یک سیگنال الکتریکی تبدیل می‌کند. حس‌گرهای

[1] Anti-aliasing

تصویربرداری مدرن مبتنی بر فنّاوری دستگاه جفت کننده بار[۱] یا نیمه‌رسانای اکسید فلزی[۲] می باشند. ضروری است که شامل یک آرایه‌ای از عناصر حسگر نوری[۳] یا پیکسل‌هایی باشند که ولتاژ خروجی متناسب با نوری که به سطح آن‌ها می‌خورد داشته باشند[۱۰]. ابعاد حس‌گر یا به عبارتی تعداد حس‌گرها بر واحد سطح تعیین‌کننده وضوح فضایی یک تصویر می‌باشند. هرچه تعداد عناصر آشکارساز بیشتر شود جزئیات بیشتری در تصویر ظاهر می‌شود. یک سیستم تصویربرداری با آشکارسازهای ناکافی تصویر وضوح پایین با اثرات مصنوعی تولید می‌کند. یک روش برای افزایش وضوح این است که ابعاد آشکارسازها کاهش یابد. با این کار چگالی آشکارسازها افزایش پیدا می‌کند و نرخ نمونه‌برداری نیز افزایش می‌یابد. به‌محض اینکه ابعاد آشکارساز کاهش پیدا کند مقدار نور رسیده به هر پیکسل کاهش پیدا می‌کند و این اتفاق سبب تولید نویز شات [۱۱] می‌شود که کیفیت تصویر را کاهش می‌دهد. بنابراین بر روی ابعاد آشکارساز در یک حس‌گر محدودیت وجود دارد. روش‌های فرا تفکیک‌پذیری به سه دسته تقسیم می‌شوند[۱۲]. ۱) روش‌های مبتنی بر درون‌یابی ۲) روش‌های فرا تفکیک‌پذیری چند فریمی (کلاسیک) ۳) روش‌های فرا تفکیک‌پذیری مبتنی بر یادگیری

۱-۳-۱ روش‌های درون‌یابی

معنی تحت الفظی این واژه الحاق کردن (جا دادن، داخل کردن) است. درواقع مقدار میانی به‌وسیله‌ی مقادیر همسایه‌ها تخمین زده می‌شود. در پردازش تصویر، درون‌یابی یک تصویر به این معنی است که یک سیگنال پیوسته به‌وسیله‌ی یک مجموعه‌ای از داده‌های تصاویر گسسته تخمین زده شود. روش‌های مبتنی بر درون‌یابی برای تخمین تصویر وضوح‌بالا از چندجمله‌ای‌ها استفاده می‌کنند. در ابتدا تصویر ورودی را بزرگ می‌کنند سپس مقادیر نمونه‌های جدید را با درون‌یابی نمونه‌هایی که در همسایگی

---

[۱] Charge couple device

[۲] Complementary metal-oxide semiconductor

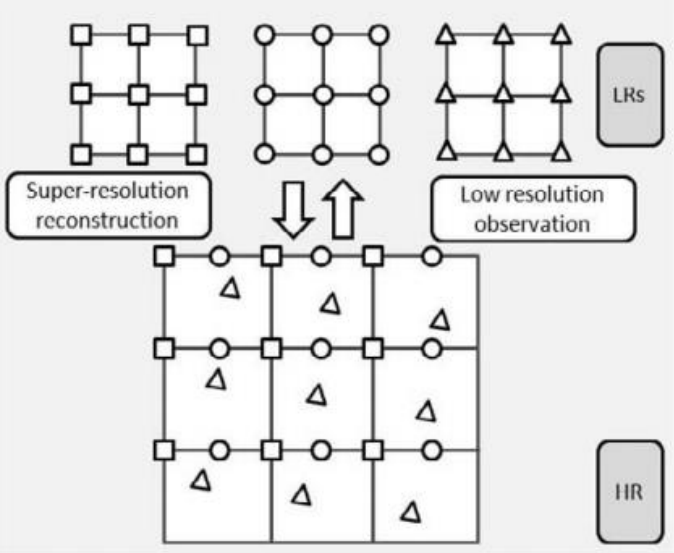
[۳] Photo-detector

قرارگرفته‌اند به دست می‌آورند. نتیجه‌ی چنین روش‌هایی در نقاط جزئیات بالای تصویر همانند لبه‌ها رضایت‌بخش نیست. به عبارتی استفاده از روش‌های درون‌یابی برای تصاویر طبیعی درست نیست و فرضیاتی که در این روش‌ها در نظر گرفته می‌شود برای تصاویر طبیعی کاربرد ندارد. یکی از نقطه‌ضعف‌های درون‌یابی bicubic این است که اثرات زیگزاکی در طول لبه‌ها به وجود می‌آورد که دلیل آن این است که به‌صورت محلی از همسایگی ۴×۴ برای تخمین مقادیر پیکسل‌های جدید استفاده می‌کند. در درون‌یابی‌های جهتی ابتدا جهت لبه را محاسبه می‌کنند و سپس با انتخاب پیکسل‌های مبدأ در طول لبه به درون‌یابی بقیه پیکسل‌ها می‌پردازند[۱۳, ۱۴]. درروش bicubic زمانی که فاکتور بزرگ‌نمایی کوچک باشد وضوح تصویر خروجی قابل‌قبول خواهد بود. اما زمانی که فاکتور وضوح افزایش پیدا کند تصویر خروجی مات خواهد شد.

۱-۳-۲ روش‌های فرا تفکیک‌پذیری چند فریمی

در این روش‌ها با استفاده از ترکیب اطلاعات چند تصویر وضوح پایین یک تصویر وضوح‌بالا به دست می‌آید. به عبارتی مؤلفه‌های فرکانس بالا افزایش پیدا می‌کنند[۱۵-۱۸]. ایده‌ی اصلی در این روش‌ها این است که از اطلاعات مفید مربوط به چند فریم که از یک صحنه گرفته شده‌اند، برای تولید تصویر وضوح‌بالا استفاده می‌شود. اطلاعات مفیدی که در این تصاویر به وجود می‌آید توسط شیفت‌های زیر پیکسلی بین فریم‌ها معرفی می‌شوند. این شیفت‌های زیر پیکسلی معمولاً به دلیل حرکات غیرقابل‌کنترل بین سیستم تصویربرداری و صحنه همانند حرکات شی و یا به‌وسیله‌ی حرکات قابل‌کنترل همانند حرکت سیستم تصویربرداری ماهواره‌ای که با سرعت مشخص و از پیش تعریف‌شده‌ای به دور زمین می‌چرخد، به وجود می‌آید. بنابراین فرا تفکیک‌پذیری چند فریمی در صورتی امکان‌پذیر است که بین پیکسل‌های موجود در فریم‌ها شیفت‌های زیر پیکسلی وجود داشته باشد تا بتواند مشکل بدحالت بودن این مسئله را حل کند[۱۵]. شکل ۱-۵ دیاگرام ساده‌ای از ایده‌ی بازسازی SR را نشان می‌دهد. روش‌های SR کلاسیک به فاکتور وضوح کوچک محدود می‌شوند[۱۹, ۲۰]. از طرفی داشتن چندین فریم از یک صحنه

برای کاربردهای عملی کارآمد نیست. درنتیجه روش‌های مبتنی بر یادگیری برای غلبه بر مشکلات SR کلاسیک به وجود آمدند.



شکل ۱-۵. ایده اصلی فرا تفکیک‌پذیری چند فریمی. شیفت‌های زیر پیکسلی اطلاعات تکمیل‌کننده فریم‌های وضوح پایین می‌باشند و به کمک آن‌ها تصویر وضوح‌بالا ساخته می‌شود.

۱-۳-۳ فرا تفکیک‌پذیری تک فریم

در این روش‌ها رابطه‌ی میان وصله‌های وضوح‌بالا و پایین با استفاده از یک پایگاه داده خارجی یاد گرفته می‌شود. سپس این رابطه به وصله‌های وضوح پایین تصویر ورودی اعمال می‌شود و یک تصویر وضوح‌بالا تولید می‌شود. این مسئله همانند حل یک دستگاه m معادله با n مجهول است به‌طوری‌که تعداد مجهولات از تعداد معادلات بیشتر می‌باشد (n>m). در این حالت بی‌نهایت جواب برای دستگاه معادله وجود خواهد داشت و همین موضوع مسئله SR را بدحالت می‌کند. برای حل این مشکل به یک سری فرضیات و اطلاعات پیشین نیاز است (همانند استفاده از یک پایگاه داده خارجی و یا استفاده از شباهت میان وصله‌های تصویر در هرم). روش‌های متفاوتی برای یادگیری رابطه‌ی میان وصله‌های

وضوح‌بالا و پایین ارائه شده‌اند. موفقیت روش‌های SR مبتنی بر مثال به دو فاکتور مهم وابسته است: ۱) جمع‌آوری یک پایگاه داده بسیار بزرگ شامل جفت وصله‌های وضوح‌بالا و پایین (HR/LR) ۲) یادگرفتن رابطه بین وصله‌های وضوح‌بالا و پایین (HR/LR). روش‌های SR مبتنی بر یادگیری به دودسته تقسیم می‌شوند ۱) مبتنی بر پایگاه داده خارجی ۲) مبتنی بر شباهت میان وصله‌های تصویر. اولین گروه مبتنی بر روش ارائه‌شده توسط آقای Freeman می‌باشد[۲۱]. در این گروه به یک پایگاه داده خارجی شامل میلیون‌ها تصویر وضوح‌بالا و پایین نیاز است. جمع‌آوری چنین پایگاه داده‌ی خارجی بسیار سخت می‌باشد. هرچه تعداد وصله‌ها زیاد باشد بار محاسباتی نیز بالا می‌رود. اگر محتویات تصویر ورودی با پایگاه داده متناسب نباشد تابع نگاشت به‌دست‌آمده ممکن است توانایی بازسازی جزئیات فرکانس بالا را نداشته باشد. به‌عنوان‌مثال اگر تابع نگاشت از پایگاه داده‌ای که شامل تصاویری از اشیا ساخته‌شده توسط انسان، یاد گرفته‌شده باشد (به‌عنوان‌مثال ماشین، ساختمان و ....) این تابع برای بازسازی تصاویر طبیعی مناسب نخواهد بود. درنتیجه این روش‌ها می‌توانند سبب هالوسینه[۱] کردن جزئیات فرکانس بالا شوند[۲۱, ۲۲]. ازاین‌رو اگر رابطه بین وصله‌های وضوح‌بالا و پایین با توجه به وصله‌های خود تصویر در مقیاس‌های متفاوتی از آن یاد گرفته شود مشکلات مطرح‌شده در بالا وجود نخواهد داشت. برای این منظور از هرم تصویر استفاده می‌کنند که با توجه به سطوح هرم می‌توان تصویر وضوح‌بالا و پایین به دست آورد. ازاین‌رو در این پایان‌نامه یک ایده‌ی جدید در فرا تفکیک‌پذیری ارائه می‌شود که به پایگاه داده خارجی نیاز نداشته باشد و با کمک هرم تصویر، جزئیات ازدست‌رفته در تصویر ورودی بازگردانی می‌شود.

۱-۴ کاربردهای فرا تفکیک‌پذیری

فرا تفکیک‌پذیری در بسیاری از زمینه‌ها کاربرد دارد:

[۱] Hallucination

۱- نظارت بر ویدیو[۲۳, ۲۴]: بزرگ‌نمایی منطقه موردنظر(ROI) در یک ویدیو برای شناسایی افراد یا خواندن پلاک خودرو

۲- سنجش از راه دور[۲۵]: به کمک چندین تصویر وضوح پایین که از یک منطقه گرفته می‌شود یک تصویر وضوح‌بالا می‌سازند.

۳- پردازش تصاویر پزشکی (CT, MRI,...)[۲۶]: با استفاده از چندین تصویر وضوح پایین یک تصویر وضوح‌بالا به وجود می‌آورند.

۴- تصاویر ماهواره‌ای

۵- شناسایی و تشخیص هدف [۲۹].

۱-۵ سهم ما

سعی ما این است که یک روش جدید و سریع در فرا تفکیک‌پذیری با استفاده از خود تصویر ورودی و بدون نیاز به جمع‌آوری پایگاه داده خارجی ارائه دهیم. اخیراً روش‌هایی که در این حوزه محبوبیت پیداکرده‌اند روش‌های مبتنی بر دیکشنری می‌باشند[۳۰, ۳۱]. ما نیز از وصله‌های تصویر ورودی در سطح‌های متفاوتی از هرم برای تولید وصله‌های وضوح‌بالا و پایین استفاده می‌کنیم و با کمک دیکشنری رابطه‌ی میان این وصله‌ها را به دست می‌آوریم و پیچیدگی محاسباتی روش‌های مبتنی بر دیکشنری را نخواهیم داشت. همانند بسیاری از روش‌های مبتنی بر خود یادگیری به پایگاه داده خارجی نیاز نداریم و با استفاده از دو سطح هرم ساخته‌شده از تصویر ورودی ارتباط بین وصله‌های وضوح‌بالا و پایین را به دست می‌آوریم. با مرور روش‌های گذشته و بررسی نقاط ضعف آن‌ها به این نتیجه رسیدیم که ایده‌ی بخش بندی تصویر[۱] برای ساخت دیکشنری بسیار مفید است چون فلسفه‌ی به وجود آمدن آن این است که هر وصله بر اساس وصله‌هایی که بسیار شبیه به آن هستند ساخته شود. سه مشکل اساسی در

[۱] Image segmentation

مبحث بخش‌بندی تصویر وجود دارد، در ادامه این مشکلات را توضیح می‌دهیم و برای از بین بردن آن‌ها ایده‌ی جدید بلوک‌بندی تصویر را ارائه می‌دهیم. این مشکلات عبارت‌اند از:

۱- هر الگوریتم بخش‌بندی تصویر برای اجرا به زمان نسبتاً زیادی احتیاج دارد که این نقطه‌ضعف روش‌های مبتنی بر بخش‌بندی هستند.

۲- هیچ‌کدام از الگوریتم‌های بخش‌بندی موجود تضمین نمی‌کنند که بخش‌بندی درستی از تصویر انجام دهند.

۳- در بخش‌بندی تصویر با وصله‌های مرزی روبرو می‌شویم که خطای بازسازی را بیشتر می‌کنند.

با توجه به مشکلات بخش‌بندی، ایده‌ی جدیدی برای ساخت هر وصله با وصله‌های مشابه محلی آن ارائه می‌دهیم که بر مشکلات بالا غلبه می‌کند. ایده‌ی بلاک بندی تصویر را ارائه می‌دهیم که یک ایده‌ی کاملاً جدید است و مشکل ساخت وصله‌های مرزی را از بین می‌برد، از طرفی بلاک‌بندی تصویر نسبت به بخش‌بندی به زمان بسیار کمتری نیاز دارد. درنتیجه با ساختن دیکشنری محلی توسط یک بلاک و ۸ بلاک همسایه، اتم‌های تشکیل‌دهنده دیکشنری به هم نزدیک‌تر خواهند شد و خطای بازسازی کمتر می‌شود. برای ساخت دیکشنری الگوریتم‌های زیادی مانند K-SVD[۱] [۳۲] ارائه‌شده‌اند که بسیار زمان‌بر می‌باشند. ایده ما در این پایان‌نامه این است که به‌طور مستقیم وصله‌های استخراج‌شده از تصویر را در قالب بردارهای ستونی کنار هم قرار داده و بدین ترتیب بدون اجرای هیچ الگوریتم اضافی به دیکشنری مبتنی بر وصله‌های تصویر دست می‌یابیم. با توجه به اینکه تعداد اتم‌های دیکشنری (ستون‌های دیکشنری) به‌دست‌آمده نسبتاً کم و اتم‌ها به هم شبیه هستند دیکشنری ساخته‌شده تفاوت چندانی با دیکشنری آموزش‌دیده ندارد نتایج تجربی ارائه شده گواه این ادعاست. برای یادگیری ارتباط بین وصله‌ی وضوح‌بالا و وضوح‌پایین از هرم تصویر استفاده می‌کنیم، اکثر روش‌ها برای یادگیری رابطه‌ی وصله‌ی

[۱] K-singular value decomposition

وضوح‌بالا و پایین از کوچک کردن[1] وصله‌ی وضوح‌بالا استفاده می‌کنند. این عملیات به دلیل زیاد بودن وصله‌ها بسیار زمان‌بر است. ما ایده‌ی جدیدی برای به دست آوردن وصله‌ی وضوح پایین متناظر با هر وصله‌ی وضوح‌بالا ارائه می‌دهیم. همچنین ایده‌ی جدیدی برای افزایش PSNR تصویر SR نهایی پیشنهاد می‌دهیم. این ایده مبتنی بر استخراج ویژگی برای ساخت دیکشنری می‌باشد. به این معنی که برای ساخت دیکشنری به‌جای آنکه مستقیم از وصله استفاده کنیم (ویژگی‌ها در این حالت همان روشنایی می‌باشند) ابتدا با استفاده از ۴ فیلتر بالا گذر لبه‌های تصویر را به دست می‌آوریم و برای ساخت دیکشنری از تصاویر فیلتر شده استفاده می‌کنیم. همچنین از تنظیم‌کننده تیخونوف در حل مسئله بهینه‌سازی استفاده می‌کنیم و درنهایت برای تولید SR نهایی از تنظیم‌کننده BTV[2] برای از بین بردن اثرات مصنوعی تصویر استفاده می‌کنیم. با انجام کارهای اشاره‌شده تصویر بازسازی‌شده‌ی SR هم ازلحاظ PSNR و هم ازلحاظ بصری نسبت به روش‌های موجود از کیفیت بهتری برخوردار است.

۱-۶ خلاصه‌ای از فصل

در این فصل دید کاملی از تعریف فرا تفکیک‌پذیری به خواننده داده شد. همچنین مفهوم پردازش سیستم‌های دیجیتال و فاکتورهایی که سبب خراب شدن وضوح تصاویر می‌شوند معرفی گردید. علاوه بر آن دلیل استفاده از روش‌های فرا تفکیک‌پذیری و همچنین تعدادی از کاربردهای فرا تفکیک‌پذیری نیز مطرح شد.

۱-۷ ساختار پایان‌نامه

در ادامه به معرفی کارهای انجام‌گرفته در این پایان‌نامه می‌پردازیم:

در فصل ۲ مروری بر روش‌های انجام‌گرفته در زمینه‌ی فرا تفکیک‌پذیری تک تصویر خواهیم داشت

---

[1] Down-sample

[2] Bilateral total variation

و روش‌هایی را که در حوزه‌ی مشابه با کار ما قرار دارند را معرفی خواهیم کرد و مزایا و معایب هرکدام از روش‌های شرح داده شده را توضیح خواهیم داد. در فصل ۳ تئوری‌های موردنیاز برای انجام فرآیند فرا تفکیک‌پذیری را مطرح خواهیم کرد. در فصل ۴ روش پیشنهادی را مطرح می‌کنیم. در فصل ۵ نتایج تجربی را نشان می‌دهیم. درنهایت در فصل ۶ نتایج کلی پایان‌نامه را مطرح می‌کنیم و پیشنهادهایی برای کارهای آتی خواهیم داشت.

۲ فصل دوم: مرور روش‌های گذشته

در سال‌های اخیر تلاش‌های زیادی در زمینه‌ی پردازش تصویر صورت گرفته است. علت اصلی این تلاش‌ها این است که اکثریت داده‌هایی که از طریق مشاهده یا فن‌های پردازش تصویر به دست انسان‌ها می‌رسند در زمینه‌های گسترده‌ای همچون تصاویر پزشکی، نظارت و کنترل از راه دور کاربرد دارند؛ بنابراین در بسیاری از زمینه‌ها نیاز به داشتن تصویری با وضوح‌بالا رو به افزایش است. ایده‌ی فرا تفکیک‌پذیری اولین بار توسط Tsai و Huang در سال ۱۹۸۴ معرفی شد [۱]. فرا تفکیک‌پذیری تک تصویر در دو دهه اخیر توجه بسیاری از محققان را به خود جلب کرده است. بیشتر این تحقیقات بر روی ترکیب چندین تصویر LR از یک صحنه برای بازسازی یک تصویر یا دنباله‌ای از تصاویر وضوح‌بالا انجام شده‌اند[۳۳, ۳۴]. ایده‌ی اصلی بیشتر فن‌های مطرح‌شده این است که از چندین تصویر که از یک صحنه تحت شرایط نوری یکسان اما در شرایط متفاوتی ازلحاظ مکان حس‌گرها گرفته شده است، استفاده می‌کنند. زمانی که دو تصویر دید متفاوتی از یک صحنه را به وجود می‌آورند حرکت یا بردار حرکت (مجموعه‌ای از جابه‌جایی‌ها میان پیکسل‌ها) میان پیکسل‌های متناظر در دو تصویر اطلاعات مفیدی را برای بازسازی تصویر به وجود می‌آورند. حرکت واقعی میان تصاویر در دسترس نیست بنابراین باید بردار حرکت تخمین زده شود که کاری بسیار سخت و پیچیده است[۳۳, ۳۴]. محدودیت دیگر این فن‌ها این است که نیاز به چندین تصویر از یک صحنه وجود دارد که برآورده کردن این خواسته در همه‌ی کاربردها میسر نمی‌باشد و از طرفی هزینه‌بر است. همچنین همه‌ی تصاویر LR شایستگی لازم برای بازسازی SR را نخواهند داشت. به عبارتی الگوریتم‌های SR چند تصویری در صورتی کارایی دارند که تصاویر ورودی نسبت به هم شیفت‌های زیر پیکسلی داشته باشند. مسئله بازسازی SR اندکی با هم‌جوشی تصاویر[۱] متفاوت است. در مسئله هم‌جوشی تصویر، یک یا چند تصویر LR با یک یا چند تصویر HR باهم ترکیب می‌شوند و تصویری با وضوح فضایی بالاتر نسبت به تصویر LR به وجود می‌آید. بنابراین در روش‌های هم‌جوشی حداقل یک تصویر HR مورد نیاز است و وضوح فضایی تصویر خروجی وابسته به‌اندازه

[۱] Image-Fusion

پیکسل‌های HR می‌باشند. دسته‌ی دیگری از فرا تفکیک‌پذیری حالت تک فریمی می‌باشد، به‌طوری‌که تنها یک تصویر وضوح پایین برای تولید تصویر وضوح‌بالا موردنیاز است. این روش در کاربردهایی که چند تصویر وضوح پایین در دسترس نباشد بسیار مؤثر خواهد بود. دقت الگوریتم‌های SR چند فریمی به‌دقت تخمین حرکت بین فریم‌های LR وابسته است. این محدودیت در کاربردهای واقعی استفاده از الگوریتم‌های SR چند فریمی را با مشکل مواجه می‌کند. زیرا ممکن است اشیای مختلف در صحنه‌های یکسان حرکت‌های پیچیده و متفاوتی داشته باشند. ازاین‌رو الگوریتم‌های SR تک تصویر برای غلبه بر این مشکلات به وجود آمده‌اند. در الگوریتم‌های SR تک تصویر برای بازیابی اطلاعات فرکانس بالا به ترم پیشین(همانند استفاده از یک پایگاه داده خارجی و یا استفاده از هرم تصویر ورودی) نیاز است. در یک مقاله مروری خلاصه‌ای از بیشتر الگوریتم‌های پیشنهادشده شامل روش‌های مبتنی بر تنظیم‌کننده [۳۵]، فن‌های نگهداری لبه[۳۶] و الگوریتم‌های بیزین[۳۷] گزارش شده است [۳۸]. الگوریتم‌های SR تک تصویر خود به دو دسته تقسیم می‌شوند: ۱) مبتنی بر پایگاه داده خارجی ۲) مبتنی بر محتوای خود تصویر

۲-۱ روش‌های SR مبتنی بر پایگاه داده خارجی

این روش‌ها الگوریتم‌های مبتنی بر یادگیری یا الگوریتم‌های هالوسینه می‌باشند که ابتدا در [۳۹] 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 ترم پیشین برای بازسازی تصویر وضوح‌بالا به‌کار برده می‌شود. پایگاه آموزشی الگوریتم‌های SR مبتنی بر پایگاه داده خارجی باید قابلیت عمومیت گرایی خوبی داشته باشند[۴۰]. پایگاه داده خارجی بزرگ لزوماً نتایج بهتری تولید نمی‌کند در مقابل زیاد بودن نمونه‌های نامرتبط نه‌تنها بار محاسباتی پیدا کردن بهترین نگاشت را زیاد می‌کنند بلکه باعث آشفتگی این عملیات نیز می‌شوند. روش‌های متفاوتی از این

الگوریتم در ادامه توضیح داده می‌شود.

۲-۱-۱ استفاده از ویژگی‌های هرم

کارهای برجسته‌ی این گروه توسط Baker و Kanade بر روی افزایش وضوح تصاویر چهره توسعه پیدا کرده‌اند [۴۱, ۴۲]. در فاز آموزش این الگوریتم‌ها، ابتدا هر تصویر HR را در چندین مرحله مات و کوچک می‌کنند تا یک هرم تصویر گوسین به وجود آورند. سپس از این هرم‌های گوسین هرم‌های لاپلاسین و بعد از آن هرم‌های ویژگی تولید می‌نمایند. ویژگی‌ها با استفاده از هرم‌های گوسین تولید شدند که شامل اطلاعات فرکانس پایین از تصاویر چهره می‌باشند. هرم‌های لاپلاسین شامل اطلاعات میان‌گذر موجود در تصاویر چهره می‌باشند. هرم‌ها شامل اطلاعات چندجهته از تصاویر چهره هستند[۴۳]. با داشتن یک سیستم آموزش‌دیده، برای یک تصویر تست وضوح پایین، شبیه‌ترین تصویر LR در میان تصاویر LR در همه‌ی هرم‌های تصویر پیدا می‌شود. در [۴۴] از فن نزدیک‌ترین همسایه[۱] برای پیدا کردن مشابه‌ترین تصاویر یا وصله مشابه استفاده شده است.

۲-۱-۲ استفاده از Belief network

کارهای برجسته در این گروه بر پایه مقاله Freeman و Pasztor [۴۵, ۴۶] می‌باشد. این الگوریتم‌ها از یک شبکه همانند شبکه مارکوف یا ساختار درخت استفاده می‌کنند. در این مورد از شبکه مارکوف هر دو تصویر HR و LR به چندین وصله‌ی تصویر تقسیم می‌شوند. سپس وصله‌های متناظر در این دو تصویر توسط تابع مشاهداتی به هم مرتبط می‌شوند.

---

[۱] Nearest -Neighbor

۲-۱-۳ نگاشت[۱]

ترم پیشین یاد گرفته‌شده در [۴۱, ۴۲]یک محدودیت محلی بر روی تصویر بازسازی‌شده به وجود می‌آورد. برای ایجاد محدودیت‌های سراسری تعدادی از الگوریتم‌های SR از روش‌های مبتنی برنگاشت برای یادگیری ترم پیشین استفاده می‌کنند. به‌عنوان‌مثال در [۴۷, ۴۸] از PCA و در [۴۹] از آنالیز اجزای مورفولوژیکی (MCA[۲]) استفاده شده است. در روش‌های مبتنی در PCA معمولاً ماتریس‌های ارائه‌دهنده هر تصویر آموزشی ابتدا به‌صورت بردار بازنمایی می‌شوند (با سازماندهی کردن همه‌ی ستون‌ها فقط در یک ماتریس) و سپس آن‌ها به یک ماتریس بزرگ تبدیل می‌شوند تا ماتریس کوواریانس داده‌ی آموزشی برای مدل کردن eigenspaceها به دست آید.

۲-۱-۴ استفاده از شبکه‌های عصبی

رهیافت روش‌ها در این گروه با روش‌های مبتنی بر belief network یکسان است اما انواع متفاوتی از شبکه عصبی را استفاده می‌کنند و در الگوریتم‌های متفاوتی از SR به کار می‌روند. نمونه‌هایی از این کار عبارت‌اند از شبکه‌های عصبی هاپلفید[۳] [۵۰, ۵۱]، شبکه عصبی احتمالاتی[۵۲]، پرسپترون چندلایه[۵۳] و توابع پایه‌ای شعاعی(RBF[۴]) [۵۴].

۲-۱-۵ استفاده از Manifold

در روش‌های مبتنی بر منیفلد [۵۵, ۵۶] [۵۷]فرض می‌شود که تصاویر HR و LR، منیفلدهایی با هندسی محلی مشابه در دو فضای ویژگی متمایز تشکیل می‌دهند[۵۸]. این روش‌ها معمولاً برای کاهش دیمانسیون به کار می‌روند. این روش‌ها معمولاً دارای مراحل زیر می‌باشند:

---

[۱] Projection
[۲] Morphological Component
[۳] Hopfield NN
[۴] Radial basis function

منیفلدهای HR/LR در فاز آموزش با استفاده از تصاویر HR و LR متناظر با آن‌ها ساخته می‌شود. در فاز تست ابتدا تصویر تست LR را به چندین وصله LR تقسیم می‌کنند. برای هر وصله‌ی LR از این تصویر (l_x) تعداد K وصله‌ی مشابه به آن را با استفاده از الگوریتم K-nearest از روی منیفلد وضوح پایین پیدا می‌کنند(l_j). برای هر کدام از این K وصله، ضرایب (w_x^L) موردنیاز برای بازسازی وصله l_x را محاسبه می‌کنند.

$$arg\min_{w_x^L} \left\| l_x - \sum_{j \in N_L} w_x^L l_j \right\| \qquad (۱-۲)$$

به‌طوری‌که جمع وزن‌ها برابر یک خواهد بود. با استفاده از وزن‌های به‌دست‌آمده و با کمک همسایه‌ها، وصله‌ی وضوح‌بالای متناظر با وصله وضوح پایین l_x طبق فرمول(۲-۲) به دست می‌آید:

$$\widehat{h_x} = \sum_{j \in N_L} w_{xj}^H h_j \qquad (۲-۲)$$

به‌طوری که N_L (x) شامل همسایه‌های l_x است و h_j ها وصله‌های وضوح‌بالای متناظر با وصله‌های همسایه وضوح پایین l_x می‌باشند. بر اساس بیشتر روش‌های مبتنی بر منیفلد فرضیه‌ای که برای حل این روش‌ها به کار می‌رود این است که $w_{xj}^H = w_{xj}^L$. این فرضیه در[۵۹, ۶۰] مطرح‌شده است. باید خاطرنشان شود که ممکن است این فرضیه درست نباشد.

۲-۱-۶ استفاده از Compressive sensing

در پردازش سیگنال پیشنهاد شده است که بین سیگنال‌های HR و پروجکشن های LR آن‌ها رابطه‌ی خطی وجود دارد. اخیراً بسیاری از محققان از پراکندگی تصاویر برای الگوریتم‌های SR استفاده کرده‌اند. در حوزه اسپارس فرض می‌شود یک دیکشنری فوق کامل[۱] شبیه $D \in R^{n \times k}$ وجود دارد که شامل k وصله تصویر می‌باشد. این فرض وجود دارد که یک تصویر HR یعنی $y_h \in R^n$ می‌تواند به‌صورت

[۱] Overcomplete

ترکیب خطی از این وصله‌ها ارائه شود.

$$y_h = Dq \tag{۳-۲}$$

به‌طوری که q بردار ضرایب اسپارس است به این معنی که $\|q\|_0 << k$.

این الگوریتم‌ها [۶۱, ۶۲] دو دیکشنری فوق کامل دارند: یکی برای وصله‌های HR که با D_h نشان داده می‌شود و دیگری برای وصله‌های وضوح پایین متناظر آن‌ها که با D_l نشان داده می‌شود. معمولاً مقادیر میانگین پیکسل‌ها از هر وصله کم می‌شود که دیکشنری ساخته‌شده بیشتر اطلاعات بافت[۱] را نسبت به اطلاعات روشنایی بازنمایی کند. با داشتن یک تصویر LR (**g**) بحث بالا برای پیدا کردن ضرایب اسپارس به‌صورت زیر مینیمم می‌شود:

$$\lambda \|q\|_1 + \frac{1}{2}\left\|\tilde{D}q - \tilde{g}\right\|_2^2 \tag{۴-۲}$$

به‌طوری که $\tilde{D} = \begin{bmatrix} FD_l \\ \beta P D_h \end{bmatrix}$ و $\tilde{g} = \begin{bmatrix} Fg \\ \beta w \end{bmatrix}$.

F اپراتور استخراج ویژگی(معمولاً فیلتر بالا گذر [۶۱, ۶۲])، P ناحیه‌ی همپوشانی[۲] میان وصله‌ها می‌باشد. β پارامتر کنترلی برای پیدا کردن مصالحه‌ای بین تصویر LR ورودی و پیدا کردن وصله HR است به‌طوری که با همسایه‌هایش همساز باشد. با پیدا کردن ضرایب q از معادله بهینه‌سازی (۲-۴) برای به دست آوردن تصویر HR بازیابی شده از فرمول (۲-۵) استفاده می‌شود.

$$y_h = D_h\, q \tag{۵-۲}$$

۲-۲ استفاده از روش‌های خود یادگیری

اگر رابطه‌ی بین وصله‌های وضوح‌بالا و پایین با توجه به وصله‌های خود تصویر در مقیاس‌های متفاوتی

---

[۱] Texture

[۲] Overlap

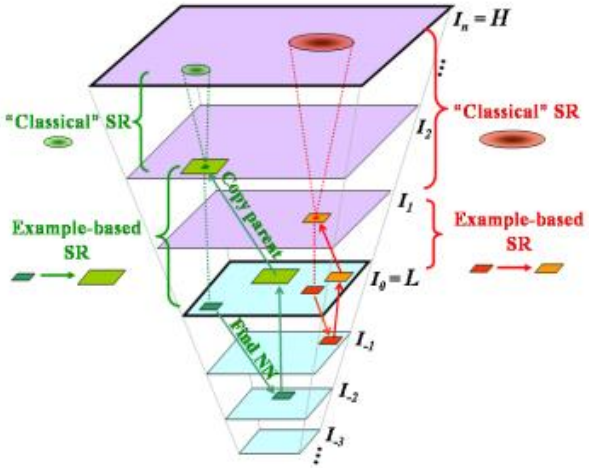
یاد گرفته شود مشکلات مطرح‌شده در روش‌های مبتنی بر پایگاه داده خارجی برطرف می‌گردد. برای این منظور از هرم تصویر استفاده می‌کنند که با توجه به سطوح هرم می‌توان تصویر وضوح‌بالا و پایین داشت. روش‌هایی که اخیراً در حوزه‌ی SR تک فریمی پیشنهاد شده‌اند مبتنی بر ارائه تُنک می‌باشند. ارائه تُنک از یک سیگنال، بر 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 به‌صورت تُنک ارائه شوند به‌طوری که توابع پایه می‌توانند از طریق مجموعه‌ای از داده‌ها یاد گرفته شوند. به‌عنوان مثال یک تصویر می‌تواند به‌صورت ترکیب وزن‌دار از الگوهای تکراری از پیکسل‌ها بیان شود. حال به بررسی مفصل‌تری از چندین تحقیق انجام گرفته‌شده در زمینه‌ی فرا تفکیک‌پذیری خواهیم پرداخت به‌طوری که با استفاده از این روش‌ها بتوانیم ایده‌ی اصلی خود در این پایان‌نامه را ارائه نماییم. در [۱۲] به این نکته اشاره شده است که تکرار وصله‌ها در یک مقیاس از تصویر، پایه و اساس SR کلاسیک است، درحالی‌که 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. 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 (به‌عنوان مثال ۵×۵) 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 همکارانش [۱۲] این فرضیه را به‌صورت آماری بررسی کردند و به این نتیجه رسیدند که به‌طور متوسط ۹۰ درصد وصله‌ها در هر تصویر ۹ وصله مشابه یا بیشتر در همان مقیاس تصویر اصلی دارند. ۸۰ درصد

---

[۱] De-noising
[۲] Inpainting
[۳] Classification

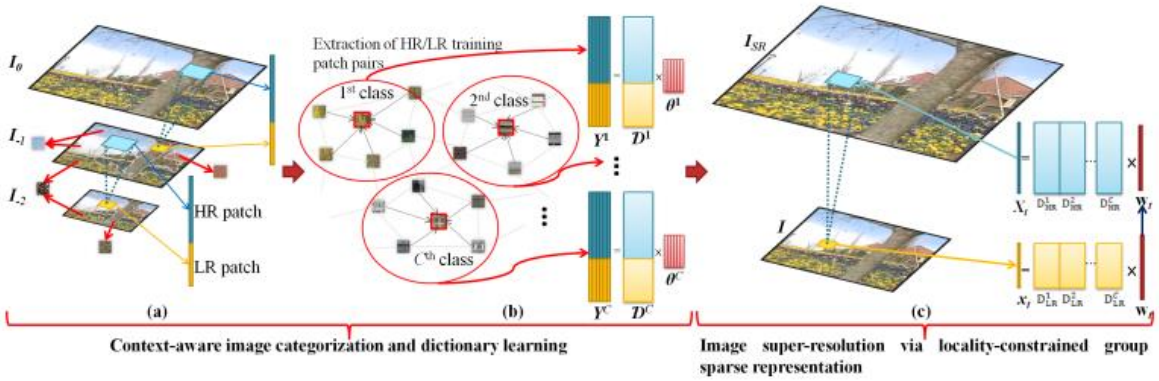
از وصله‌های ورودی ۹ وصله مشابه یا بیشتر در ۰,۴۱ از ابعاد تصویر اصلی و ۷۰ درصد آن‌ها ۹ وصله مشابه در ۰,۲۶ از ابعاد تصویر ورودی دارند. همان‌طور که در این مقاله بررسی شد وصله‌های مشابه در سرتاسر تصویر پخش شده‌اند. به همین ترتیب یک هرم تصویر ساخته شد و با ترکیب SR تک فریمی و چند فریمی محدودیت SR مبتنی بر مثال، که داشتن یک پایگاه داده خارجی ضرورت داشت برطرف گردید. اما این روش بسیار هزینه‌بر است.

Tsai و همکارانش[۶۳] برای به دست آوردن وصله‌های مشابه پیشنهاد دادند که وصله‌ها از بخش‌هایی از تصویر که محتوای مشابه‌ای دارند استخراج شوند. برای به دست آوردن وصله‌های مشابه به‌جای جستجو در کل تصویر، ابتدا تصویر ورودی را ناحیه بندی کرده و وصله‌های مشابه در ناحیه‌ی مربوط به وصله‌ی ذکرشده جست‌وجو می‌شود. نظریه آن‌ها بر این فرض استوار بود که با بخش‌بندی تصویر خطای بازسازی کمتر می‌شود. بنابراین وصله‌هایی که در هر ناحیه از تصویر استخراج شد بسیار به هم مشابه بودند. در این روش برای هر بخش از تصویر یک دیکشنری وضوح‌بالا و یک دیکشنری وضوح پایین ساخته می‌شود. درواقع برای به دست آوردن دیکشنری وضوح پایین از سطوح پایین هرم تصویر استفاده



شکل ۲-۱. ترکیب SR مبتنی بر مثال با SR چند فریمی.

می‌شود. دیکشنری‌های ساخته‌شده شامل وصله‌های وضوح‌بالا و وضوح پایین متناظر می‌باشد که با استفاده از هرم تصویر ساخته شده‌اند.



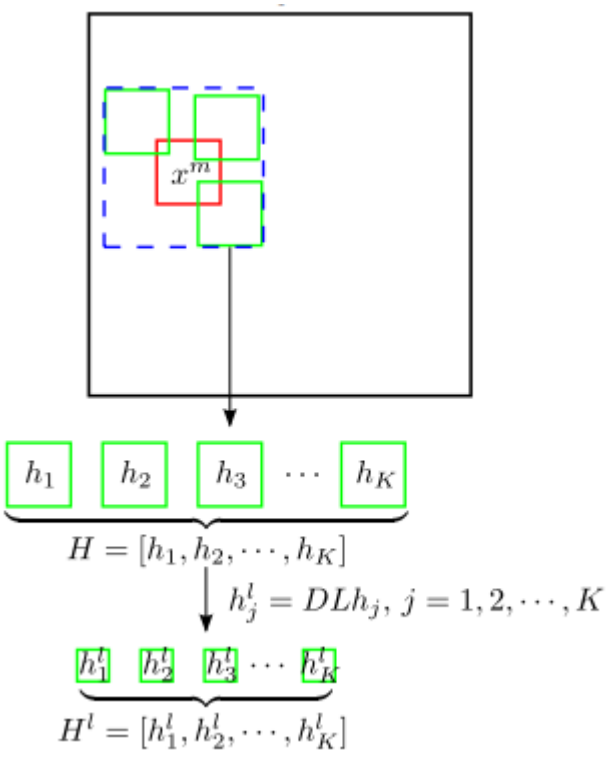
شکل ۲-۲. روش پیشنهادی برای SR تک تصویر در [۶۳]

همان‌طور که درشکل ۲-۲ مشاهده می‌شود ابتدا تصویر ورودی با فاکتور کاهش مقیاس معینی تا دو سطح کوچک می‌شود و بدین ترتیب یک هرم تصویر به وجود می‌آید. سپس بر روی سطح I_{-1} و I_{-2} با استفاده از الگوریتم جابه‌جایی میانگین[۱] [۶۴] عملیات بخش‌بندی انجام می‌گیرد. وصله‌های هر ناحیه از تصویر I_{-2} با اندازه‌ی مشخصی استخراج شده و متناظر هر وصله که از I_{-2} استخراج گردید، وصله‌ای از تصویر I_{-1} با ابعادی که با فاکتور کاهش مقیاس رابطه عکس دارد استخراج می‌شود. به همین ترتیب این کار برای وصله‌های موجود در سطح I_{-1} و I_0 انجام داده می‌شود. سپس این وصله‌های متناظر به‌صورت بردار ستونی بازنمایی می‌شود و برای هر ناحیه از تصویر یک دیکشنری ساخته می‌شود. سپس برای هر وصله‌ی وضوح پایین در تصویر ورودی با متصل کردن تمام دیکشنری‌های وضوح پایین به هم و با استفاده از دو شرط محلی بودن و تنک بودن گروهی[۶۵] ضرایب اسپارس به دست آمدند. با این فرض که ضرایب اسپارس برای دیکشنری‌های وضوح‌بالا و پایین یکسان است[۲۲] ضرایب اسپارس به‌دست‌آمده در دیکشنری وضوح بالایی که از متصل کردن تمام دیکشنری‌های وضوح‌بالا ساخته شده بود ضرب شدند و وصله‌ی وضوح‌بالای متناظر با آن به دست آمدند. این کار را برای تمام وصله‌های تصویر ورودی انجام می‌دهند و تصویر SR نهایی را به دست می‌آورند. فایده این روش این است که

[۱] Meanshift

وصله‌ی وضوح‌بالای هر وصله از تصویر ورودی توسط وصله‌های مشابه آن ساخته می‌شود. اما این کار هزینه محاسباتی بالایی دارد.

chen و همکارانش[۶۶] 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 overlap با سایز B×B تقسیم می‌شود. برای هر وصله در تصویر LR وصله‌ی متناظر آن در تصویر X_m با سایز sB×sB در همان مکان استخراج می‌شود. برای هر وصله در تصویر X_m در یک همسایگی مشخص، تمام وصله‌های اطراف آن استخراج می‌شوند و تمام وصله‌های استخراج شده به‌صورت برداری کنار هم قرار می‌گیرند تا ماتریس hypothesis ساخته شود. تولید این ماتریس hypothesis در شکل ۲-۳ نشان داده شده است.



شکل ۲-۳. تولید ماتریس hypothesis در داخل یک همسایگی مشخص [۶۶]

هم‌زمان با تولید ماتریس hypothesis تمام وصله‌ها به ابعاد sB×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[۶۷] 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[۶۸] یک روش سریع برای SR تک فریمی ارائه دادند. آن‌ها یک پیاده‌سازی کارا بر پایه الگوریتم K-SVD پیشنهاد کردند. در واقع بدون نیاز به پایگاه داده خارجی بزرگ و با استفاده از تصویر ورودی روشی سریع در فرا تفکیک‌پذیری ارائه دادند. 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، فن SMSE[۱] می‌باشد. از معیار بهینه‌سازی مبتنی بر L0[۲] که بسیار سریع‌تر از L1[۳] می‌باشد برای هر دو فاز آموزش و بازسازی استفاده شده است. روش کار به این صورت است که در ابتدا به‌صورت تصادفی وصله‌هایی از تصویر استخراج می‌شوند. تعداد وصله‌هایی که به‌صورت تصادفی استخراج می‌شوند برابر با ۳۰۰۰ وصله است. با استفاده از معیار SMSE وصله‌هایی که اطلاعات بالایی دارند برای ساخت دیکشنری استفاده می‌شود. سپس با استفاده از الگوریتم بهبودیافته K-SVD یک دیکشنری وضوح‌بالا ساخته می‌شود. سپس دیکشنری وضوح پایین با کاهش نمونه‌ها از دیکشنری وضوح‌بالا ساخته می‌شود. برای هر وصله در تصویر ورودی با استفاده از دیکشنری وضوح پایین ضرایب مورد نیاز برای ساخت وصله‌ی SR به دست می‌آید. سرانجام این ضرایب در دیکشنری وضوح‌بالا ضرب می‌شوند تا تصویر SR ساخته شود.

در بیشتر روش‌هایی که برای SR تک فریمی پیشنهاد شده است از شدت روشنایی تصویر برای

---

[۱] Sample mean square error

[۲] نرم L_0 که به صورت $\|\ \|_0$ نشان داده می شود تعداد عناصر غیر صفر یک بردار را نشان می دهد.

[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 می کند:

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

ساخت دیکشنری استفاده می‌شود. یا به عبارتی مؤلفه‌های فرکانس بالا و پایین به‌طور هم‌زمان بازیابی می‌شوند. اما روش‌های دیگری نیز وجود دارند که از تصویر ورودی ویژگی استخراج می‌کنند و به دنبال یادگیری اطلاعات فرکانس بالای موجود در تصویر هستند [۶۹]. این روش بر اساس تحقیق صورت گرفته در مرجع [۶۲] است که روشی سریع برای SR تک فریمی می‌باشد. روش پیشنهاد شده در مقاله [۶۲] شامل دو فاز است: ۱) فاز آموزش ۲) فاز بازسازی

در فاز آموزش ابتدا یک پایگاه داده خارجی شامل تصاویر وضوح‌بالا جمع‌آوری می‌شود و با استفاده از فیلتر کردن و کاهش نمونه‌ها تصاویر وضوح پایین متناظر با آن‌ها ساخته می‌شود. سپس از این تصاویر وصله‌های وضوح‌بالا و پایین متناظر با یکدیگر استخراج می‌شوند. سپس بر روی هر کدام از زوج وصله‌های متناظر یک سری عملیات پیش پردازش انجام می‌شود به‌طوری که از وصله‌های وضوح‌بالا مؤلفه‌های فرکانس بالا حذف می‌شود و از وصله‌های وضوح پایین با استفاده از ۴ فیلتر بالا گذر، ویژگی استخراج می‌شود. سپس از PCA برای کاهش ابعاد ویژگی‌های استخراج شده از وصله‌های وضوح پایین استفاده می‌شود که فرآیند ساخت دیکشنری را سریع‌تر می‌کند. برای وصله‌های وضوح پایین و وضوح‌بالا هرکدام دیکشنری جداگانه‌ای ساخته می‌شود. در فاز بازسازی ابتدا تصویر ورودی با استفاده از روش درون‌یابی bicubic به ابعاد مشخصی بزرگ می‌شود و فیلترهای اعمال‌شده در فاز آموزش بر روی تصویر ساخته شده از مرحله‌ی قبل اعمال می‌شوند. از این چهار تصویر یک وصله در مکان مشخصی استخراج می‌شود و این ۴ وصله به‌صورت برداری به هم متصل می‌شوند و با استفاده از دیکشنری وضوح پایین ضرایب بازسازی هر وصله به دست می‌آید. ضرایب به دست آمده در مرحله قبل در دیکشنری وضوح‌بالا ضرب می‌شود و وصله‌های بازیابی شده در ناحیه‌هایی که باهم همپوشانی دارند متوسط گیری می‌شوند تا تصویر SR نهایی به دست آید.

۳ فصل سوم: تئوری

در این فصل از پایان‌نامه تئوری روش پیشنهادی مطرح می‌شود. ازآنجایی‌که روش پیشنهادی مبتنی بر دیکشنری است ابتدا نحوه‌ی ساخت دیکشنری توضیح داده خواهد شد. سپس راه‌حل جبری برای حل مسئله بهینه‌سازی که روش‌های مبتنی بر تنظیم‌کننده نامیده می‌شوند، معرفی می‌شود. همچنین برای حل مشکلات موجود در زمینه‌ی تنظیم‌کننده‌ها روش‌های بازسازی سراسری مطرح خواهد شد.

۳-۱ هرم تصویر

در روش‌های مبتنی بر خود یادگیری برای ساخت تصویر وضوح‌بالا با استفاده از تصویر ورودی وضوح‌پایین، به تشکیل هرم تصویر نیاز است. اگر رابطه‌ی بین تصویر ورودی و نسخه‌ی کاهش مقیاس یافته‌ی آن یاد گرفته شود می‌توان از این رابطه برای ساخت نسخه‌ی SR تصویر ورودی نیز استفاده کرد. بنابراین اولین گام در فرا تفکیک‌پذیری مبتنی بر خود یادگیری ساخت هرم تصویر است. لازم به ذکر است با هر فاکتوری که تصویر ورودی در هرم کوچک می‌شود، تصویر SR نهایی با عکس همان فاکتور بزرگ می‌شود. مثلاً اگر تصویر ورودی با فاکتور ۰,۵ کوچک شود تصویر SR تولید شده ۲ برابر تصویر اصلی خواهد بود.

۳-۲ دیکشنری

در چند دهه اخیر استفاده از دیکشنری در فرا تفکیک‌پذیری نظر بسیاری از محققان را به خود جلب کرده است. ایده‌ی این روش بر این اساس استوار است که می‌توان یک وصله از تصویر را به‌صورت ترکیب خطی از وصله‌های دیگر بیان کرد. بنابراین باید ضرایب بازسازی هر وصله که در ترکیب خطی شرکت دارد محاسبه شود. برای محاسبه‌ی ضرایب بازسازی هر وصله به یک دیکشنری نیاز است. برای ساخت دیکشنری روش‌های زیادی ارائه شده است. یکی از الگوریتم‌هایی که برای این منظور استفاده می‌شود الگوریتم K-SVD [۷۰]می‌باشد. هزینه محاسباتی این الگوریتم زیاد است و هر چه تعداد اتم‌ها بیشتر باشد زمان ساخت دیکشنری نیز بالاتر می‌رود. راه دیگر این است که مستقیم از وصله‌های تصویر استفاده

شود. به‌عبارتی‌دیگر هر وصله برداری می‌شود و به‌عنوان یک اتم در داخل دیکشنری قرار می‌گیرد. دیکشنری که بر این اساس ساخته می‌شود دیکشنری مبتنی بر وصله نامیده می‌شود. روش دیگر برای ساخت دیکشنری استفاده از ویژگی است. در مقاله [۶۹] با استفاده از ۴ فیلتر از تصویر ورودی ویژگی استخراج شده است. سپس بردار ویژگی‌های به‌دست‌آمده به‌عنوان اتم در داخل دیکشنری قرارگرفته‌اند. اگر هر وصله به‌صورت ترکیب خطی از اتم‌ها (وصله‌هایی که به بردار تبدیل شده‌اند) بیان شود و n تعداد اتم‌ها در دیکشنری باشد، خواهیم داشت:

$$p_i = \propto_1 p_1 + \propto_2 p_2 + \propto_3 p_3 + \cdots = \sum_{j=1}^{n} \propto_j p_j \quad (\mathrm{i} \neq \mathrm{j}) \tag{۱-۳}$$

$$p_i = D_l q \tag{۲-۳}$$

به‌طوری که p_i وصله‌ی وضوح پایین i –ام و $q = (\propto_1 \propto_2 \ \dots \ \propto_n)^T$ بردار ضرایب است که اهمیت هر اتم را برای ساخت نسخه‌ی وضوح‌بالای هر وصله‌ی LR تعیین می‌کند. $D_l = (p_1\, p_2 \ \dots \ p_n)$ دیکشنری وضوح پایین می‌باشد. وصله‌ی بازسازی شده باید تا حد ممکن به وصله‌ی موردنظر نزدیک باشد. بردار ضرایب q به وسیله‌ی حل معادله‌ی بهینه‌سازی (۳-۳) به دست می‌آید:

$$\mathrm{q} = \underset{q}{\operatorname{argmin}} \parallel D_l q - p_i \parallel \tag{۳-۳}$$

۳-۳ معیار SMSE برای فیلتر کردن وصله‌ها

از آنجایی که تعداد وصله‌های زیاد فرآیند ساخت دیکشنری را کُند می‌کند به منظور سرعت بخشیدن به فرآیند ساخت دیکشنری فقط از وصله‌هایی استفاده می‌شود که حاوی جزئیات بالایی باشند. مشابه [۶۸] از معیار SMSE برای انتخاب مؤثر وصله‌ها برای ساخت دیکشنری استفاده می‌شود. معیار SMSE جزئیات هر وصله را اندازه می‌گیرد. فرض می‌شود که تعداد وصله‌ها در هر بلاک برابر N و ابعاد هر وصله $A\times A$ باشد، معیار SMSE از طریق فرمول زیر محاسبه می‌شود:

$$SMSE(p_i) = \frac{\sum_{j=1}^{A^2}(p_i(j) - \sum_{j=1}^{A^2}\frac{p_i(j)}{\mathrm{A}\times\mathrm{A}})^2}{A^2-1} \quad (۴-۳)$$

به‌طوری که $p_i(j)$ عنصر j–ام از وصله‌ی i–ام می‌باشد. این فرمول واریانس روشنایی در وصله‌ی i – ام تصویر را اندازه می‌گیرد. اگر معیار SMSE برای یک وصله، از مقدار آستانه‌ی از پیش تعریف شده بالاتر باشد، از وصله برای ساخت دیکشنری استفاده می‌شود در غیر این صورت از وصله‌ی موردنظر صرف‌نظر می‌شود.

۳-۴ بهینه‌سازی

در حالت کلی هدف از فرا تفکیک‌پذیری تولید تصویر SR (y_h) از تصویر ورودی (z_l) می‌باشد. طبق فرمول (۳-۵) z_l به y_h مربوط می شود:

$$z_l = SHy_h \quad (۵-۳)$$

به‌طوری که S و H به ترتیب اپراتور کاهش نرخ نمونه‌ها و مات‌کنندگی می‌باشند. دیکشنری HR و LR از تصویر ورودی و نسخه‌ی کاهش مقیاس یافته‌ی آن تولید می‌شوند. ازآنجایی‌که اتم‌ها در دیکشنری وضوح‌پایین D_l و دیکشنری وضوح‌بالا D_h به هم مرتبط هستند و از مکان‌های متناظر هم استخراج می‌شوند انتظار می‌رود بردار ضرایب q برای نسخه‌ی HR و LR یک وصله‌ی تصویر تغییر نکند. بنابراین نسخه‌ی SR یک وصله‌ی وضوح پایین با استفاده از دیکشنری HR به‌صورت زیر تولید می‌شود:

$$x_i = D_h q \quad (۶-۳)$$

به‌طوری که x_i نسخه‌ی وضوح‌بالای وصله‌ی p_i می‌باشد.

ازآنجایی‌که در معادله (۳-۳) ماتریس D_l یک ماتریس مربعی نیست مسئله بدحالت است و بی‌نهایت جواب خواهد داشت. برای حل این مسئله از تنظیم‌کننده تیخونوف[۶۷] استفاده شده است.

$$\vec{q} = \underset{q}{arg\min} \| p - D_l q \|_2^2 + \lambda_{tik} \| q \|_2^2 \quad (۷-۳)$$

به‌طوری که λ_{tik} پارامتر تنظیم‌کننده می‌باشد. راه‌حل برای معادله (۳-۷) به‌صورت زیر به دست می‌آید:

$$\vec{q} = ((D_l)^t D_l + \lambda_{tik} I)^{-1} (D_l)^t p \quad (۸-۳)$$

از ضرایب q برای ساخت x_i استفاده می‌شود:

$$x_i = D_h q \quad (۹-۳)$$

به‌طوری که ماتریس پروجکشن برابر خواهد بود با:

$$P_G = ((D_l)^t D_l + \lambda_{tik} I)^{-1} (D_l)^t \quad (۱۰-۳)$$

ماتریس P_G فقط یک بار محاسبه می‌شود، یعنی در طول اجرای الگوریتم SR فقط نیاز است که ماتریس پروجکشن در ورودی LR ضرب شود.

در معادله (۳-۷) به همه‌ی اتم‌ها وزن مشابه داده می‌شود، به این معنی که وصله‌ی ورودی نسبت به تمام اتم‌های دیکشنری با وزن یکسان سنجیده می‌شود. این روش خیلی دقیق نیست و فقط سرعت محاسبه ضرایب q را بالا می‌برد. روش دوم برای حل معادله (۳-۳) استفاده از ماتریس Γ است. Γ ماتریسی است که برحسب فاصله‌ی وصله‌ی ورودی به اتم‌های دیکشنری، به اتم‌ها وزن‌های متفاوتی می‌دهد. بنابراین معادله‌ی (۳-۷) به صورت زیر بازنویسی می‌شود:

$$\vec{q} = \underset{q}{arg\min} \| D_l q - \widetilde{p_l} \|_2^2 + \lambda_{tik} \| \Gamma q \|_2^2 \quad (۱۱-۳)$$

به طوری که λ_{tik} پارامتر تنظیم‌کننده می‌باشد. همان‌طور که در [۷۱] پیشنهاد شده است ماتریس Γ به‌صورت زیر تعریف می‌شود.

$$\Gamma_{jj} = \| p - d_j^L \|_2^2 \quad (۱۲-۳)$$

به‌طوری که d_j^L ستون j-ام دیکشنری D_l می‌باشد. درنتیجه ضرایب q به‌صورت زیر محاسبه می‌شود:

$$\vec{q} = ((D_l)^t D_l + \lambda_{tik} \Gamma \Gamma^t)^{-1} (D_l)^t \widetilde{p_l} \quad (۱۳-۳)$$

نتایج حاصل از اجرای این دو تنظیم‌کننده روی یک تصویر درشکل ۳-۱ نشان داده شده است. همان‌طور که در شکل پیداست اگر از ماتریس Γ استفاده شود اثرات بلوکی کمتری به وجود خواهد آمد. علاوه بر دید بصری این دو تصویر ازنقطه‌نظر PSNR و زمان اجرا نیز مقایسه شده‌اند.

(الف). PSNR=29.8,time= 0.2s (ب). PSNR=30.43,time= 5.6s

شکل ۳-۱. نتایج مقایسه تنظیم‌کننده تیخونوف. (الف) بدون محاسبه ماتریس Γ (ب). با استفاده از ماتریس Γ

با توجه به مطالب بالا در روش پیشنهادی از ماتریس Γ در تنظیم‌کننده تیخونوف استفاده می‌شود.

۳-۵ روش بازسازی سراسری

باید ذکر شود که معادله (۳-۳) به دلیل وجود نویز و این فرض که هر وصله بر اساس وصله‌های

همسایه قابل بیان است، جواب واقعی برای وصله‌ی وضوح‌بالای p_h وجود نخواهد داشت. اگر تصویر به‌دست‌آمده با X_0 نشان داده شود این اختلاف به‌وسیله‌ی پروجکت X_0 بر روی فضای $z_l = SHy_h$ محاسبه می‌شود:

$$y_h{}^* = argmin_{y_h} \ \|SHy_h - z_l\|_2^2 + c\|y_h - X_0\|_2^2 \qquad (۱۴-۳)$$

به طوری که S و H به ترتیب اپراتور کاهش نرخ نمونه‌ها و ماتی می‌باشند. y_h تصویر وضوح بالا، z_l تصویر وضوح پایین، c عددی ثابت است. راه‌حلی که برای این معادله استفاده می‌شود گرادیان نزولی است. بنابراین خواهیم داشت:

$$y_{h_{(t+1)}} = y_{h_{(t)}} + \beta[H^T S^T(z_l - SHy_h) + c(y_h - X_0)] \qquad (۱۵-۳)$$

به‌طوری که $y_{h_{(t)}}$ تخمین تصویر وضوح‌بالا در تکرار t-ام است و گام گرادیان نزولی برابر β خواهد بود. به $y_h{}^*$ تخمین نهایی تصویر SR گفته می‌شود. این تصویر تا حد ممکن به تصویر تخمین اولیه (X_0) نزدیک است. برای افزایش دقت بازسازی سراسری و از بین بردن نویز به معادله (۳-۱۵) ترم پیشین نیز اضافه می‌شود.

$$y_h{}^* = argmin_{y_h} \ \|SHy_h - z_l\|_2^2 + c\|y_h - X_0\|_2^2 + \tau\rho(y_h) \qquad (۱۶-۳)$$

$\rho(y_h)$ تابع جریمه‌ای است که ترم پیشینی را درباره‌ی تصویر y_h به وجود می‌آورد. این تابع می‌تواند Huber، MRF، TV[۱][۷۲]، BTV[۲][۷۳] باشد.

[۱] Total Variation

[۲] Bilateral Total Variation

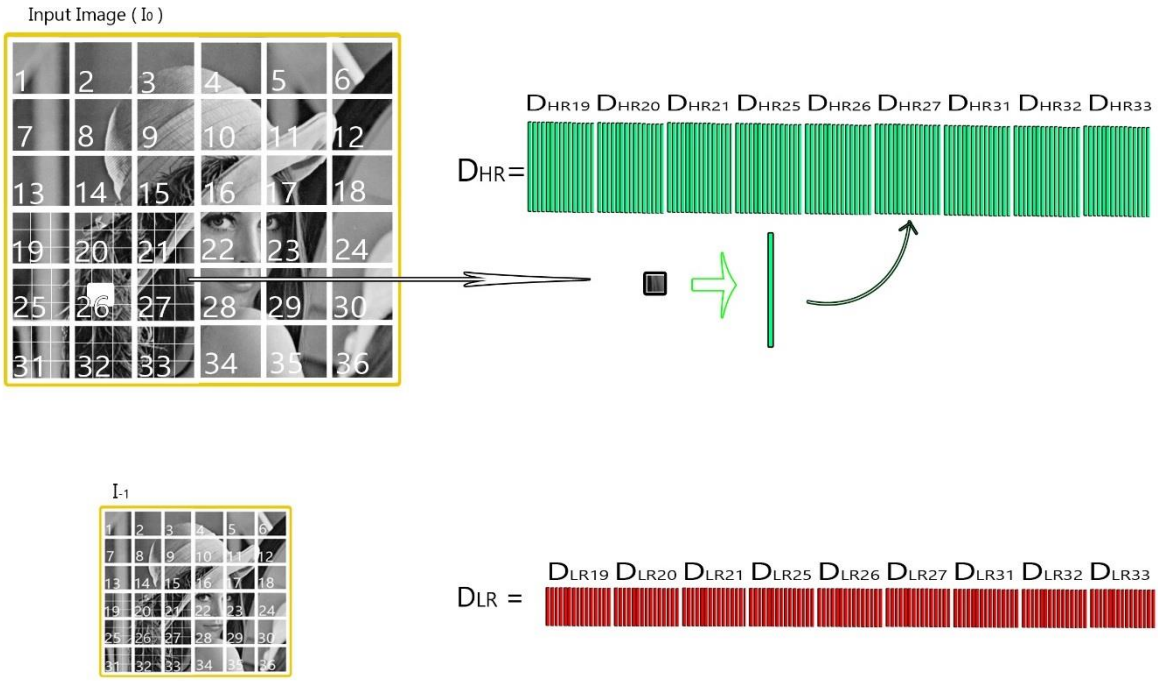
۴ فصل چهارم: روش پیشنهادی

با الهام گرفتن از روش‌های ارائه‌شده در فصل دوم در این بخش از پایان‌نامه، یک روش خود یادگیری سریع برای فرا تفکیک‌پذیری ارائه می‌شود به‌طوری که پیچیدگی محاسباتی روش‌های ذکر شده کاهش پیدا کند. مشابه سایر روش‌های خود یادگیری دیگر، نیازی به جمع‌آوری پایگاه داده خارجی برای یادگیری رابطه‌ی جفت وصله‌های HR/LR وجود ندارد. این رابطه فقط با استفاده از دو سطح از هرم تصویر ورودی یاد گرفته می‌شود. به‌منظور ساخت نسخه وضوح‌بالای هر وصله از تصویر ورودی، یک مجموعه از وصله‌ها که به بلوک‌های همسایه متعلق هستند به‌عنوان دیکشنری وضوح‌بالا جمع‌آوری می‌شود و برای ساخت دیکشنری وضوح پایین از نسخه کاهش مقیاس یافته‌ی آن‌ها استفاده می‌شود. فرض می‌شود شبیه‌ترین وصله‌ها به وصله‌ی موردمطالعه وصله‌های همسایه آن هستند. این فرض در بسیاری از موارد منطقی است. برای ساخت وصله SR به یک دیکشنری نیاز است که تا حد امکان به وصله‌ی موردمطالعه نزدیک باشد. این کار باعث می‌شود خطای بازسازی و پیچیدگی محاسباتی کم باشد. با مرور روش‌های گذشته و بررسی نقاط ضعف آن‌ها می‌توان به این نتیجه دست پیدا کرد که ایده‌ی بخش‌بندی تصویر برای ساخت دیکشنری بسیار مفید است چون فلسفه‌ی به وجود آمدن آن این است که هر وصله بر اساس وصله‌هایی که بسیار شبیه به آن هستند ساخته شود.

با توجه به مشکلات بخش‌بندی، ایده‌ی جدیدی برای ساخت هر وصله با وصله‌های مشابه محلی آن ارائه می‌شود که بر مشکلات غلبه کند، ایده‌ی ما بلاک بندی تصویر است که یک ایده کاملاً جدید است و مشکل وصله‌های مرزی را از بین می‌برد، از طرفی بلاک کردن نسبت به بخش‌بندی زمان کمی می‌خواهد. درنتیجه با ساختن دیکشنری محلی توسط یک بلاک و ۸ بلاک همسایه، اتم‌های تشکیل‌دهنده دیکشنری به هم نزدیک‌تر می‌شوند و خطای بازسازی کمتر می‌شود. برای ساخت دیکشنری الگوریتم‌های زیادی مثل K-SVD [۳۲] ارائه شده‌اند که بسیار زمان‌بر می‌باشند. ایده ما در این پایان‌نامه این است که به‌طور مستقیم، وصله‌های استخراج‌شده از تصویر را به شکل بردار ستونی بازنمایی و کنار هم قرار داده و بدین ترتیب بدون اجرای هیچ الگوریتم اضافی دیکشنری ساخته می‌شود، با توجه به اینکه تعداد اتم‌های دیکشنری ساخته‌شده کم است و اتم‌ها به هم نزدیک هستند دیکشنری

ساخته‌شده تفاوت چندانی با دیکشنری آموزش دیده ندارد، نتایج تجربی ارائه شده گواه این ادعاست. برای یادگیری ارتباط بین وصله‌ی وضوح‌بالا و وضوح پایین، از هرم تصویر استفاده می‌شود. اکثر روش‌ها برای یادگیری رابطه‌ی وصله‌ی وضوح‌بالا و پایین از کاهش نمونه‌برداری وصله‌ی وضوح‌بالا استفاده می‌کنند که این عملیات به دلیل زیاد بودن وصله‌ها بسیار زمان‌بر می‌باشد. درنتیجه ایده‌ی جدیدی برای به دست آوردن وصله وضوح پایین متناظر با هر وصله‌ی وضوح‌بالا ارائه می‌شود. برای حل مشکل ساخت وصله‌های مرزی، از یک همسایگی مشخصی از وصله‌ی موردنظر استفاده می‌شود، درنتیجه هالوسینه کمتری رخ خواهد داد. روش پیشنهادی با جزئیات بیشتر در بخش بعد توضیح داده خواهد شد.

شکل ۴-۱ فرآیند ساخت دیکشنری‌های HR/LR، برای ساخت نسخه‌ی وضوح‌بالای یک وصله‌ی تصویر ورودی را نشان می‌دهد. در این شکل D_{HR} و D_{LR} به ترتیب دیکشنری وضوح‌بالا و وضوح پایین می‌باشند. در این دیکشنری‌ها هر ستون، نشان‌دهنده روشنایی پیکسل‌های وصله‌ی متعلق به بلوکی است که شماره آن در شکل ۴-۱ نشان داده است. همانطور که در شکل ۴-۱ مشخص است برای ساخت وصله بازیابی شده، متناظر با وصله‌ی سفید که در بلوک ۲۶-ام قرار گرفته است از وصله‌های بلوک ۲۶-ام و هشت همسایه اطراف آن استفاده می‌شود. از وصله‌هایی که با یکدیگر همپوشانی دارند استفاده می‌شود اما برای اینکه شکل شلوغ نشود وصله‌ها بدون همپوشانی نشان داده شده است. همچنین بلاک‌ها نیز بدون هم‌پوشانی هستند. هر وصله‌ی استخراج شده (به‌عنوان مثال وصله‌ای که در بلوک ۲۷-ام قرارگرفته و با فلش نشان داده شده است) به یک بردار ستونی تبدیل شده است و با در نظر گرفتن معیار SMSE مستقیم در دیکشنری قرار می‌گیرد.



شکل ۴-۱. روش پیشنهادی برای مهیاکردن دیکشنری HR/LR برای یک وصله تصویر

روش پیشنهادی شامل دو مرحله می‌باشد: ۱) مهیاکردن دیکشنری با استفاده از وصله‌های HR و LR (رویه ۱) و ۲) ساخت تصویر SR (رویه ۲).

رویه ۱: ساخت یک جفت دیکشنری وضوح‌بالا و پایین برای هر بلاک از تصویر
۱- ساخت یک هرم دو سطحی از تصویر ورودی ۲- تصویر HR (I_0) به بلوک‌های بدون همپوشانی تقسیم می‌شود و بلوک‌های متناظر با هرکدام از بلوک‌های HR در تصویر LR (I_{-1}) به دست می‌آید. ۳- برای هر بلاک از تصویر HR با در نظر گرفتن معیار SMSE وصله‌های با جزئیات بالا استخراج می‌شود. ۴- وصله‌های LR متناظر با وصله‌های HR (که از مرحله قبل به دست آمده‌اند) از تصویر LR استخراج می‌شود. ۵- برای هر بلاک HR یک دیکشنری توسط وصله‌های انتخابی از بلاک موردنظر و هشت همسایه اطراف آن ساخته می‌شود. ۶- برای هر بلاک LR یک دیکشنری LR مشابه دیکشنری HR (در مرحله ۵) ساخته می‌شود.

رویه ۲: ساخت تصویر SR
۱- تصویر ورودی I_0 به تعدادی وصله که باهم همپوشانی دارند تقسیم می‌شود. برای هر وصله‌ی تصویر مراحل ۲ و ۳ تکرار می‌شود: ۲- با استفاده از دیکشنری متناظر با هر وصله و تنظیم‌کننده تیخونوف ضرایب هر اتم در دیکشنری محاسبه می‌شود. ۳- ضرایبی که از مرحله ۲ به‌دست آمده‌اند در دیکشنری وضوح‌بالا ضرب می‌شود تا نسخه‌ی SR هر وصله به دست آید.

رویه‌های بالا با جزئیات در زیر توضیح داده می‌شود.

۱-۴ تصاویر LR/HR برای یادگیری جزئیات

ازآنجایی‌که روش پیشنهادی بر اساس خود یادگیری است و هیچ پایگاه داده خارجی وجود ندارد، وجود یک هرم تصویر ضروری است که رابطه‌ی میان وصله‌های وضوح‌بالا و وضوح‌پایین یاد گرفته شود. فرض می‌شود تصویر وضوح پایین (z_l) با استفاده از فرمول زیر به تصویر وضوح‌بالا (y_h) رابطه دارد:

$$z_l = SHy_h \qquad (۱-۴)$$

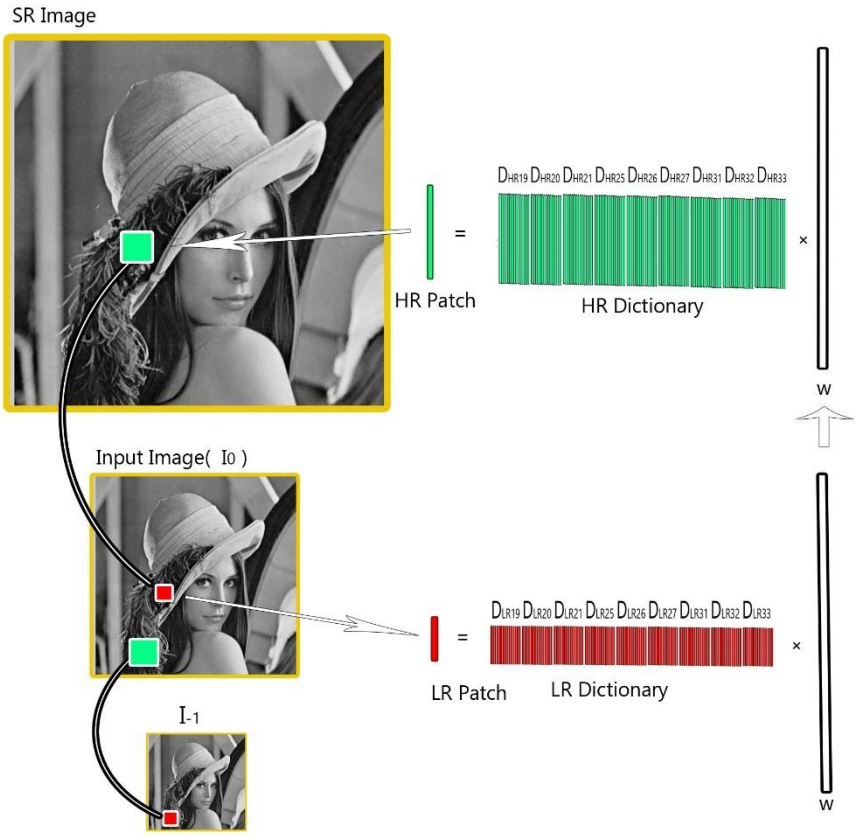
به‌طوری که S و H به ترتیب اپراتور کاهش نمونه‌برداری و مات‌کنندگی می‌باشند. یک هرم تصویر با سه سطح $\{I_{-1}, I_0, I_1\}$ ساخته می‌شود. I_0 تصویر ورودی و I_{-1} از فرمول (۴-۱) به دست می‌آید. تصویر I_1 نسخه‌ی SR است که توسط الگوریتم پیشنهادی به دست می‌آید. ابعاد تصویر ورودی I_0 برابر با m×n است. اگر فرض شود تصویر خروجی و ورودی توسط فرمول(۴-۱) به هم مرتبط می‌شوند می‌توان این رابطه را بین I_0 و I_{-1} یاد گرفت و سپس این رابطه را به I_0 و I_1 نیز اعمال کرد. اگر هر وصله از

تصویر ورودی به‌صورت ترکیب خطی از اتم‌های یک دیکشنری وضوح پایین بیان شود این منطقی است که همین رابطه بین نسخه‌ی وضوح‌بالای SR و اتم‌های دیکشنری HR نیز وجود داشته باشد. همان‌طور که قبلاً هم توضیح داده شد یک دیکشنری وضوح‌بالا و وضوح‌پایین برای هر بلاک از تصویر با استفاده از وصله‌های درون بلاک مربوطه و نسخه وضوح پایین آن، ساخته می‌شود. در این کار تصویر ورودی به‌عنوان تصویر HR و نسخه‌ی وضوح پایین آن به‌عنوان تصویر LR در نظر گرفته می‌شود. این پیشنهاد در تضاد با روش ارائه‌شده در[۶۶] است به‌طوری که تصویر ورودی با استفاده از درون‌یابی bicubic برای ساخت نسخه‌ی HR بزرگ می‌شود. در حقیقت در روش پیشنهادی به‌جای بزرگ کردن تصویر ورودی برای یادگیری رابطه‌ی بین تصاویر LR/HR، تصویر ورودی با فاکتور S کوچک می‌شود، زیرا نسخه وضوح‌بالای HR تصویر ورودی با استفاده از bicubic دارای اثرات مصنوعی می‌باشد. به‌منظور بازسازی نسخه‌ی SR برای هر وصله‌ی تصویر به دو دیکشنری وضوح بالا و پایین از وصله‌های HR و LR نیاز است که به کمک آن‌ها بتوان رابطه‌ی بین وصله‌های LR و HR را یاد گرفت. برای این منظور وصله‌ها در یک دیکشنری (اتم‌ها)، تا حد ممکن باید به وصله‌ی تصویری که قرار است نسخه وضوح‌بالایش ساخته شود، نزدیک باشند. به‌عبارتی‌دیگر، هزینه محاسباتی برای مدل کردن رابطه‌ی بین وصله‌های HR/LR باید قابل‌کنترل باشد. همان‌طور که در بخش بعد خواهید دید هر وصله از تصویر به‌صورت ترکیب خطی از وصله‌های همسایه‌هایش بیان می‌شود. برای داشتن وصله به تعداد کافی و همچنین افزایش دقت، هر وصله در داخل هر بلاک به‌وسیله‌ی تمام وصله‌های بلاک موردنظر و ۸ همسایه اطراف آن بیان می‌شود. وصله‌های هر بلاک به‌صورت جداگانه استخراج می‌شود. لازم به ذکر است برای ساخت دیکشنری فقط از وصله‌هایی که حاوی جزئیات بالاتری هستند استفاده شده است. تصویر به تعداد مشخصی بلاک که باهم همپوشانی ندارند تقسیم می‌شود. برای این منظور قبل از اینکه عملیات تقسیم‌بندی انجام شود ابتدا به‌اندازه‌ی ابعاد یک بلاک در اطراف تصویر از عملیات آینه کردن استفاده می‌شود. تصویر ورودی و نسخه کاهش مقیاس یافته‌ی آن به ترتیب با y_h و z_l بیان می‌شود. استخراج بلاک از تصویر به‌صورت زیر فرمول‌بندی می‌شود:

$$B_i^H = Q_i^H y_h \quad (۲-۴)$$

$$B_i^L = Q_i^L z_l \quad (۳-۴)$$

به‌طوری که B_i^H بلوک i–ام در تصویر HR و B_i^L بلاک متناظر آن در تصویر LR می‌باشد. Q_i^H و Q_i^L اپراتور استخراج بلوک i–ام برای تصاویر HR و LR می‌باشند. ابعاد ماتریس B_i^H برابر است با $\frac{m}{b}\times\frac{n}{b}$ و ابعاد ماتریس B_i^L برابر است با $\frac{m}{s\times b}\times\frac{n}{s\times b}$ به‌طوری که تعداد بلاک‌ها $b\times b$ می‌باشد.



شکل ۴-۲. روش پیشنهادی SR تک تصویر در فاز بازسازی. D_{LRi} و D_{HRi} دیکشنری HR و LR هستند که به بلاک i–ام تعلق دارند. رابطه‌ی بین وصله‌ها در $\{I_0\}$ و $\{I_{-1}\}$ برای ساخت تصویر SR $\{I_1\}$ استفاده می‌شود.

۲-۴ استخراج وصله‌ها

در قسمت قبل در مورد نحوه‌ی تقسیم‌بندی تصویر ورودی و نسخه‌ی کاهش مقیاس یافته‌ی آن به تعدادی بلوک‌های غیر هم‌پوشان صحبت شد. برای هر بلاکی از تصویر HR و LR وصله‌هایی که با یکدیگر همپوشانی دارند استخراج می‌شود. بنابراین اگر پیکسل بالا سمت چپ هر وصله، درون هر بلاکی بیفتد آن وصله متعلق به آن بلاک خواهد بود. در [۶۶] یک وصله‌ی وضوح پایین به‌وسیله کاهش مقیاس دادن وصله‌ی HR متناظرش به دست می‌آید. ازآنجایی‌که تعداد وصله‌ها زیاد است هزینه محاسباتی این فرآیند بسیار زیاد می‌باشد. به‌منظور کاهش پیچیدگی محاسباتی این فرآیند پیشنهاد می‌شود به‌جای اینکه تک‌تک وصله‌های تصویر کوچک شود تصویر ورودی و نسخه کاهش مقیاس یافته‌ی آن شیفت پیدا کند. رویه‌ی ۳ فن پیشنهادی برای این فرآیند را توضیح می‌دهد. در این رویه ابعاد وصله‌های LR و HR به ترتیب برابر است با a×a و A×A ($A = a \times s$). تعداد همپوشانی بین دو وصله‌ی HR با W برابر است.

رویه ۳: استخراج وصله‌ها

۱- ابتدا ماتریس Z ساخته می‌شود.

$$Z = \begin{bmatrix} 0 & 0 \\ 0 & A-W \\ A-W & 0 \\ A-W & A-W \end{bmatrix}$$

۲- K=1 و i=1 تنظیم می‌شود.

۳- y_h به اندازه‌ی $Z(i, :)$ شیفت داده می‌شود و حاصل در T^H قرار می‌گیرد.

۴- T^H با فاکتور S کاهش نمونه داده می‌شود و حاصل در T^L قرار می‌گیرد.

۵- مراحل ۶ تا ۸ تا زمانی که تمام وصله‌ها استخراج شوند تکرار می‌شود:

۶- تمامی وصله ها با سایز $A \times A$ که به بلاک K-ام تعلق دارند از T^H استخراج می‌شود و با در نظر گرفتن معیار SMSE در دیکشنری وضوح بالای K-ام قرار می‌گیرد.

$D_k^H = \{p_{1_k}^H, p_{2_k}^H, \ldots\}$

۷- تمامی وصله‌ها با سایز $a \times a$ متناظر با وصله‌های استخراج شده در مرحله‌ی قبل از T^L استخراج می‌شود و دیکشنری وضوح پایین مربوط به K-امین بلاک ساخته می‌شود.

$D_k^L = \{p_{1_k}^L, p_{2_k}^L, \ldots\}$

۸- K=K+1

۹- i=i+1 و اگر $i < ۵$ به مرحله‌ی ۳ برو.

Z (I, :) برداری است که تعداد شیفت تصویر در طول ستون‌ها (در جهت راست) و در طول سطرها (در جهت پایین) را نشان می‌دهد. D_k^H و D_k^L K-امین دیکشنری HR و LR می‌باشند که متعلق به K-امین بلاک تصویر می‌باشند. $p_{1_k}^H$ و $p_{1_k}^L$ اولین وصله‌ی HR و LR می‌باشند که متعلق به بلاک K -ام هستند و به‌صورت برداری در دیکشنری HR و LR قرار گرفته‌اند. درنتیجه‌ی اجرای این الگوریتم تعداد دفعاتی که از عملیات کاهش مقیاس استفاده می‌شود به ۴ مرتبه می‌رسد.

۳-۴ مقایسه سرعت و PSNR روش پیشنهادی با نتایج مرجع[۶۶]

اگر دو فن ارائه‌شده در این بخش، بر روی الگوریتم پیشنهادشده در مقاله [۶۶] اعمال شود، به‌طور چشمگیری باعث افزایش سرعت اجرای برنامه می‌شود. در این قسمت مقایسه‌ای بین این دو روش انجام می‌گیرد.

PSNR = 29.97, Time = 4.65s

الف

PSNR = 30.62, Time = 174,2s

ب

شکل ۴-۳. مقایسه روش پیشنهادی و روش ارائه‌شده در [۶۶]. (الف) روش پیشنهادی، (ب)روش ارائه‌شده در [۶۶]

PSNR=28.96,Time=4.46

الف

PSNR=29.48,Time=174.07

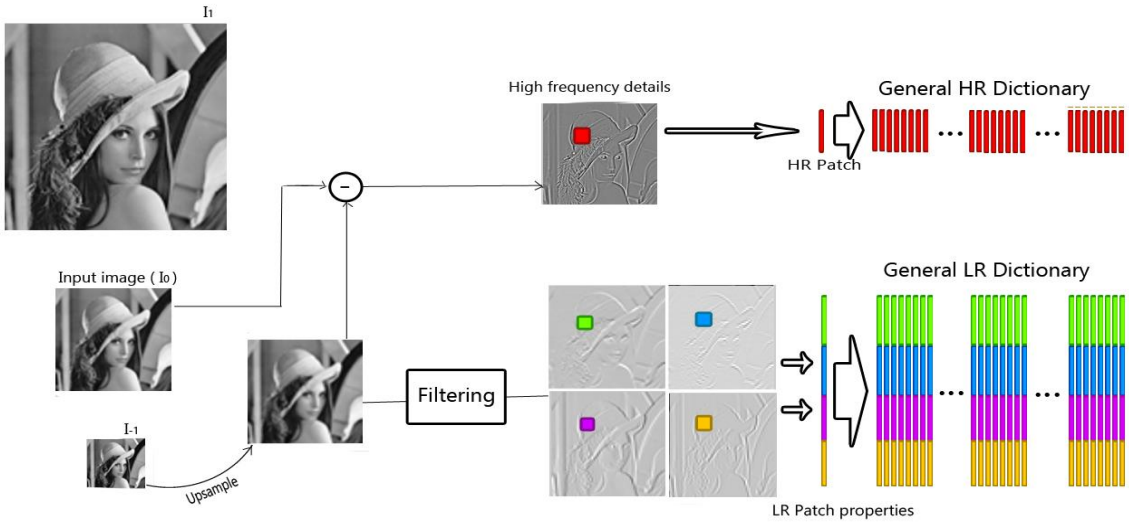
ب

شکل ۴-۴. مقایسه روش پیشنهادی و روش ارائه‌شده در [۶۶]. (الف) روش پیشنهادی، (ب)روش ارائه‌شده در [۶۶]

همان‌طور که مشخص است با الگوریتم پیشنهادی زمان از ۱۷۰ ثانیه به ۴ ثانیه می‌رسد. در بخش بعدی سعی می‌کنیم نتایج حاصل از الگوریتم پیشنهادی را به لحاظ بصری بهبود دهیم.

۴-۴ بهبود روش پیشنهادی ازنقطه‌نظر PSNR و کیفیت بصری

در بخش قبل روش پیشنهادی ازلحاظ سرعت بررسی شد. اما در این بخش، روش پیشنهادی ازنقطه‌نظر PSNR و وضوح بصری بهبود داده می‌شود. یکی از روش‌هایی که بر روی بهبود روش پیشنهادی تأثیر می‌گذارد ساخت دیکشنری با استفاده از ویژگی است. یکی از ویژگی‌هایی که معمولاً استفاده می‌شود مؤلفه‌های روشنایی می‌باشند. ویژگی دیگری که در بیشتر مقاله‌ها استفاده شده است، مشتق مرتبه اول و دوم وصله‌ها می‌باشد. برای استخراج ویژگی در روش پیشنهادی از ویژگی‌هایی که zeyde و همکارانش[۶۹] پیشنهاد داده‌اند، استفاده شده است. بیشتر روش‌های SR شبیه [۶۶, ۷۴] مؤلفه‌های وضوح‌بالا و وضوح‌پایین وصله‌ها را به‌طور هم‌زمان می‌سازند. هدف از این بخش یادگیری جزئیات فرکانس بالاست. به همین منظور مؤلفه‌های فرکانس پایین از تصویر ورودی حذف می‌شود و سپس از تصویر به‌دست‌آمده برای ساخت دیکشنری استفاده می‌شود. این کار سبب می‌شود روال دقیقی برای مینیمم کردن تابع هدف به وجود آید. بنابراین روش پیشنهادی بر پایه ساخت دیکشنری است، اما برای ساخت دیکشنری برخلاف بخش قبل که مستقیم از خود وصله‌های تصویر استفاده می‌شد از ویژگی‌های استخراج‌شده از هرکدام از وصله‌ها استفاده می‌شود. در این روش ابتدا تصویر ورودی به تعدادی مشخصی بلوک تقسیم می‌شود. برای هر وصله در تصویر ورودی یک دیکشنری محلی وضوح‌بالا از ۸ بلوک همسایه اطراف آن و خود بلوک موردنظر درست می‌شود. همچنین یک دیکشنری وضوح پایین از ویژگی‌های استخراج‌شده از تصاویر LR در مکان متناظر با وصله‌های موجود در دیکشنری وضوح‌بالا، ساخته می‌شود. سپس با یادگیری رابطه‌ی میان ویژگی‌ها از وصله‌ی LR و دیکشنری LR، وصله‌ی وضوح‌بالا با استفاده از دیکشنری وضوح‌بالا به دست می‌آید. شمای کلی از الگوریتم پیشنهادی در شکل ۴-۵ نشان داده شده است.



شکل ۴-۵. فاز آموزش: برای ساخت دیکشنری وضوح پایین وصله‌های LR به‌وسیله‌ی بزرگ‌نمایی نسخه‌ی کاهش مقیاس یافته‌ی تصویر ورودی به دست می‌آید $\{I_{-1}\}$. سپس تصویر به‌دست‌آمده توسط ۴ فیلتر بالا گذر فیلتر می‌شود.

روش پیشنهادی شامل دو فاز است : ۱) فاز آموزش(رویه ۴) ۲) فاز بازیابی(رویه ۵). مراحل موجود در هر فاز به‌صورت زیر بیان می‌شود.

رویه ۴: فاز آموزش
۱- ساخت مجموعه آموزشی: تصویر ورودی به‌عنوان تصویر HR در نظر گرفته می‌شود $\{y_h\}$. تصویر ورودی را مات کرده و سپس آن را کوچک کرده تا نسخه وضوح پایین آن به دست آید. دوباره آن را بزرگ کرده تا به اندازه تصویر ورودی برسد. با این کار تصویر وضوح پایین $\{y_l\}$ به دست می‌آید.
۲- پیش‌پردازش: مؤلفه‌های فرکانس پایین از تصویر $\{y_h\}$ حذف می‌شود تا تصویر $\{X_h\}$ به دست آید. تصویر $\{y_l\}$ توسط ۴ فیلتر بالاگذر فیلتر می‌شود بنابراین تصاویر $\{X_l\}_i$ تولید می‌شوند، به طوری که $i = 1,2,\dots,4$.
۳- ساخت دیکشنری: تصویر $\{X_h\}$ به تعداد مشخصی بلوک (بدون همپوشانی) تقسیم می‌شود و بلوک‌های متناظر با هرکدام از این بلوک‌ها از $\{X_l\}_i$ به دست می‌آید.
۴- وصله‌های هر بلاک از تصویر $\{X_h\}$ استخراج می‌شود.
۵- وصله‌های LR متناظر با وصله‌های HR (که از مرحله قبل به دست آمده‌اند) از تصاویر $\{X_l\}_i$ استخراج می‌شود.

۶- برای هر بلاک در تصویر $\{X_l\}_i$ یک دیکشنری توسط وصله‌های آن بلاک و هشت همسایه اطراف آن ساخته می‌شود $\{D_l\}$.
۷- برای هر بلاک موجود در تصویر $\{X_h\}$ یک دیکشنری HR مشابه دیکشنری که در مرحله ۵ تولید شده است به دست می‌آید $\{D_h\}$.

رویه ۵: فاز بازیابی

۱- پیش‌پردازش: تصویر $\{y_h\}$ بزرگ می‌شود و سپس توسط ۴ فیلتر بالا گذر فیلتر می‌شود تا تصویر $\{W_l\}_i$ به دست آید به‌طوری که $i = 1,2, \dots ,4$.
۲- از $\{W_l\}_i$ وصله‌های LR متناظر استخراج می‌شود، آن‌ها را به هم متصل کرده تا وصله‌ی $\widetilde{P_l}$ به دست آید.
۳- ضرایب بازسازی هر وصله با توجه به دیکشنری $\{D_l\}$ به دست می‌آید.
۴- ضرایب به‌دست‌آمده از مرحله‌ی قبل در دیکشنری D_h ضرب می‌شود تا وصله‌ی وضوح‌بالای متناظر با آن به دست آید.
۵- سرانجام وصله‌های HR در نقاطی که با یکدیگر همپوشانی دارند متوسط گیری می‌شوند و تصویر SR اولیه به دست می‌آید.
۶- با استفاده از تنظیم‌کننده BTV و back projection تصویر SR نهایی تولید می‌شود.

جزئیات مراحل فوق در ادامه توضیح داده می‌شود.

۴-۴-۱ ساخت مجموعه آموزشی

در فاز آموزش تصویر ورودی به‌عنوان تصویر وضوح‌بالا در نظر گرفته می‌شود $\{y_h\}$. تصویر را مات کرده و با فاکتور S کوچک می‌کنیم. این عملیات سبب تولید تصویر $\{z_l\}$ می‌شود. سپس این تصویر با استفاده از روش درون‌یابی bicubic به ابعاد تصویر اصلی برگردانده می‌شود که آن را با اپراتور Q نشان می‌دهیم. درنتیجه نسخه وضوح پایین تصویر ورودی در ابعاد تصویر اصلی به دست می‌آید. خواهیم داشت:

$$y_l = Qy_l = Qz_l = Q(SHy_h + v) = QSHy_h + Qv = Ly_h + \tilde{v} \quad (۴-۴)$$

به‌طوری‌که z_l و y_h به ترتیب تصویر وضوح‌پایین و وضوح‌بالا، Q اپراتور درون‌یابی، S فاکتور کاهش-مقیاس، H اپراتور مات کنندگی و v نویز می‌باشد. این نکته مهم است که اپراتور S، H و Q باید در هر دو فاز آموزش و بازیابی یکسان باشند.

۴-۴-۲ پیش‌پردازش و استخراج ویژگی

به‌منظور ساخت نسخه‌ی SR تصویر ورودی، یادگیری لبه و بخش‌های فرکانس بالا برای بازیابی جزئیات و اطلاعات بافت بسیار ضروری است[۷۰, ۷۵]. می‌خواهیم بر روی جزئیاتی تمرکز کنیم که درون‌یابی bicubic قادر به بازسازی آن‌ها نمی‌باشد. بنابراین مؤلفه‌های فرکانس پایین به‌وسیله‌ی محاسبه تفاوت تصاویر $\{y_h\}$ و $\{y_l\}$ به دست می‌آید. مرحله بعدی درروش پیشنهادی پیش‌پردازش است که توسط ۴ فیلتر بالاگذر انجام می‌گیرد. به‌جای اینکه هر وصله به صورت جداگانه فیلتر شود مرحله‌ی پیش‌پردازش مستقیم بر روی کل تصویر اعمال می‌شود، سپس وصله‌ها استخراج می‌شوند. پیش‌پردازشی که بر روی تصاویر وضوح‌بالا اعمال می‌شود شامل حذف مؤلفه‌های فرکانس پایین است که به‌وسیله‌ی تفاوت تصاویر y_h و y_l محاسبه می‌شود. در نتیجه:

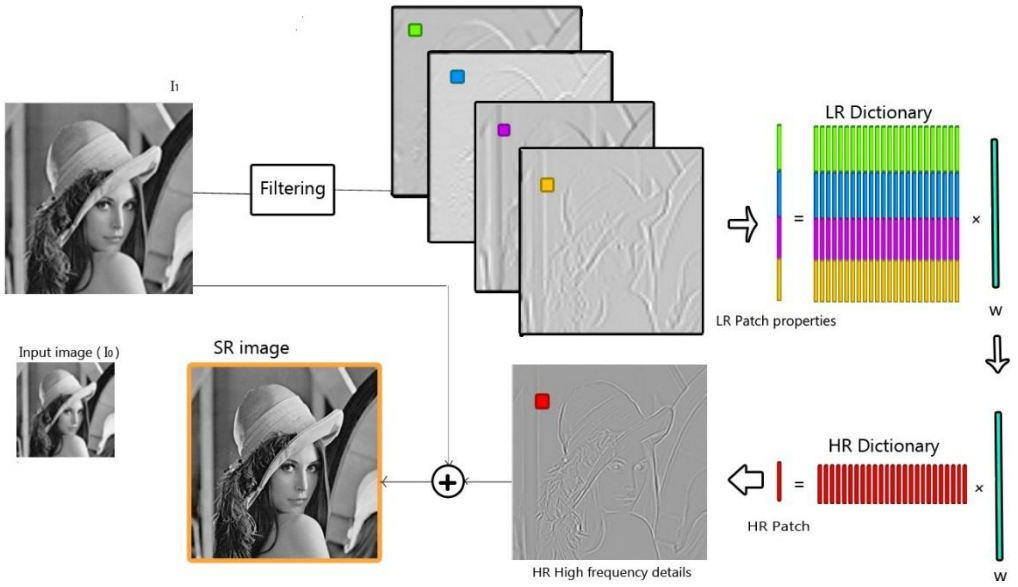
$$e = y_h - y_l \quad (۵-۴)$$

برای انجام مرحله پیش‌پردازش بر روی تصاویر LR، از ۴ فیلتر بالاگذر استفاده می‌شود. بنابراین تصویر وضوح پایین y_l به‌صورت زیر فیلتر می‌شود:

$$\{X_l\}_i = f_i * y_l \quad (۶-۴)$$

به‌طوری که i =1,2,…,4 است. اپراتور * نشان‌دهنده‌ی کانولوشن می‌باشد. فیلترهایی که معمولاً استفاده می‌شوند فیلتر لاپلاسین و گوسی می‌باشند. سپس همانند روشی که در بخش قبل توضیح داده شد تصاویر LR و HR بلوک‌بندی می‌شوند. سپس وصله‌های درون هر بلاک برای ساخت دیکشنری

وضوح‌بالا استخراج می‌شوند.



شکل ۴-۶. فاز بازیابی: وصله‌ها از تصاویر فیلتر شده استخراج می‌شوند و ضرایب بازسازی برای ساخت وصله‌ی SR به دست می‌آید.

۴-۴-۳ استخراج وصله‌ها و ساخت دیکشنری وضوح‌بالا

بعد از مرحله‌ی پیش‌پردازش جفت وصله‌های HR/LR استخراج می‌شوند. وصله‌های HR از $\{X_h\}$ در مکان (r,t) با ابعاد a×a استخراج می‌شود. وصله‌های LR از ۴ تصویر وضوح پایین $\{X_l\}_i$ (i=1,2,..,4) با ابعاد a×a و در مکان (r,t) استخراج می‌شوند. وصله‌ها در تصویر HR/LR با یکدیگر همپوشانی دارند. برای ساخت دیکشنری LR، ۴ وصله‌ی LR برداری شده و به هم متصل می‌شوند. با توجه به اینکه وصله‌ها در کدام بلوک قرار می‌گیرند از هشت همسایه‌ی اطراف آن نیز وصله استخراج می‌شود و همه‌ی آن‌ها کنار هم قرارگرفته و دیکشنری LR ساخته می‌شود. هم‌زمان دیکشنری HR به‌وسیله‌ی برداری-کردن وصله‌های متناظر با هر وصله‌ی وضوح پایین ساخته می‌شود، به این معنی که i-امین اتم در دیکشنری LR به i-امین اتم در دیکشنری HR مرتبط می‌شود. لازم به ذکر است که برای استخراج وصله‌ها از الگوریتم ارائه‌شده در بخش ۴-۲ استفاده می‌شود.

۴-۴-۴ فاز بازیابی

هدف از فاز بازیابی این است که تصویر وضوح پایین z_l بزرگ شود. فرض بر این است که این تصویر از y_h با همان اپراتور مات‌کنندگی و کاهش مقیاسی که در فاز آموزش استفاده شد، ساخته می‌شود. در این فاز مراحل زیر انجام می‌گیرد:

تصویر ورودی با فاکتور S و اپراتور درون‌یابی bicubic بزرگ می‌شود. تصویر حاصله $\{y_l\}$ نامیده می‌شود. تصویر y_l توسط ۴ فیلتر بالاگذری که در فاز آموزش استفاده شد فیلتر می‌شود. در این ۴ تصویر به‌دست‌آمده وصله‌های متناظر هم استخراج می‌شوند. این وصله‌ها برداری شده و زیر هم قرار داده می‌شوند $\{\widetilde{P_l}\}$. سپس با استفاده از دیکشنری LR ساخته‌شده در فاز آموزش ضرایب بازسازی وصله موردنظر به دست می‌آید.

$$y_l = SHy_h \tag{۷-۴}$$

به‌طوری که S و H اپراتور کاهش مقیاس و مات‌کنندگی می‌باشند. سپس می‌توان هر وصله‌ی LR را به‌صورت ترکیب خطی از اتم‌های دیکشنری LR بیان کرد:

$$p_l = D_l q \tag{۸-۴}$$

به‌طوری که p_l وصله‌ی LR و q بردار ضرایب بازسازی و D_l دیکشنری وضوح پایین می‌باشد. ازآنجایی‌که اتم‌ها در D_l و D_h به هم مرتبط هستند خواهیم داشت:

$$p_h = D_h q \tag{۹-۴}$$

p_h وصله‌ی وضوح‌بالا و D_h دیکشنری وضوح بالا می‌باشد. برای حل معادله (۴-۸) و به دست آوردن ضرایب q از فرموله بهینه‌سازی زیر استفاده می‌شود.

$$q = argmin_q \|D_l q - p_l\|_2 \tag{۱۰-۴}$$

ازآنجایی‌که در معادله (۴-۱۰) D_l ماتریس مربعی نیست، بی‌نهایت جواب وجود خواهد داشت. برای حل این مسئله به تنظیم‌کننده نیاز داریم. برای این منظور از تنظیم‌کننده تیخونوف استفاده می‌شود. همان‌طور که در فصل ۳ توضیح داده شد از ماتریس Γ استفاده می‌شود. از طرفی برای استفاده از شرط بازسازی سراسری و از بین بردن اثرات بلوکی از تنظیم‌کننده BTV استفاده می‌شود[۷۳]. بنابراین معادله‌ی (۴-۱۰) به‌صورت زیر بازسازی می‌شود:

$$y_h{}^* = argmin_{y_h} \ \|SHy_h - z_l\|_2^2 + c\|y_h - X_0\|_2^2 + \tau BTV \qquad (۴-۱۱)$$

۵ فصل پنجم: نتایج تجربی

۵-۱ آزمایش‌ها

در این بخش عملکرد الگوریتم پیشنهادشده با سایر روش‌های ارتقای وضوح موجود مقایسه می‌شود. بازدهی روش پیشنهادی بر روی تصاویری که تاکنون در فرا تفکیک‌پذیری به کار رفته است بررسی می‌شود. برای این منظور از تصاویر مختلفی که در پایگاه داده USC-SIPI وجود دارد استفاده شده است. از ۱۸ تصویر متفاوت در این پایگاه داده برای ارزیابی استفاده شده است. برای تصاویر رنگی الگوریتم پیشنهادی فقط بر روی کانال روشنایی اعمال شده است. تصویر ورودی به‌عنوان نسخه‌ی وضوح‌بالا در نظر گرفته‌شده و برای ساخت نسخه‌ی وضوح پایین، تصویر ورودی با فاکتور کاهش مقیاس S و با استفاده از درون‌یابی bicubic کوچک‌شده است. از معیار پیک سیگنال به پیک نویز[۱] و همچنین RMSE[۲]،SSIM[۳] به‌عنوان معیار سنجش کیفیت تصاویر تخمین زده‌شده استفاده شده است. همچنین برای بررسی پیچیدگی محاسباتی روش پیشنهادی نسبت به روش‌های موجود، زمان موردنیاز برای اجرای هرکدام از الگوریتم‌های ذکرشده باهم مقایسه می‌شوند.

هرکدام از معیارهای PSNR ، RMSE و SSIM به‌صورت زیر محاسبه می‌شوند:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2 mn}{\left\|\widehat{\mathrm{X}} - \mathrm{y}_h\right\|^2}\right) \qquad (۱-۵)$$

که mn تعداد کل پیکسل‌های تصویر HR است. $\widehat{X}$ و y_h به ترتیب تصویر HR ساخته‌شده و تصویر اصلی می‌باشند.

SSIM یک معیار مبتنی بر مشاهده است که خرابی‌های تصویر را به‌عنوان تغییراتی که در اطلاعات ساختاری تصویر روی می‌دهد اندازه می‌گیرد. اطلاعات ساختاری به این معنی است که پیکسل‌ها باید

---

[۱] PSNR

[۲] Root mean squared error

[۳] Structural similarity

وابستگی قوی به هم داشته باشند به‌ویژه اینکه ازلحاظ فضایی به هم نزدیک باشند. این وابستگی، اطلاعات مهمی درباره‌ی ساختار اشیا در طبیعت هستند. SSIM بر روی پنجره‌های متفاوتی از یک تصویر محاسبه می‌شود. بر روی دو پنجره‌ی x و y به ابعاد N×N به‌صورت زیر محاسبه می‌شود:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (۲-۵)$$

به‌طوری که μ_x متوسط تصویر x، μ_y متوسط تصویر y، σ_x^2 واریانس تصویر x و σ_y^2 واریانس تصویر y می‌باشد. σ_{xy} کوواریانس x و y می باشد.

$$C_1 = (K_1 L)^2 \quad (۳-۵)$$

$$C_2 = (K_2 L)^2 \quad (۴-۵)$$

L رنج دینامیکی از مقادیر پیکسل‌ها است که به‌طورمعمول برابر است با:

$$L = 2^{\#bits\ per\ pixel} - 1 \quad (۵-۵)$$

به‌طوری که K_1 و K_2 به ترتیب ۰,۰۱ و ۰,۰۳. می‌باشد.

پارامتر دیگر RMSE می‌باشد. درواقع RMSE تفاوت میان مقادیر تخمین زده‌شده توسط مدل و مقادیر واقعی را اندازه می‌گیرد. RMSE تخمین‌گر $\hat{\theta}$ نسبت به پارامتر تخمین زده‌شده‌ی θ به‌صورت زیر تعریف می‌شود.

$$RMSE(\hat{\theta}) = \sqrt{MSE(\hat{\theta})} = \sqrt{E(\hat{\theta} - \theta)^2} \quad (۶-۵)$$

البته پارامترهای PSNR، RMSE، SSIM به‌تنهایی نمی‌توانند معیار مناسبی برای بررسی عملکرد الگوریتم‌های SR باشند 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 bicubic و NEDI همچنین چندین روش SR همانند روش‌های [۶۶] و [۶۸]،مقایسه شده است. برای مقایسه، مستقیم از کد برنامه منتشرشده توسط نویسندگان مقاله‌های ذکرشده استفاده شده است. تصویر ورودی با فاکتور ۲ بزرگ شده است. در تصویر وضوح‌پایین و وضوح‌بالا وصله‌هایی با ابعاد ۴×۴ استخراج شده‌اند. در جدول ۵-۱ نتایج مقایسه PSNR نشان داده شده است. همان‌طور که در این جدول مشاهده می‌شود ازنظر PSNR نسبت به سایر روش‌ها بالاتر هستیم. می‌توان بیان کرد از نقطه‌نظر PSNR نسبت به روش موجود در [۶۶] و [۶۸] بالاتر هستیم و ازلحاظ بصری کاملاً مشخص است که روش پیشنهادی لبه‌های تیزتری تولید می‌کند و اثرات مصنوعی کمتری دارد. در جدول ۵-۲ و جدول ۵-۳ روش پیشنهادی با استفاده از معیار RMSE و SSIM با روش‌های نامبرده مقایسه شده است. همان‌طور که در این دو جدول دیده می‌شود نسبت به روش‌های نام‌برده روش پیشنهادی نتایج رضایت بخشی دارد. با توجه به جدول ۵-۴ ازنقطه‌نظر زمان نیز نتایج رضایت بخشی دارد. همان‌طور که دیده می‌شود از روش‌های موجود در [۶۶] و [۶۸]، روش پیشنهادی سریع‌تر است. درنتیجه روش پیشنهادی می‌تواند در بسیاری از زمینه‌ها کاربرد داشته باشد. یکی از مزیت‌های روش پیشنهادی این است که به‌راحتی قابل پیاده‌سازی است و پیچیدگی محاسباتی ندارد درحالی‌که در [۶۶] و [۶۸]، پیچیدگی محاسباتی بسیار بالاتر است. همه‌ی روش‌ها بر روی یک سیستم با مشخصات 3GHZ پردازنده و 4GB حافظه اجراشده است.

جدول ۵-۱. نتایج PSNR برای ۱۸ تصویر متفاوت.تصویر ورودی ۱۲۸×۱۲۸.

	Bicubic	Chen[66]	Zhu[68]	NEDI[76]	روش پیشنهادی	بیشترین
barbara	۲۹,۳۷	۳۰,۶۲	۳۰,۴۲	۲۶,۸۶	**۳۰,۷۱**	**۳۰,۷۱**
man	۲۹,۶۸	۳۱,۱۷	۳۰,۸۵	۲۷,۰۴	**۳۱,۳۱**	**۳۱,۳۱**
boat	۲۸,۲۶	۲۹,۴۸	۲۹,۳۰	۲۶,۰۳	**۲۹,۶۲**	**۲۹,۶۲**
mandrill	۲۵,۷۸	۲۶,۳۳	۲۶,۳۸	۲۴,۴۸	**۲۶,۴۰**	**۲۶,۴۰**
lena	۳۰,۹۴	۳۲,۸۷	۳۲,۴۱	۲۷,۷۵	**۳۲,۹۷**	**۳۲,۹۷**
peppers	۳۱,۱۶	۳۲,۵۷	۳۲,۴۶	۲۷,۶۳	**۳۳,۲۸**	**۳۳,۲۸**
barbara2	۲۸,۷۵	۲۹,۳۷	۲۹,۰۱	۲۶,۴۲	**۲۹,۵۷**	**۲۹,۵۷**
cameraman	۲۸,۸۳	۳۰,۷۱	۳۰,۲۸	۲۵,۹۴	**۳۰,۸۵**	**۳۰,۸۵**
clown	۲۹,۱۹	۲۹,۳۴	۲۹,۸۷	۲۵,۸۵	**۳۰,۱۹**	**۳۰,۱۹**
couple	۲۸,۲۲	۲۹,۳۳	۲۹,۲۵	۲۶,۲۲	**۲۹,۴۰**	**۲۹,۴۰**
crowd	۲۸,۳۸	۲۹,۹۸	۲۹,۹۱	۲۵,۲۷	**۳۰,۰۹**	**۳۰,۰۹**
goldhill	۳۰,۶۱	۳۱,۷۰	۳۱,۷۳	۲۸,۲۵	**۳۱,۹۷**	**۳۱,۹۷**
nature	۲۶,۶۶	۲۷,۴۵	۲۷,۴۹	۲۴,۵۴	**۲۷,۵۴**	**۲۷,۵۴**
tank	۳۰,۵۱	۳۱,۴۰	۳۰,۲۰	۲۸,۱۳	**۳۱,۵۵**	**۳۱,۵۵**
elaine	۳۴,۷۱	۳۶,۸۲	۳۶,۰۶	۳۰,۱۱	**۳۷,۴۴**	**۳۷,۴۴**
girl	۳۲,۵۸	۳۴,۹۲	۳۴,۴۳	۲۹,۲۸	**۳۴,۹۴**	**۳۴,۹۴**
brick	۲۶,۹۲	۲۸,۴۷	۲۷,۲۰	۲۶,۴۸	**۲۸,۷۱**	**۲۸,۷۱**
testpat	۲۳,۶۱	۲۵,۲۳	۲۵,۵۱	۲۱,۹۶	**۲۵,۶۸**	**۲۵,۶۸**
average	۲۹,۱۲	۳۰,۴۳	۳۰,۱۵	۲۶,۵۷	**۳۰,۹۷**	**۳۰,۹۷**

همان‌طور که در جدول فوق مشاهده می‌شود روش پیشنهادی در همه‌ی تصاویر ازلحاظ PSNR نسبت به سایر روش‌ها بالاتر است. قابل مشاهده است که روش پیشنهادی به ترتیب نسبت به روش [۶۶] و [۶۸] به‌طور متوسط ازنظر معیار PSNR ۰,۵۴ و ۰,۸۲ بیشتر می‌باشد.

جدول ۵-۲. نتایج RMSE برای ۱۸ تصویر متفاوت، تصویر ورودی ۱۲۸×۱۲۸.

	Bicubic	Chen[66]	Zhu[68]	NEDI[76]	روش پیشنهادی	کمترین
barbara	۸,۶۷	۷,۵	۷,۶۷	۱۱,۵۷	**۷,۴۲**	**۷,۴۲**
man	۸,۳۵	۷,۰۴	۷,۳۰	۱۱,۳۲	**۶,۹۲**	**۶,۹۲**
boat	۹,۸۴	۸,۵۵	۸,۷۳	۱۲,۷۲	**۸,۴۲**	**۸,۴۲**
mandrill	۱۳,۱۰	۱۲,۳۰	۱۲,۲۲	۱۵,۲۱	**۱۲,۱۹**	**۱۲,۱۹**
lena	۷,۲۲	۵,۷۹	۶,۱۰	۱۰,۴۴	**۵,۷۲**	**۵,۷۲**
peppers	۷,۰۵	۵,۹۹	۶,۰۷	۱۰,۵۹	**۵,۵۲**	**۵,۵۲**
Barbara2	۹,۳۰	۸,۶۶	۹,۲۳	۱۲,۱۶	**۸,۴۶**	**۸,۴۶**
cameraman	۹,۲۱	۷,۴۳	۷,۸۰	۱۲,۸۵	**۷,۳۰**	**۷,۳۰**
clown	۸,۸۴	۸,۰۴	۸,۱۷	۱۳,۰۰	**۷,۸۸**	**۷,۸۸**
couple	۹,۸۹	۸,۷۰	۸,۷۸	۱۲,۴۵	**۸,۶۴**	**۸,۶۴**
crowd	۹,۷۱	۸,۰۷	۸,۱۴	۱۳,۸۸	**۷,۹۸**	**۷,۹۸**
goldhill	۷,۵۰	۶,۶۲	۶,۶۰	۹,۸۵	**۶,۴۲**	**۶,۴۲**
nature	۱۱,۸۳	۱۰,۸۰	۱۰,۷۶	۱۵,۱۰	**۱۰,۶۹**	**۱۰,۶۹**
tank	۷,۵۹	۶,۸۵	۷,۰۲	۹,۹۹	**۶,۷۴**	**۶,۷۴**
elaine	۴,۶۸	۳,۶۷	۴,۰۱	۷,۹۵	**۳,۴۲**	**۳,۴۲**
girl	۵,۹۹	۴,۵۷	۴,۸۳	۸,۷۵	**۴,۵۶**	**۴,۵۶**
brick	۱۱,۴۸	۹,۶۰	۱۰,۰۳	۱۲,۰۷	**۹,۳۴**	**۹,۳۴**
testpat	۱۶,۸۲	۱۳,۲۷	۱۳,۵۰	۲۰,۳۲	**۱۳,۲۵**	**۱۳,۲۵**
average	۹,۲۸	۷,۹۶	۸,۱۶	۱۲,۲۳	**۷,۵۰**	**۷,۵۰**

ازآنجایی‌که معیار RMSE خطا را محاسبه می‌کند بنابراین باید در جدول فوق به دنبال کمترین خطا باشیم. همان‌طور که دیده می‌شود از نقطه‌نظر RMSE نیز روش پیشنهادی نسبت به سایر روش‌های ذکرشده بهتر است.

جدول ۵-۳. نتایج SSIM برای ۱۸ تصویر متفاوت، تصویر ورودی ۱۲۸×۱۲۸.

	Bicubic	Chen[66]	Zhu[68]	NEDI[76]	روش پیشنهادی	بیشترین
barbara	۰,۸۷	**۰,۸۹**	۰,۸۸	۰,۸۲	**۰,۸۹**	**۰,۸۹**
man	۰,۸۸	**۰,۹۱**	۰,۹۰	۰,۸۱	**۰,۹۱**	**۰,۹۱**
boat	۰,۸۶	۰,۸۹	۰,۸۹	۰,۷۹	**۰,۹۰**	**۰,۹۰**
mandrill	۰,۷۴	**۰,۸۰**	۰,۸۰	۰,۶۶	**۰,۸۰**	**۰,۸۰**
lena	۰,۹۲	**۰,۹۵**	۰,۹۴	۰,۸۷	۰,۹۴	**۰,۹۵**
peppers	۰,۹۴	**۰,۹۶**	۰,۹۵	۰,۹۰	**۰,۹۶**	**۰,۹۶**
Barbara2	۰,۸۳	**۰,۸۶**	۰,۸۳	۰,۷۳	**۰,۸۶**	**۰,۸۶**
cameraman	۰,۹۲	**۰,۹۴**	۰,۹۳	۰,۸۶	**۰,۹۴**	**۰,۹۴**
clown	۰,۹	**۰,۹۱**	۰,۹۰	۰,۸۳	۰,۹۰	**۰,۹۱**
couple	۰,۸۴	**۰,۸۸**	۰,۸۷	۰,۷۷	**۰,۸۸**	**۰,۸۸**
crowd	۰,۸۹	**۰,۹۳**	۰,۹۲	۰,۸۲	**۰,۹۳**	**۰,۹۳**
goldhill	۰,۸۷	**۰,۹۰**	۰,۹۰	۰,۸۰	**۰,۹۰**	**۰,۹۰**
nature	۰,۸۱	**۰,۸۵**	**۰,۸۵**	۰,۷۱	**۰,۸۵**	**۰,۸۵**
tank	۰,۸۵	**۰,۸۸**	۰,۸۵	۰,۷۷	**۰,۸۸**	**۰,۸۸**
elaine	۰,۹۴	**۰,۹۶**	۰,۹۶	۰,۹۱	**۰,۹۶**	**۰,۹۶**
girl	۰,۹۰	**۰,۹۴**	۰,۹۳	۰,۸۳	**۰,۹۴**	**۰,۹۴**
brick	۰,۷۰	۰,۸۰	۰,۸۳	۰,۷۵	**۰,۸۴**	**۰,۸۴**
testpat	۰,۹۰	۰,۹۰	۰,۹۱	۰,۸۶	**۰,۹۲**	**۰,۹۲**
average	۰,۸۶	**۰,۸۹**	**۰,۸۹**	۰,۸۰	**۰,۸۹**	**۰,۸۹**

همان‌طور که در جدول فوق دیده می‌شود از لحاظ معیار SSIM روش پیشنهادی بسیار نزدیک به روش ارائه شده در [۶۶] است. اما در بعضی از تصاویر روش پیشنهادی مقادیر بزرگتری دارد.

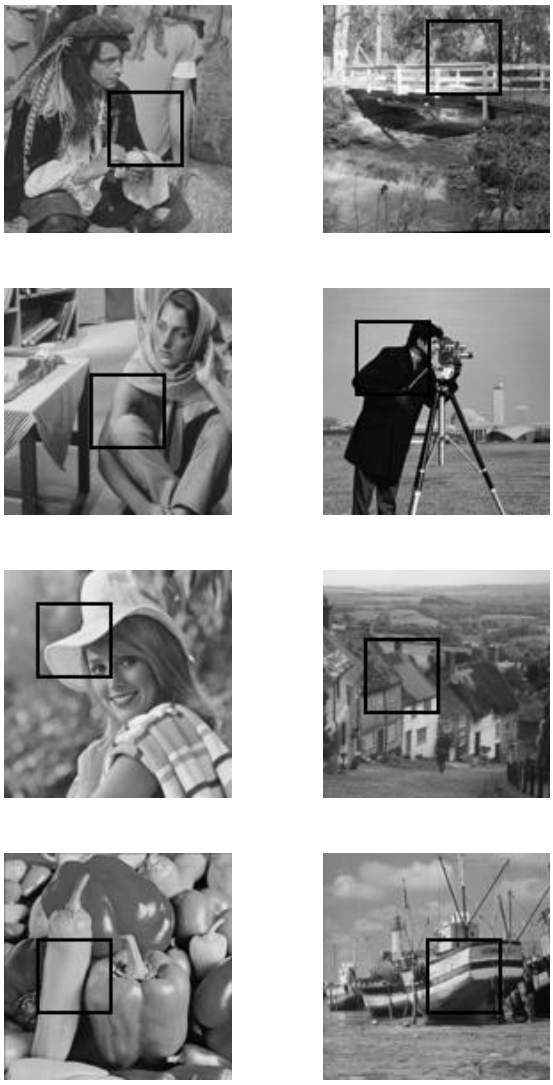
جدول ۵-۴. نتایج زمان برای ۱۸ تصویر.تصویر ورودی ۱۲۸×۱۲۸.

	Chen[66]	Zhu[68]	روش پیشنهادی	کمترین
barbara	۱۷۴,۲	**۳,۰۱**	۵,۱۹	**۳,۰۱**
man	۱۷۳,۹۳	**۲,۷۸**	۵,۴۳	**۲,۷۸**
boat	۱۷۴,۰۷	**۲,۴۴**	۵,۲۲	**۲,۴۴**
mandrill	۱۷۳,۹۴	**۲,۴۳**	۵,۲۶	**۲,۴۳**
lena	۱۷۳,۵۵	**۳,۹۲**	۵,۱۸	**۳,۹۲**
peppers	۱۷۳,۵	**۴,۵۷**	۵,۲۰	**۴,۵۷**
Barbara2	۱۷۴,۰۶	۶,۰۲	**۵,۲۲**	**۵,۲۲**
cameraman	۱۷۳,۶۳	۷,۳۹	**۵,۲۱**	**۵,۲۱**
clown	۱۸۳,۴۹	۷,۶۵	**۵,۲۱**	**۵,۲۱**
couple	۱۷۳,۵۱	۶,۳۰	**۵,۲۶**	**۵,۲۶**
crowd	۱۷۳,۸۵	۷,۶۳	**۵,۲۱**	**۵,۲۱**
goldhill	۱۷۳,۹۱	**۲,۴۸**	۵,۲۱	**۲,۴۸**
nature	۱۷۳,۷۷	**۵,۰۹**	۵,۲۴	**۵,۰۹**
tank	۱۷۳,۸۳	۵,۲۴	**۵,۲۱**	**۵,۲۱**
elaine	۱۷۳,۷۹	**۲,۵۱**	۵,۲۷	**۲,۵۱**
girl	۱۷۴,۱۷	**۳,۰۲**	۵,۲۳	**۳,۰۲**
brick	۱۷۳,۷۴	۶,۶۷	**۵,۲۳**	**۵,۲۳**
testpart	۱۹۳,۴۷	۱۵,۶۹	**۵,۲۷**	**۵,۲۷**
average	۱۷۵,۴۶	۵,۲۶	**۵,۲۳**	**۵,۲۳**

همان‌طور که در جدول فوق مشاهده می‌شود روش پیشنهادی نسبت به روش ارائه شده در [۶۶] بسیار سریع‌تر می‌باشد. اما به طور متوسط از لحاظ زمانی به روش ارائه شده در [۶۸] بسیار نزدیک است.

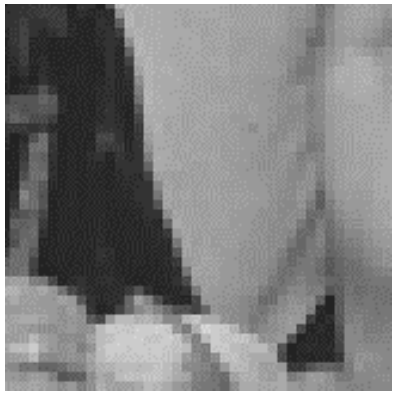
درشکل ۵-۱، ۸ تصویر وضوح پایین مشاهده می‌شود. می‌خواهیم الگوریتم پیشنهادی خود را بر روی این تصاویر پیاده کنیم. قسمتی از این تصاویر با مربع مشکی بریده‌شده است تا بهتر اثرات بزرگ‌نمایی روش‌های مطرح‌شده و همچنین روش پیشنهادی را نشان دهد. همان‌طور که درشکل ۵-۲ تا شکل ۵-۹ دیده می‌شود روش پیشنهادی نسبت به روشbicubic ، و همچنین [۶۶] و [۶۸]، لبه‌های تیزتری دارد. همچنین اثرات بلوکی کمتری تولید می‌شود. همان‌طور که ملاحظه می‌شود در شکل ۵-۲ لبه‌های لباس در روش پیشنهادی تیزتر است درحالی‌که قسمت (د) و (ه) همین شکل دارای اثرات بلوکی است. همچنین درشکل ۵-۴ لبه‌های مربوط به لباس در قسمت (د) و (و) دارای اثرات بلوکی و دندانه‌ای است

درحالی‌که لبه‌های تولیدشده در روش پیشنهادی اعوجاج کمتری دارد.



شکل ۵-۱. ۸ تصویر وضوح پایین برای مقایسه بصری.

الف



ب

ج

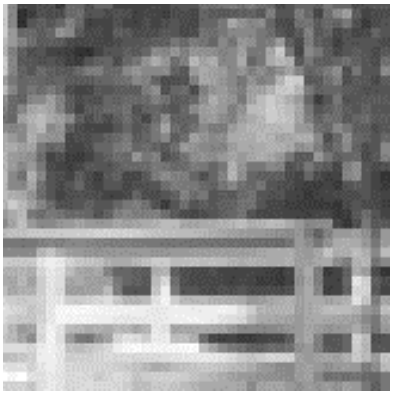
د

ه

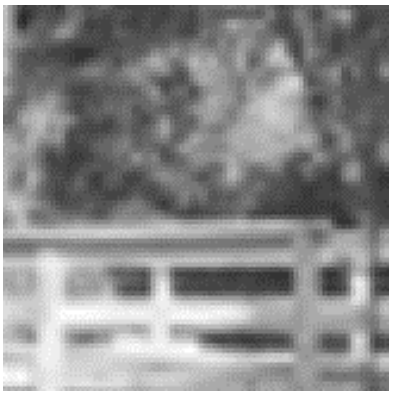
و

شکل ۵-۲. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف

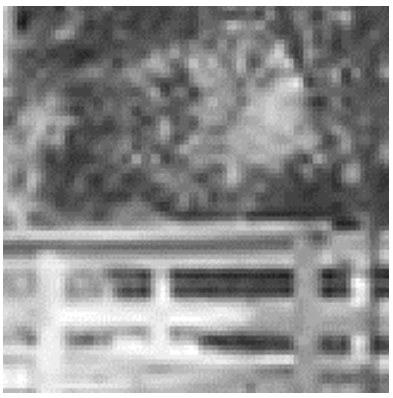


ب



ج

د

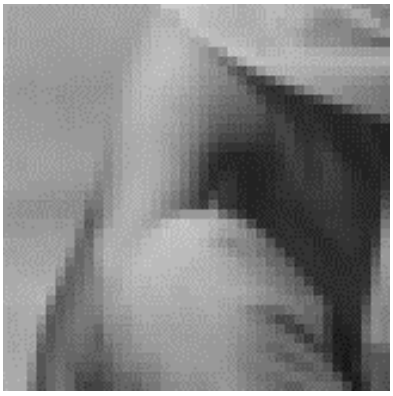


ه

و

شکل ۵-۳. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف



ب

ج

د

ه

و

شکل ۵-۴. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف

ب

ج

د

ه

و

شکل ۵-۵. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف

ب

ج

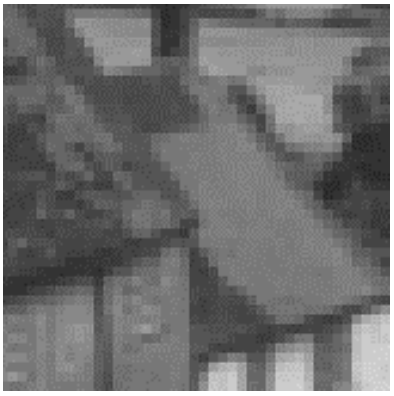
د

ه

و

شکل ۵-۶. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف



ب

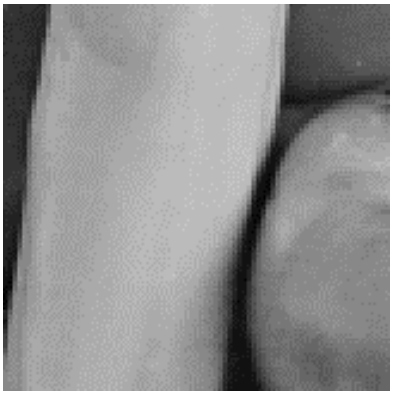
ج

د

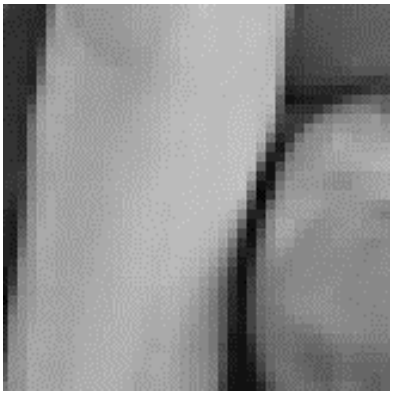
ه

و

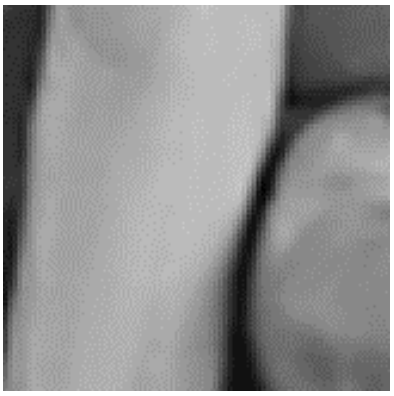
شکل ۵-۷. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.



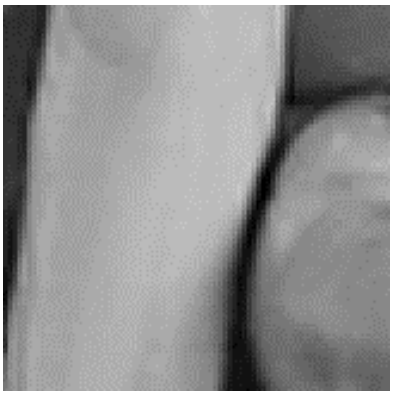
الف



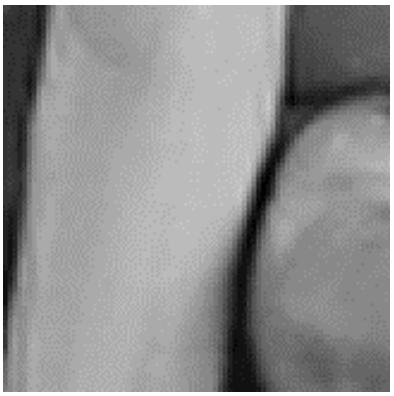
ب



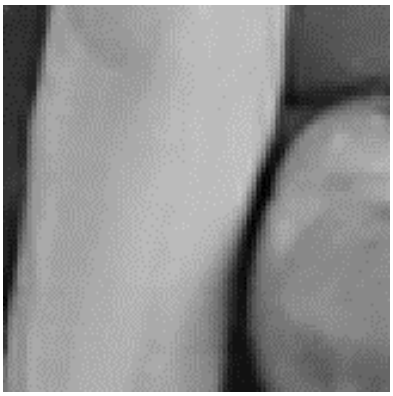
ج



د



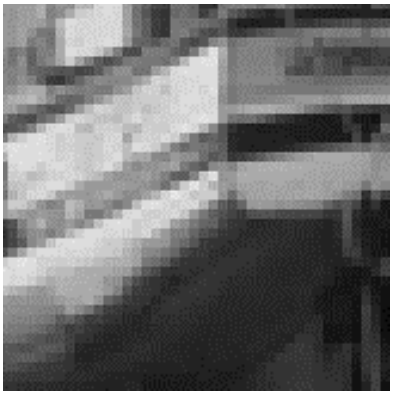
ه



و

شکل ۵-۸. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر [۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

الف



ب

ج

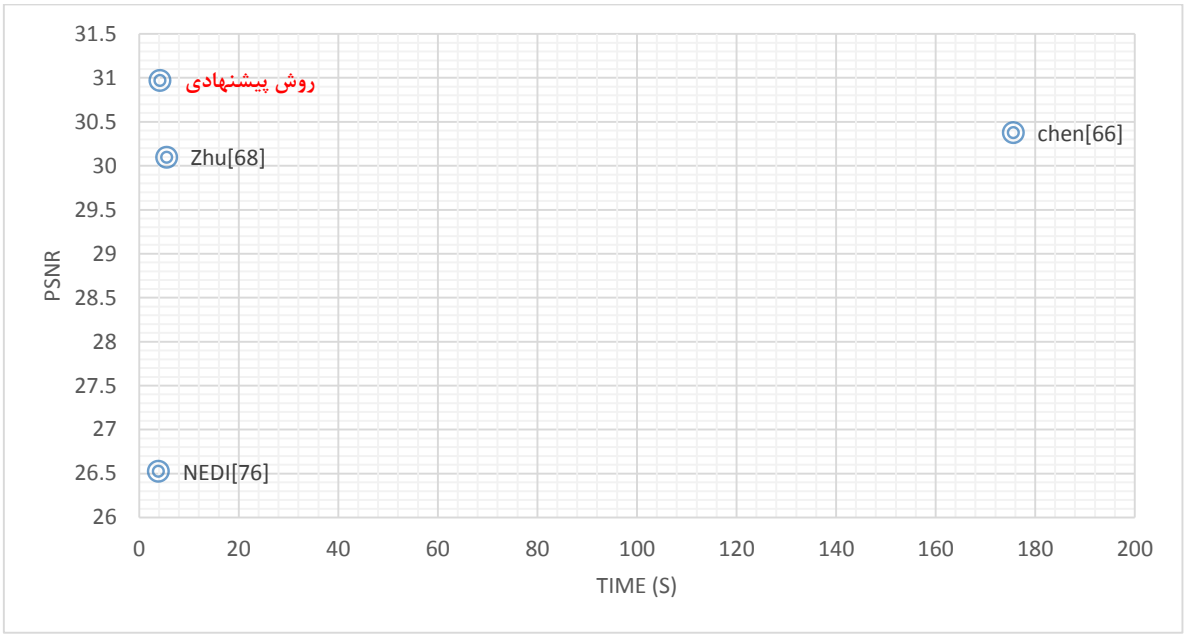
د

ه

و

شکل ۵-۹. (الف) تصویر اصلی، (ب) تصویر وضوح پایین، (ج) تصویر بزرگ‌شده توسط درون‌یابی bicubic، (د) روش مبتنی بر[۶۶]، (ه) روش مبتنی بر [۶۸]، (و) روش پیشنهای.

شکل ۵-۱۰ نتیجه‌ی مقایسه روش پیشنهادی با سایر روش‌های مطرح‌شده در این پایان‌نامه را با استفاده از PSNR و زمان نشان می‌دهد. همان‌طور که در شکل دیده می‌شود روش پیشنهادی ازنقطه‌نظر PSNR از روش‌های مطرح‌شده بالاتر است. همچنین ازنقطه‌نظر زمان نسبت به [۶۶] بسیار سریع‌تر است اما به روش ارائه‌شده در [۶۸] بسیار نزدیک می‌باشد.



شکل ۵-۱۰. مقایسه روش‌های SR با روش پیشنهادی از نقطه نظر زمان و PSNR.

۶ فصل ششم:جمع‌بندی و نتیجه‌گیری

۶-۱ نتیجه‌گیری:

در این پایان‌نامه روشی برای افزایش وضوح تصاویر خاکستری ارائه شد. رهیافت ارائه‌شده یک روش مبتنی بر خود یادگیری است. در این روش بدون نیاز به پایگاه داده خارجی و با استفاده از هرم ساخته‌شده توسط تصویر ورودی، نسخه‌ی SR تصویر ورودی ساخته می‌شود. از شباهت میان وصله‌های وضوح‌بالا و وضوح‌پایین برای یادگیری رابطه‌ی بین تصاویر استفاده شده است. در فصل سوم مباحث تئوری مربوط به فرا تفکیک‌پذیری تک فریمی مطرح گردید. در فصل چهارم ابتدا روشی برای استخراج وصله برای تولید دیکشنری و همچنین ایده بلاک بندی مطرح شد که سبب پایین آمدن پیچیدگی محاسباتی روش پیشنهادی گردید. در روش پیشنهادی هر وصله بر اساس هشت بلاک همسایه اطراف آن ساخته شد. ایده‌ی این روش، از این قضیه نشات می‌گیرد که هر وصله بر اساس وصله‌هایی که به آن نزدیک‌تر هستند ساخته شود. درنتیجه خطای بازسازی کمتر می‌شود. در این روش مستقیم از خود وصله‌ها استفاده شد. ازآنجایی‌که روشنایی یک تصویر تحت تأثیر عوامل زیادی قرار می‌گیرد روش پیشنهادی ازلحاظ معیار PSNR نتایج خوبی نداشت. در نتیجه از ۴ فیلتر بالاگذر برای استخراج ویژگی استفاده شد که دقت ساخت هر وصله را بالاتر برد. بنابراین دیکشنری ساخته‌شده نسبت به دیکشنری مبتنی بر وصله بسیار کارآمدتر است. همچنین برای از بین بردن اثرات بلوکی تولیدشده در تصویر نهایی از BTV و الگوریتم back projection استفاده شد. در فصل پنجم نتایج تجربی بازسازی مطلوب تصویر توسط روش پیشنهادی در مقایسه با سایر روش‌های مطرح شده بررسی گردید. نتایج تجربی نشان می‌دهد که روش پیشنهادی از نقطه‌نظر سرعت و PSNR به نتایج رضایت بخشی رسیده است.

۶-۲ 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.

- تاکنون کارهای متفاوتی در زمینه‌ی انتخاب جمله‌ی تنظیم‌کننده انجام شده است. یکی از کارها این است که از جمله‌ی تنظیم‌کننده مناسب‌تری در مسئله مینیمم‌سازی استفاده شود. درواقع به جمله‌ی تنظیم‌کننده‌ای نیاز است که لبه‌های تیزتری تولید کند و درعین‌حال اثرات بلوکی را نیز کاهش دهد.
- برای ساخت دیکشنری از ویژگی‌های دیگر استفاده شود. همچنین می‌توان برای افزایش سرعت، ابعاد ویژگی‌های استخراج‌شده را با استفاده از PCA کاهش داد.
- یکی از راهکارها برای افزایش PSNR این است که هم‌زمان از وصله‌های خود تصویر و همچنین یک پایگاه داده خارجی استفاده کرد.
- در حوزه فرا تفکیک‌پذیری تاکنون مقالات زیادی منتشرشده است. بااین‌حال 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.

مراجع:

[١] R. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in computer vision and Image Processing,* vol. 1, pp. 317-339, 1984.

[٢] S. Chaudhuri and M. V. Joshi, "Research on image super-resolution," *Motion-Free Super-Resolution* ,pp. 15-31, 2005.

[٣] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE,* vol. 20, pp. 21-36, 2003.

[٤] V. Patanavijit, "Super-resolution reconstruction and its future research direction," *AU J,* vol. 12, pp. 149-163, 2009.

[٥] 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. 471-474.

[٦] K. Nguyen, S. Sridharan, S. Denman, and C. Fookes, "Feature-domain super-resolution framework for Gabor-based face and iris recognition," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012 ,pp. 2642-2649.

[٧] 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.

[٨] H. Nyquist, "Certain topics in telegraph transmission theory," *American Institute of Electrical Engineers, Transactions of the,* vol. 47, pp. 617-644, 1928.

[٩] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review,* vol. 5, pp. 3-55, 2001.

[١٠] R. J. Schalkoff, *Digital image processing and computer vision* vol. 286: Wiley New York, 1989.

[١١] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *JOSA A,* vol. 6, pp. 1715-1726, 1989.

[١٢] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Computer Vision, 2009 IEEE 12th International Conference on*, 2009, pp. 349-356.

[١٣] V. R. Algazi, G. E. Ford, and R. Potharlanka, "Directional interpolation of images based on visual properties and rank order filtering," in *Acoustics, Speech, and Signal*

Processing, 1991. ICASSP-91., 1991 International Conference on, 1991, pp. 3005-3008.

[۱۴] W. Kwok and H. Sun, "Multi-directional interpolation for spatial error concealment," *Consumer Electronics, IEEE Transactions on,* vol. 39, pp. 455-460, 1993.

[۱۵] P. Milanfar, *Super-resolution imaging*: CRC Press, 2010.

[۱۶] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *Image Processing, IEEE Transactions on,* vol. 18, pp. 36-51, 2009.

[۱۷] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *Image processing, IEEE Transactions on,* vol. 13, pp. 1327-1344, 2004.

[۱۸] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *Image Processing, IEEE Transactions on,* vol. 16, pp. 349-366, 2007.

[۱۹] K. Kim and Y. Kwon, "Example-based learning for singleimage SR and JPEG artifact removal," *MPI-TR,(173),* vol. 8, 2008.

[۲۰] Z. Lin and H.-Y. Shum, "Fundamental limits of reconstruction-based superresolution algorithms under local translation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 26, pp. 83-97, 2004.

[۲۱] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *Computer Graphics and Applications, IEEE,* vol. 22, pp. 56-65, 2002.

[۲۲] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on,* vol. 19, pp. 2861-2873, 2010.

[۲۳] M. Cristani, D. S. Cheng, V. Murino, and D. Pannullo, "Distilling information with super-resolution for video surveillance," in *Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks*, 2004, pp. 2-11.

[۲۴] F. C. Lin, C. B. Fookes, V. Chandran, and S. Sridharan, "Investigation into optical flow super-resolution for surveillance applications," 2005.

[۲۵] F .Li, X. Jia, and D. Fraser, "Universal HMT based super resolution for remote sensing images," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 333-336.

[۲۶] J. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical*

image analysis, vol. 2, pp. 1-36, 1998.

[۲۷] T. Bauer, "Super-resolution imaging: The use case of optical astronomy," in *Proceedings of the IADIS International Conference Computer Graphics, Visualization, Computer Vision and Image Processing, CGVCVIP 2011, Edited by Yingcai Xiao, Rome, Italy, 2011, p. 49-59, ISBN 978-972-8939-48-9*, 2011, pp. 49-59.

[۲۸] K. Puschmann and F. Kneer, "On super-resolution in astronomical imaging," *Astronomy & Astrophysics*, vol. 436, pp. 373-378, 2005.

[۲۹] L. M. Novak, G. J. Owirka, and A. L. Weaver, "Automatic target recognition using enhanced resolution SAR data," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 35, pp. 157-175, 1999.

[۳۰] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *Image Processing, IEEE Transactions on*, vol. 20, pp. 1838-1857, 2011.

[۳۱] J. Zhang, C. Zhao, R. Xiong, S. Ma, and D. Zhao, "Image super-resolution via dual-dictionary learning and sparse representation," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, 2012, pp. 1688-1691.

[۳۲] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *Signal Processing, IEEE Transactions on*, vol. 54, pp. 4311-4322, 2006.

[۳۳] S. Kim, N. K. Bose, and H. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, pp. 1013-1027, 1990.

[۳۴] S. P. Kim and W.-Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *Image Processing, IEEE Transactions on*, vol. 2, pp .534–539 , 1993.

[۳۵] N. B. Karayiannis and A. N. Venetsanopoulos, "Image interpolation based on variational principles," *Signal Processing*, vol. 25, pp. 259-288, 1991.

[۳۶] K. Xue, A. Winans, and E. Walowit, "An edge-restricted spatial interpolation algorithm," *Journal of Electronic Imaging*, vol. 1, pp. 152-161, 1992.

[۳۷] R. R. Schultz and R. L. Stevenson, "A Bayesian approach to image expansion for improved definition," *Image Processing, IEEE Transactions on*, vol. 3, pp. 233-242, 1994.

[۳۸] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *Image Processing, IEEE Transactions on*, vol. 5, pp. 996-1011, 1996.

[۳۹] E. Mjolsness, "Fingerprint Hallucination," California Institute of Technology, 1985.

[۴۰] D. Kong ,M. Han, W. Xu, H. Tao, and Y. Gong, "A conditional random field model for video super-resolution," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 2006, pp. 619-622.

[۴۱] S. Baker and T. Kanade, "Hallucinating faces," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 83-88.

[۴۲] S. Baker and T. Kanade, "Super-resolution: Reconstruction or recognition," in *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing* ,2001.

[۴۳] C. Su, Y. Zhuang, L. Huang, and F. Wu, "Steerable pyramid-based face hallucination," *Pattern Recognition,* vol. 38, pp. 813-824, 2005.

[۴۴] S. Baker and T. Kanade, *Super-resolution optical flow*: Carnegie Mellon University, The Robotics Institute, 1999.

[۴۵] W. T. Freeman and E. C. Pasztor, "Learning to estimate scenes from images," *Advances in Neural Information Processing Systems,* pp. 775-781, 1999.

[۴۶] W. T. Freeman and E. Pasztor, "Markov networks for low-level vision," *Mitsubishi Electric Research Laboratory Technical, Report TR99-08,* 1999.

[۴۷] H. Yan, J. Liu, J. Sun, and X. Sun, "ICA based super-resolution face hallucination and recognition," in *Advances in Neural Networks–ISNN 2007,* ed: Springer, 2007, pp. 1065-1071.

[۴۸] J. Liu, J. Qiao, X. Wang, and Y. Li, "Face hallucination based on independent component analysis," in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, 2008, pp. 3242-3245.

[۴۹] Y. Liang, J.-H. Lai, X. Xie, and W. Liu, "Face hallucination under an image decomposition perspective," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 2158-2161.

[۵۰] S. Zhang and Y. Lu, "Image resolution enhancement using a Hopfield neural network," in *Information Technology, 2007. ITNG ۰۷'Fourth International Conference on*, 2007, pp. 224-228.

[۵۱] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Super-resolution target

identification from remotely sensed images using a Hopfield neural network," *Geoscience and Remote Sensing ,IEEE Transactions on,* vol. 39, pp. 781-796, 2001.

[۵۲] C. Miravet and F. B. Rodríguez, "A hybrid MLP-PNN architecture for fast image superresolution," in *Artificial Neural Networks and Neural Information Processing—ICANN/ICONIP 2003*, ed: Springer, 2003 ,pp. 417-424.

[۵۳] G. Pan, S. Han, and Z. Wu, "Hallucinating 3D facial shapes," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1-8.

[۵۴] X. Zeng and H. Huang, "Super-resolution method for multiview face recognition from a single image per person using nonlinear mappings on coherent features," *Signal Processing Letters, IEEE,* vol. 19, pp. 195-198, 2012.

[۵۵] C. Su and L. Huang, "Facial expression hallucination," in *Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops on*, 2005, pp. 93-98.

[۵۶] K. Su, Q. Tian, Q. Xue, N. Sebe, and J. Ma, "Neighborhood issue in single-frame image super-resolution," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, 2005, p. 4 pp.

[۵۷] 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.

[۵۸] B. V. Kumar and R. Aravind, "Face hallucination using OLPP and kernel ridge regression," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 353-356.

[۵۹] B. Li, H. Chang, S. Shan, and X. Chen, "Aligning coupled manifolds for face hallucination," *Signal Processing Letters, IEEE,* vol. 16, pp. 957-960, 2009.

[۶۰] B. Li, H. Chang, S. Shan, and X. Chen, "Locality preserving constraints for super-resolution with neighbor embedding," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, 2009, pp. 1189-1192.

[۶۱] J. Yang, H. Tang, Y. Ma, and T. Huang, "Face hallucination via sparse coding," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 1264-1267.

[۶۲] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 2008, pp. 1-8.

[۶۳] C.-Y. Tsai, D.-A. Huang, M.-C. Yang, L.-W. Kang, and Y.-C. F. Wang, "Context-

aware single image super-resolution using locality-constrained group sparse representation," in *Visual Communications and Image Processing (VCIP), 2012 IEEE*, 2012, pp. 1-6.

[۶۴] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 603-619, 2002.

[۶۵] Y.-W. Chao, Y.-R. Yeh, Y.-W. Chen, Y.-J. Lee, and Y.-C. F. Wang, "Locality-constrained group sparse representation for robust face recognition," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, 2011, pp. 761-764.

[۶۶] C. Chen and J. E. Fowler, "Single-image super-resolution using multihypothesis prediction," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 608-612.

[۶۷] A. N. Tikhonov and V. I. A. k. Arsenin, *Solutions of ill-posed problems*: Vh Winston, 1977.

[۶۸] Z. Zhu, F. Guo, H. Yu, and C. Chen, "Fast Single Image Super-Resolution via Self-Example Learning and Sparse Representation," *Multimedia, IEEE Transactions on*, vol. 16, pp. 2178-2190, 2014.

[۶۹] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, ed: Springer, 2012, pp. 711-730.

[۷۰] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation Technion—Israel Inst. of Technology, 2005," *Tech. Ref*.

[۷۱] E. W. Tramel and J. E. Fowler, "Video compressed sensing with multihypothesis," in *Data Compression Conference (DCC), 2011*, 2011, pp. 193-202.

[۷۲] 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.

[۷۳] S. Farsiu, "A fast and robust framework for image fusion and enhancement," Citeseer, 2005.

[۷۴] R. Timofte, V. De, and L. Van Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Computer Vision (ICCV), 2013 IEEE International Conference on*, 2013, pp. 1920-1927.

[۷۵] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," *CS Technion*, vol. 40, pp. 1-15,

2008.

[۷۶] X. Li and M. T. Orchard, "New edge-directed interpolation," *Image Processing, IEEE Transactions on,* vol. 10, pp. 1521-1527, 2001.

Abstract:

Super-resolution is a promising technique of digital imaging which attempts to generate a raster image with a higher resolution than its source. The source may consist of one or more images of a scene. This thesis focuses on single-frame super-resolution i.e. the source is a single raster image. In interpolation methods increasing resolution is performed by upsampling the image. In these methods missing details are not reconstructed properly. In comparison to other image enhancement techniques, super-resolution image reconstruction technique not only improves the quality of under-sampled, low-resolution images by increasing their spatial resolution but also attempts to filter out distortions. Increasing resolution in image means increasing the number of pixels in it. To increase the resolution of the input image we propose a new approach based on self learning technique. In proposed method for learning relation between low/high resolution patches, we use image pyramid that can estimate high frequncy details of the image without need for external database. Using the proposed method in this thesis, user can overcome to problems imaging systems using moderate cost hardware. In other hand user can generate high resolution image without collecting external database that includes high and low resolution images. In this thesis different methods of super resolution will be explained. The proposed method is fast so that for generating high resolution image the average time is 5 second. Also in terms of PSNR a satisfactory result is obtained. Experimental results show that the proposed method can remove artificial effects and also produce sharp edges. The proposed method in terms of time, PSNR and SSIM criteria is better than the other methods intruduced in this thesis.

Key words: super-resolution, tikhonov-regularization, dictionary, imag pyramid, BTV regularization.

Shahrood University of Technology

Faculty of Electrical and Robotic Engineering

MSc Thesis in Electronic Engineering

Image super resolution by learning image details from levels of image pyramid

By

Azade Mokari

Supervisor:

Dr .Alireza Ahmadyfard

February 2016