

بسم الله الرحمن الرحيم



دانشگاه حکیم بسزوری

دانشکده ریاضی و علوم کامپیوتر

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته علوم تصمیم و مهندسی دانش

## فرا تفکیک پذیری با نمایش تُنگ

استاد راهنما

دکتر محمود امین طوسی

استاد مشاور

دکتر مهدی زعفرانیه

پژوهشگر:

آلاله عصاران دربان

زمستان ۱۳۹۵



جلسه دفاع از پایان نامه آقای /خانم **آلاله عصاران** در بان دانشجوی رشته علوم تصمیم و مهندسی دانش به شماره دانشجویی ۹۳۱۳۱۳۷۰۳۱ با عنوان:

### فرا تفکیک پذیری با نمایش تنگ

در مورخه ..... در دانشکده ریاضی و علوم کامپیوتر تشکیل و توسط هیات داوران مورد ارزشیابی قرار گرفت و نمره ..... برابر درجه ..... برای آن تعیین گردید .  
به این ترتیب از این تاریخ آقای / خانم آلاله عصاران در بان به عنوان کارشناس ارشد در رشته مذکور شناخته می شود .

نمره کسب شده	حداکثر نمره	موارد	موارد ارزشیابی
	۴	رعایت اصول نگارش انسجام در تنظیم بخشهای مختلف، کیفیت تصاویر، جداول و اشکال، تنظیم فهرست ها، منابع و ماخذ	۱- کیفیت نگارش
	۱۰	بررسی تاریخچه و سابقه تجربی و نظری موضوع انسجام منطقی در بخش های مختلف پایان نامه، ابتکار و نوآوری، اهمیت و ارزش علمی پایان نامه، استفاده از منابع معتبر و جدید، کیفیت تجزیه و تحلیل یافته ها و نتیجه گیری، روشن بودن روش کار، هدف ها و فرضیه های تحقیق، جدید بودن روش تحقیق	۲- کیفیت علمی
	۴	تسلط بر موضوع و بیان واضح و تفهیم آن، توانایی در پاسخگویی به سوالات مطرح شده در جلسه، رعایت زمان ارائه و روش ارائه	۳- کیفیت ارائه در جلسه دفاع
	۱	گزارش های دوره ای پیشرفت کار (حداقل ۴ مورد)	۴- ارزشیابی گزارشات
	۱	مقاله مستخرج از پایان نامه: این نمره به صورت زیر اختصاص می یابد (۱) چکیده کنفرانسی هر مورد ۰/۲۵ نمره تا سقف ۰/۵ نمره (۲) مقاله کامل در مجموع مقالات همایشهای معتبر یا مقاله در مجلات علمی-ترویجی معتبر پذیرفته شده یا چاپ شده هر مورد ۰/۵ نمره تا سقف ۱ نمره (۳) مقاله پذیرفته شده یا چاپ شده در مجلات علمی پژوهشی معتبر ۱ نمره (۴) مقاله ارسال شده به مجلات علمی پژوهشی معتبر هر مورد ۰/۲۵ نمره تا سقف ۰/۵ نمره (۵) دستگاه ساخته شده دارای گواهی ثبت اختراع یا به سفارش سازمان ها تا سقف ۱ نمره (۶) دستگاه ساخته شده کاربردی که به تایید رئیس دانشکده رسیده باشد تا سقف ۰/۵ نمره	۵- خروجی پایان نامه
جمع			

درجه معادل کسب شده: (از ۲۰ تا ۲۰ عالی)  از ۱۸ تا ۱۸/۹۹ بسیار خوب  از ۱۶ تا ۱۷/۹۹ خوب  از ۱۴ تا ۱۵/۹۹ قابل قبول  کمتر از ۱۴ غیر قابل قبول

### مشخصات هیات دوران

ردیف	نام و نام خانوادگی	سمت	مرتبۀ علمی	محل کار	امضا
۱	دکتر محمود امین طوسی	استاد راهنما	استادیار	دانشگاه حکیم سبزواری	
۲	دکتر مهدی زعفرانیه	استاد مشاور	استادیار	دانشگاه حکیم سبزواری	
۳	دکتر علی اکبر عارفی جمال	استاد داور	دانشیار	دانشگاه حکیم سبزواری	
۴	دکتر مهدی مهدی زاده	نماینده تحصیلات تکمیلی	استادیار	دانشگاه حکیم سبزواری	

امضا

رئیس دانشکده

امضا

مدیر گروه



## سوگند نامه دانش آموختگان دانشگاه حکیم سبزواری

به نام خداوند جان و خرد      کزین برتر اندیشه بر نگذرد

اینک که به خواست آفریدگار پاک، کوشش خویش و بهره گیری از دانش استادان و سرمایه‌های مادی و معنوی این مرز و بوم، توشه‌ای از دانش و خرد گردآورده‌ام، در پیشگاه خداوند بزرگ سوگند یاد می‌کنم که در به کارگیری دانش خویش، همواره بر راه راست و درست گام بردارم. خداوند بزرگ، شما شاهدان، دانشجویان و دیگر حاضران را به عنوان داورانی امین گواه می‌گیرم که از همه دانش و توان خود برای گسترش مرزهای دانش بهره‌گیرم و از هیچ کوششی برای تبدیل جهان به جایی بهتر برای زیستن، دریغ نورزم. پیمان می‌بندم که همواره کرامت انسانی را در نظر داشته باشم و هموعان خود را در هر زمان و مکان تا سر حد امکان یاری دهم. سوگند می‌خورم که در به کارگیری دانش خویش به کاری که باره و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مبادت دارد دست نیازم. همچنین در سایه اصول جهان شمول انسانی و اسلامی، پیمان می‌بندم از هیچ کوششی برای آبادانی و سرافرازی میهن و هم میهنانم فروگذاری نکنم و خداوند بزرگ را به یاری طلبم تا همواره در پیشگاه او و در برابر وجدان بیدار خویش و ملت سرافراز، بر این پیمان تا ابد استوار بمانم.

نام و نام خانوادگی:      آلاله عصاران دربان

تاریخ و امضا:

## تأییدی صحت و اصالت نتایج

باسمه تعالی

اینجانب آلاله عصاران دربان به شماره دانشجویی ۹۳۱۳۱۳۷۰۳۱ دانشجوی رشته علوم تصمیم و مهندسی دانش مقطع تحصیلی کارشناسی ارشد تأیید می‌نمایم که کلیه نتایج این پایان‌نامه حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. در صورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ... ) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض در خصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذی صلاح (اعم از اداری و قضایی) به عهده ی اینجانب خواهد بود و دانشگاه هیچ گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: آلاله عصاران دربان

تاریخ و امضا:

## مجوز بهره برداری از پایان نامه

بهره برداری از این پایان نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما به شرح زیر

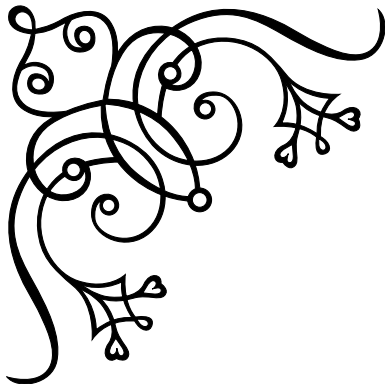
تعیین می شود، بلامانع است:

- بهره برداری از این پایان نامه برای همگان بلامانع است.
- بهره برداری از این پایان نامه با اخذ مجوز از استاد راهنما، بلامانع است.
- بهره برداری از این پایان نامه تا تاریخ ..... ممنوع است.

استاد راهنما: دکتر محمود امین طوسی

تاریخ و امضا:

تقدیم به:



پدر و مادر

و

همسرم



سپاس خداوندگار حکیم را که با لطف بی کران خود، آدمی را زیور عقل آراست. در آغاز وظیفه خود می دانم از زحمات بی دریغ استاد راهنمای خود، جناب آقای دکتر محمود امین طوسی، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی های ارزنده ایشان، این مجموعه به انجام نمی رسید. از جناب آقای دکتر مهدی زعفرانی که زحمت مطالعه و مشاوره این رساله را تقبل فرمودند و در آماده سازی این رساله، به نحو احسن اینجانب را مورد راهنمایی قرار دادند، کمال امتنان را دارم. همچنین لازم می دانم از گروه پارسی لاتک در پاسخگویی به مشکلات کاربران کمال قدردانی را داشته باشم. در پایان، بوسه می زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می کنم وجود مقدس شان را و تشکر می کنم از خانواده عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودشان، که بهترین پشتیبان من بودند.

آلاله عصاران دربان

زمستان ۱۳۹۵

# فهرست مطالب

د	فهرست جداول
ه	فهرست تصاویر
و	فهرست الگوریتم‌ها
۱	چکیده
۲	پیش‌گفتار
۳	فصل ۱: نمایش تنگ در پردازش تصویر
۳	۱-۱ مقدمه . . . . .
۵	۲-۱ فراتفکیک پذیری تصویر . . . . .
۶	۳-۱ ردیابی بصری . . . . .
۷	۱-۳-۱ ردیابی شی در ویدئو . . . . .
۸	۴-۱ طبقه‌بندی تصاویر . . . . .
۹	۱-۴-۱ چالش‌های اصلی در طبقه‌بندی تصاویر با استفاده از نمایش تنگ . . . . .
۱۰	۲-۴-۱ طبقه‌بندی تصاویر ابرطیفی . . . . .
۱۱	۳-۴-۱ تشخیص چهره . . . . .
۱۲	۵-۱ تخمین حالت انسان . . . . .
۱۳	۶-۱ تصاویر پزشکی . . . . .
۱۴	۷-۱ ترکیب تصاویر با هدف بهبود کیفیت . . . . .
۱۵	۸-۱ مقایسه روش‌های بهینه‌سازی تنگ در زمینه‌های کاربردی . . . . .
۱۷	فصل ۲: روش‌های بهینه‌سازی تنگ
۱۸	۱-۲ روش جست و جوی تطابقی . . . . .

۱۹	جست و جوی تطابقی متعامد	۲-۲
۲۰	روش KSVD	۳-۲
۲۲	روش LIQP	۴-۲
۲۲	الگوریتم KSVD تُنک	۵-۲
۲۳	الگوریتم بروز کردن موازی اتم‌ها	۶-۲
۲۴	نمایش تُنک مضاعف	۷-۲

### فصل ۳: فراتفکیک پذیری با نمایش تُنک

۲۷	مقدمه	۱-۳
۲۸	روش یانگ	۲-۳
۲۸	فراتفکیک پذیری تصاویر از منظر تُنک بودن	۱-۲-۳
۳۰	فراتفکیک پذیری برای تصاویر عمومی	۲-۲-۳
۳۱	القای محدودیت بازسازی سراسری	۳-۲-۳
۳۲	شرح بهینه‌سازی سراسری	۴-۲-۳
۳۳	فراتفکیک پذیری برای تصاویر صورت	۵-۲-۳
۳۳	آموزش لغت‌نامه برای فراتفکیک پذیری	۶-۲-۳
۳۵	آموزش لغت‌نامه	۷-۲-۳
۳۵	آموزش لغت‌نامه منفرد	۸-۲-۳
۳۷	آموزش جفت لغت‌نامه	۹-۲-۳
۳۸	نمایش ویژگی برای قطعه‌های با تفکیک پذیری پایین	۱۰-۲-۳
۳۸	تأثیر محدودیت سراسری	۱۱-۲-۳
۳۹	بررسی معیار شباهت ساختاری بجای میانگین مربعات خطا	۳-۳

### فصل ۴: نتایج آزمایشات

۴۳	بکارگیری فیلتر موجک بجای گرادیان	۱-۴
۴۴	مشکل روش فراتفکیک پذیری با نمایش تُنک	۲-۴
۴۴	تغییر پارامترها	۳-۴
۴۵	تأثیر میزان همپوشانی در افزایش تفکیک پذیری	۱-۳-۴
۴۵	اثر تعداد اتم‌های لغت‌نامه در افزایش تفکیک پذیری	۲-۳-۴
۴۷	اثر اندازه قطعه‌های تصویر در افزایش تفکیک پذیری	۳-۳-۴
۴۸	اثر تعداد نمونه‌ها برای ساخت لغت‌نامه	۴-۳-۴



۴۸	ساخت لغت‌نامه از روی تصاویر صورت	۴-۴
۵۰	ساخت لغت‌نامه بر اساس پایگاه داده اعداد mnistAll	۵-۴
۵۱	تغییر روش حل مسئله‌ی بهینه‌سازی تَنک	۶-۴
۵۱	جایگزینی روش OMP بجای روش L1QP	۱-۶-۴
۵۲	جایگزینی روش KSVD بجای L1QP	۲-۶-۴
۵۳	نتیجه‌گیری	۷-۴

۵۴ فهرست منابع

۵۷ پیوست آ: نمونه برنامه‌ی های نوشته شده

# فهرست جداول

- ۱-۱ مقایسه روش‌های بهینه‌سازی تُتک در زمینه‌های کاربردی . . . . . ۱۵
- ۱-۳ بررسی معیار SSIM بجای MSE در فراتفکیک پذیری [۲۵] . . . . . ۴۲
- ۱-۴ بکارگیری فیلتر موجک بجای فیلتر گرادیان افقی و عمودی در ساخت لغت نامه و استخراج ویژگی . . . . . ۴۴
- ۲-۴ مقایسه میزان همپوشانی بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۱۰۲۴ و اندازه قطعات ۵ و تعداد قطعات ۱۰۰۰) . . . . . ۴۵
- ۳-۴ مقایسه زمان اجرا برای تصویر لنا (تعداد اتم‌های لغت‌نامه ۱۰۲۴ و اندازه قطعات ۵ و تعداد قطعات ۱۰۰۰) . . . . . ۴۵
- ۴-۴ مقایسه تعداد اتم‌های لغت‌نامه بر روی تصویر لنا (اندازه قطعات ۵ و همپوشانی ۴ و تعداد قطعات ۱۰۰۰) . . . . . ۴۶
- ۵-۴ مقایسه زمان اجرا بر روی تصویر لنا (اندازه قطعات ۵ و همپوشانی ۴ و تعداد قطعات ۱۰۰۰) . . . . . ۴۶
- ۶-۴ مقایسه اندازه قطعات لغت‌نامه بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۲۵۶ و همپوشانی ۴ و تعداد قطعات ۱۰۰۰) . . . . . ۴۸
- ۷-۴ مقایسه زمان اجرا بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۲۵۶ و همپوشانی ۴ و تعداد قطعات ۱۰۰۰) . . . . . ۴۸
- ۸-۴ مقایسه تعداد نمونه قطعات بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۵۱۲ و همپوشانی ۴) . . . . . ۴۸
- ۹-۴ مقایسه زمان اجرا بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۵۱۲ و همپوشانی ۴) . . . . . ۴۹
- ۱۰-۴ مقایسه روش برای آموزش لغت‌نامه و عمل فراتفکیک پذیری بر روی تصویر لنا (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۱
- ۱۱-۴ مقایسه زمان اجرا بر روی تصویر لنا (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۲
- ۱۲-۴ مقایسه روش برای آموزش لغت‌نامه و عمل فراتفکیک پذیری بر روی تصویر لنا (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۲
- ۱۳-۴ مقایسه زمان اجرا بر روی تصویر لنا (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۳

# فهرست تصاویر

- ۱-۱ نمونه تصویر برای طبقه‌بندی . . . . . ۹
- ۲-۱ (آ) تصویر اصلی ، (ب) تصویر مرجع ، (پ) خروجی قطعه‌بندی با روش SVM و (ت) خروجی قطعه‌بندی با روش SDL . . . . . ۱۱
- ۳-۱ ورودی ها با فوکوس متفاوت . . . . . ۱۵
- ۴-۱ تصویر خروجی با کیفیت بالا ترکیب شده از تصاویر ورودی . . . . . ۱۵
- ۱-۴ از چپ به راست تعداد اتم‌ها ۲۰۴۸، ۱۰۲۴، ۵۱۲، ۲۵۶ برای لغت‌نامه . . . . . ۴۶
- ۲-۴ داده های ورودی برای لغت‌نامه . . . . . ۴۹
- ۳-۴ ورودی و خروجی فرافتفکیک پذیری با لغت‌نامه تصاویر صورت . . . . . ۴۹
- ۴-۴ داده های ورودی برای لغت‌نامه . . . . . ۵۰
- ۵-۴ ورودی و خروجی فرافتفکیک پذیری با لغت‌نامه تصاویر اعداد . . . . . ۵۰
- ۶-۴ ورودی و خروجی تفکیک پذیری با روش OMP (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۱
- ۷-۴ ورودی و خروجی تفکیک پذیری با روش KSVD (لغت‌نامه با اندازه ۶۴ اتم و اندازه قطعات ۵) . . . . . ۵۲

# فهرست الگوریتم‌ها

جست و جوی تطابقی با ورودی‌های ماتریس $A$ و بردار $b$ و حد آستانه $\epsilon$ . . . . .	۱۹	۱-۲
جست و جوی تطابقی متعامد با ورودی‌های ماتریس $A$ و بردار $b$ و حد آستانه $\epsilon$ . . . . .	۲۰	۲-۲
الگوریتم L1QP با ورودی‌های لغت‌نامه $M$ و بردار ویژگی $p$ . . . . .	۲۲	۳-۲
الگوریتم فراتفکیک پذیری برای تصاویر طبیعی . . . . .	۳۲	۱-۳
الگوریتم آموزش لغت‌نامه . . . . .	۳۶	۲-۳
الگوریتم معیار شباهت ساختاری مبتنی بر OMP . . . . .	۴۱	۳-۳
الگوریتم فراتفکیک پذیری مبتنی بر شباهت ساختاری . . . . .	۴۱	۴-۳



دانشگاه سبزوار

## فرم چکیده ی پایان نامه ی دوره ی تحصیلات تکمیلی

مدیریت تحصیلات تکمیلی

نام خانوادگی دانشجو: عصاران دربان	نام: آلاله	ش. دانشجویی: ۹۳۱۳۱۳۷۰۳۱
استاد راهنما: دکتر محمود امین طوسی		
استاد مشاور: دکتر مهدی زعفرانیه		
دانشکده ریاضی و علوم کامپیوتر	رشته: علوم تصمیم و مهندسی دانش	
مقطع: کارشناسی ارشد	تاریخ دفاع: زمستان ۱۳۹۵	تعداد صفحات: ۶۵
عنوان پایان نامه: فراتفکیک پذیری با نمایش تُنک		
کلید واژه ها: نمایش تُنک، آموزش لغت نامه، فراتفکیک پذیری، استخراج ویژگی		
<p>چکیده: پردازش تُنک سیگنال ها به عنوان ابزاری قدرتمند و جایگزینی کارا برای تبدیلات کلاسیک کامل طی دهه ی اخیر به شدت مورد توجه قرار گرفته است. بر پایه ی نمایش تُنک می خواهیم از بین تعداد زیادی سیگنال پایه، کمترین تعداد را برای نمایش یک سیگنال انتخاب کنیم. هر سیگنال پایه یک «اتم» و مجموعه ی این اتم ها یک «لغت نامه» نامیده می شود. در سال های اخیر با معرفی الگوریتم های عملی نشان داده شده است که تُنک ترین نمایش یک سیگنال در یک لغت نامه فوق کامل تحت شرایطی یکتا بوده، به این ترتیب این بحث به سرعت در کاربردهای گوناگون پردازش سیگنال به خصوص تصویر از جمله نویز زدایی تصویر، فراتفکیک پذیری، ردیابی بصری، ردیابی شی در ویدئو، طبقه بندی تصاویر، تخمین حالت انسان، تشخیص چهره، تصویر برداری پزشکی، ترکیب تصاویر با هدف بهبود کیفیت مورد استفاده قرار گرفت که در فصل اول بیان شده است. دو مسئله ی مهم در نمایش تُنک وجود دارد. یکی از این مسائل، پیدا کردن یک لغت نامه ی فوق کامل مناسب برای یک مجموعه مشخص از داده هاست، یعنی لغت نامه ای که بتواند برای همه ی سیگنال های آن مجموعه، یک نمایش به اندازه ی کافی تُنک ارائه دهد از جمله این الگوریتم ها می توان به K-SVD و MOD اشاره کرد. مسئله ی دوم داشتن یک الگوریتم کارا برای بدست آوردن تُنک ترین نمایش سیگنال (یا کدینگ تُنک) است از جمله این الگوریتم ها می توان به الگوریتم تعقیب تطابق (MP)، تعقیب تطابق متعامد (OMP) اشاره کرد. در این پایان نامه روشی که یانگ ارائه کرده است مبنا قرار داده شده است که در فصل سوم به تفصیل به آن پرداخته شده است. سپس به منظور یافتن نتیجه ی بهتر از دیگر روش ها برای استخراج ویژگی و ساخت لغت نامه بهره برده شده است و تاثیر تغییر پارامترها بررسی شده است.</p>		

## پیش‌گفتار

پردازش تُنک سیگنال‌ها به عنوان ابزاری قدرتمند و جایگزینی کارا برای تبدیلات کلاسیک کامل طی دهه‌ی اخیر به شدت مورد توجه قرار گرفته‌است. بر پایه‌ی نمایش تُنک می‌خواهیم از بین تعداد زیادی سیگنال پایه، که در حالت کلی تعدادشان خیلی بیشتر از بُعدشان است، کمترین تعداد را برای نمایش یک سیگنال انتخاب کنیم. هر سیگنال پایه یک «اتم» و مجموعه‌ی این اتم‌ها یک «لغت‌نامه» نامیده می‌شود. این عمل در حالت کلی دشوار بوده و جزء مسائل NP-hard است؛ چرا که نیازمند یک جستجوی ترکیبیاتی است. در سال‌های اخیر اما با ارائه‌ی پشتوانه‌های تنوریک و معرفی الگوریتم‌های عملی نشان داده شده‌است که تُنک‌ترین نمایش یک سیگنال در یک لغت‌نامه فوق کامل تحت شرایطی یکتا بوده و می‌توان این جواب را در زمان محدود بدست آورد. به این ترتیب این مبحث به سرعت در کاربردهای گوناگون پردازش سیگنال به خصوص تصویر از جمله نويز زدایی تصویر، فراتفکیک پذیری، ردیابی بصری، ردیابی شی در ویدئو، طبقه‌بندی تصاویر، تخمین حالت انسان، تشخیص چهره، تصویر برداری پزشکی، ترکیب تصاویر با هدف بهبود کیفیت مورد استفاده قرار گرفت که در فصل دوم بیان شده‌است. دو مسئله‌ی مهم در پردازش تُنک وجود دارد. یکی از این مسائل، پیدا کردن یک لغت‌نامه‌ی فوق کامل مناسب برای یک مجموعه مشخص از داده‌هاست، یعنی لغت‌نامه‌ای که بتواند برای همه‌ی سیگنال‌های آن مجموعه، یک نمایش به اندازه‌ی کافی تُنک ارائه دهد از جمله این الگوریتم‌ها می‌توان به MOD، K-SVD اشاره کرد. این موضوع منجر به توسعه‌ی الگوریتم‌های آموزش لغت‌نامه شده است. مسئله‌ی دوم داشتن یک الگوریتم کارا برای بدست آوردن تُنک‌ترین نمایش سیگنال (یا کدینگ تُنک سیگنال) است از جمله این الگوریتم‌ها می‌توان به الگوریتم تعقیب تطابق (MP)، تعقیب تطابق متعامد (OMP)، لاسو (Lasso) و فوکوس (FOCUSS) اشاره کرد. در روش‌های مبتنی بر نمایش تُنک به قطعات با تفکیک‌پذیری بالا و پایین نیاز داریم که به ما اطلاعاتی خارج از اطلاعات موجود در تصویر داده شده با تفکیک‌پذیری پایین می‌دهد. این قطعات برای آموزش دو لغت‌نامه که برای نمایش قطعات با تفکیک‌پذیری پایین و بالا بکار می‌روند استفاده می‌شود [۱]. در این پایان‌نامه روشی که یانگ ارائه کرده است مبنا قرار داده شده‌است که در فصل چهارم به تفصیل به آن پرداخته شده است. سپس به منظور یافتن نتیجه‌ی بهتر از دیگر روش‌ها برای استخراج ویژگی و ساخت لغت‌نامه بهره برده شده‌است و تاثیر تغییر پارامترها بررسی شده‌است.

# فصل ۱

## نمایش تَنک در پردازش تصویر

### ۱-۱ مقدمه

نمایش سیگنال‌ها در یک پایه‌ی <sup>۲</sup> مناسب همواره به عنوان یک گام اساسی در شناخت ویژگی‌ها و نیز استخراج و تفسیر اطلاعات سیگنال‌ها مورد توجه بوده است [۲]. تبدیل‌های کلاسیک متعامد یکه <sup>۳</sup> به دلیل سادگی محاسبات و بعلاوه، داشتن جوابی یکتا برای نمایش سیگنال‌ها، در جامعه پردازش سیگنال مورد توجه قرار داشته و دارد. از این دست تبدیل‌ها می‌توان به تبدیل فوریه، تبدیل کسینوسی گسسته و تبدیل موجک اشاره کرد. این دسته از تبدیل‌ها اصطلاحاً «کامل» هستند، یعنی تعداد سیگنال‌های پایه با بعد سیگنال برابر است. عیبی که این تبدیل‌ها دارند این است که فقط برای نمایش دسته‌ی محدودی از سیگنال‌ها مناسب هستند. یک تبدیل یا و یا نمایش «مناسب» این ویژگی را دارد که برای ساختن یا بسط سیگنال با یک خطای مشخص، تنها از تعداد کمی سیگنال پایه استفاده می‌شود. این همان مفهوم «سادگی نمایش» سیگنال‌ها است. چرا که ما همواره به دنبال ساده‌ترین نمایش برای سیگنال تحت بررسی خود هستیم که با کمترین تعداد سیگنال پایه توصیف شود. آنچه در پردازش تَنک اهمیت بسیاری دارد و در کارایی آن در کاربردهای مختلف نیز مستقیماً تأثیر می‌گذارد انتخاب مناسب سیگنال‌های پایه «اتم‌ها» از یک سو و به دست آوردن نمایش مناسب برای سیگنال با استفاده از مجموعه این اتم‌ها یا «لغت‌نامه» است. مورد اول «یادگیری لغت‌نامه» <sup>۴</sup> و مورد دوم را «کدینگ تَنک» <sup>۵</sup> می‌نامند. در ادامه تعاریف مورد نیاز بیان شده است.

تعریف ۱-۱-۱ ( ماتریس رتبه کامل). یک ماتریس رتبه کامل <sup>۶</sup> نامیده می‌شود اگر تمام سطرها و ستون‌های آن مستقل خطی باشد.

تعریف ۱-۱-۲ (اتم). به بردارهای پایه در رویکرد نمایش تَنک، اتم <sup>۷</sup> می‌گویند.

تعریف ۱-۱-۳ (لغت‌نامه). به ماتریسی که ستون‌های شامل پایه‌ها یا اتم‌ها است، لغت‌نامه <sup>۸</sup> می‌گویند.

<sup>۲</sup>Basis    <sup>۳</sup>Orthogonal    <sup>۴</sup>Dictionary Learning    <sup>۵</sup>Sparse Coding    <sup>۶</sup>Full rank    <sup>۷</sup>Atom

<sup>۸</sup>Dictionary

تعریف ۱-۱-۴ (کدینگ تُتک). کدینگ تُتک<sup>۱</sup>، بردارهای داده را بصورت ترکیب خطی مینیمم تعداد از اتم‌های لغت‌نامه بیان می‌کند.

موفقیت نمایش تُتک از آنجا ناشی می‌شود که اکثر سیگنال‌های طبیعی مانند تصویر یا صدا با در نظر گرفتن پایه‌های مشخصی دارای نمایش تُتک هستند. سیگنال‌های طبیعی غالباً کل فضا را پوشش نمی‌دهند و روی یک زیرفضا قرار می‌گیرند. با فرض اینکه سیگنال می‌تواند مشخص‌کننده پیکسل‌های یک تصویر باشد. فرض تُتک بودن بیان می‌کند که این داده را می‌توان به شکل ترکیب خطی تعداد کمی از پایه‌ها که از قبل در نظر گرفته شده بیان کرد. تعداد پایه‌هایی که در نمایش  $x$  موثر هستند با  $k$  نشان داده می‌شود. این تعداد باید به شکل قابل توجهی از بعد فضای اصلی داده کوچکتر باشد. پایه‌هایی که برای نمایش داده‌ها استفاده می‌شود در یک ماتریس قرار داده می‌شود. این ماتریس لغت‌نامه نامیده می‌شود. عموماً این ماتریس کل فضای برداری مربوط به داده‌ها را پوشش می‌دهد. به عبارت دیگر ستون‌های این ماتریس برای ساخت کل یا قسمتی از داده‌هایی که از فضای برداری  $R^n$  هستند استفاده می‌شوند. هر کدام از ستون‌های ماتریس لغت‌نامه، که پایه‌های مورد نیاز برای مدل کردن داده هستند یک اتم نامیده می‌شود. اگر تعداد اتم‌های لغت‌نامه به اندازه بعد فضای برداری باشد و پایه‌ها کل فضا را پوشش دهند ماتریس لغت‌نامه کامل نامیده می‌شود. در این حالت هر داده یک نمایش یکتا توسط اتم‌های لغت‌نامه خواهد داشت. اگر ماتریس لغت‌نامه کامل باشد نمایش هر داده توسط این ماتریس یکتاست و لزوماً تُتک نخواهد بود. اگر تعداد اتم‌های لغت‌نامه از حالت کامل بیشتر شود و تعدادی اتم به آن اضافه شود در این حالت لغت‌نامه را فراکامل می‌نامند. اگر  $D$  نشان دهنده لغت‌نامه باشد در عبارت  $x = D\alpha$  داده  $x$  توسط ترکیب خطی اتم‌ها نوشته شده است. چون ماتریس ما فراکامل است دستگاه معادلات برای تعیین  $\alpha$  نامعین خواهد بود. این دستگاه بی‌شمار پاسخ خواهد داشت که این پاسخ‌ها خود بر روی یک فضای برداری قرار دارند. از آنجایی که به دنبال تُتک‌ترین نمایش از بین نمایش‌های موجود هستیم شرط دیگری اضافه می‌کنیم. باید قیدی به مسئله اضافه شود که تعداد درایه‌های غیر صفر نمایش تُتک را محدود کند:

$$\arg \min_{\alpha} \|\alpha\|_0, \text{ s.t. } x = D\alpha \quad (1-1)$$

پاسخ عبارت بهینه‌سازی ۱-۲ نمایش تُتک سیگنال  $x$  را می‌دهد. در این عبارت  $\alpha$  برداری شامل نمایش تُتک سیگنال است. سوالات زیادی وجود دارد که باید پاسخ داده شود. اول اینکه آیا اصولاً تُتک‌ترین جواب این دستگاه معادلات یکتا است؟ چرا که اگر چندین جواب متمایز به عنوان تُتک‌ترین نمایش یا ساده‌ترین توصیف برای سیگنال مورد بررسی ما وجود داشته باشد، برای ما سودی ندارد. دوم اینکه، اگر تُتک‌ترین نمایش سیگنال یکتا باشد، با چه الگوریتمی این جواب را به دست آوریم؟ یا اصلاً الگوریتمی وجود دارد که در زمان محدود و با حجم محاسبات نه چندان بالا این جواب را پیدا کند؟ در پاسخ باید گفت که انبوه مقالاتی که در زمینه پردازش تُتک سیگنال‌ها طی دهه‌ی اخیر منتشر شده است نشان می‌دهد که اولاً تُتک‌ترین جواب تحت شرایطی یکتا بوده و ثانياً الگوریتم‌های بسیار کارایی برای پیدا کردن این جواب وجود دارد. اما اهمیت نمایش تُتک چیست؟ برای پاسخ به این سوال، ابتدا اهمیت داشتن یک «مدل» برای سیگنال‌ها بیان می‌شود. تقریباً تمامی کاربردهای پردازش سیگنال مبتنی بر اتخاذ یک مدل مناسب برای سیگنال‌ها است [۳، ۴]. مدل، در حقیقت مشخصات

<sup>۱</sup>Sparse Coding



سیگنال را توضیح می‌دهد و به نوعی هویت یک سیگنال است. به عنوان یک مثال نسبتاً ساده، اگر مسئله‌ی نوینزدایی از تصویر در نظر گرفته شود. مسئله به این قرار است که یک نسخه‌ی آغشته به نویز از یک سیگنال در اختیار است و هدف این است که سیگنال تمیز مطلوب بازیابی شود. این کار مستلزم این است که یک مدل برای سیگنال مطلوب فرض شود، این مدل می‌تواند به اینصورت باشد که سیگنال «هموار» است. به این ترتیب سیگنال تمیز از نویز که ناهموار است متمایز می‌شود. نمایش تُنک کاربردهای مختلفی دارد که چند نمونه از آن در ادامه معرفی خواهد شد.

## ۲-۱ فراتفکیک پذیری تصویر<sup>۱</sup>

استفاده از نمایش تُنک در سال‌های اخیر در حوزه‌ی پردازش سیگنال بسیار مورد توجه قرار گرفته است. در بسیاری از مسائل معکوس ریاضی نشان داده شده است که از فرض تُنک بودن می‌توان به عنوان یک معیار کارا برای حل استفاده کرد. یکی از مسائلی که اخیراً برای حل آن از نمایش‌های تُنک استفاده شده است، فراتفکیک پذیری تصاویر است. فراتفکیک پذیری به معنی یافتن یک تصویر با تفکیک پذیری (رزولوشن) بالا از روی یک یا چند تصویر با تفکیک پذیری پایین است. هدف مسئله به دست آوردن تصویر با تفکیک پذیری بالا از روی یک تصویر با تفکیک پذیری پایین است که در ادبیات به فراتفکیک پذیری بر مبنای نمونه معروف است. دو دسته روش کلی برای حل این مسئله وجود دارد: یکی استفاده از روش‌های نگاشتن همسایه و دیگری استفاده از نمایش‌های تُنک که جدیدترند و از نظر سرعت و حافظه‌ی مورد نیاز بسیار بهتر عمل می‌کنند. در حل مسئله‌ی فراتفکیک پذیری بر مبنای نمونه برای افزایش تفکیک پذیری به اطلاعاتی خارج از اطلاعات موجود در تصویر داده شده با تفکیک پذیری پایین نیاز است. این اطلاعات از نمونه‌هایی از قطعات به دست آمده از زوج تصاویر با تفکیک پذیری بالا و پایین که از قبل در اختیار است به دست می‌آیند. در روش‌های مبتنی بر نگاشتن همسایه، از این نمونه‌ها مستقیماً استفاده می‌شود، در حالی که در روش‌های مبتنی بر نمایش تُنک، از این قطعات برای آموزش دو لغت‌نامه که برای نمایش قطعات با تفکیک پذیری بالا و پایین به کار می‌روند، استفاده می‌شود [۶]. در راستای بهبود سرعت روش‌های فراتفکیک پذیری که مشکل اصلی این روش‌ها است هنوز کار زیادی انجام نگرفته است. مهمترین عامل کند کننده در این روش‌ها مرحله‌ی به دست آوردن نمایش‌های تُنک است. اگر بتوان روشی به دست آورد که نمایش‌های تُنک را سریعتر (حتی با مقداری تقریب) به دست آورد، راه برای کاربرد بیشتر این روش‌ها باز خواهد شد. یک راه برای افزایش سرعت استفاده از الگوریتم SL است. مشکل این روش این است که در نمایش تُنکی که به دست می‌دهد هیچ کدام از ضرایب صفر نیستند و باید راهی یافت تا ضرایب صفر را پیدا کرد. در این صورت می‌توان امیدوار بود که سرعت الگوریتم بالاتر رود.

در [۷] روش ارائه شده بر این اساس است که قطعه‌های تصویر می‌توانند با استفاده از ترکیب خطی تُنک از مولفه‌های یک لغت‌نامه مناسب به خوبی نمایش داده شوند. در این شیوه، از ضرایب نمایش تُنک هر قطعه از ورودی‌های با تفکیک پذیری پایین برای تولید خروجی‌های تفکیک پذیری بالا استفاده می‌شود. نتایج تئوری از سنجش فشرده<sup>۲</sup> بیانگر این است که تحت شرایط عادی نمایش تُنک می‌تواند به درستی از سیگنال‌های با نرخ نمونه‌ی پایین<sup>۳</sup> بازیابی شود. بوسیله‌ی آموزش

<sup>۱</sup>Sparse Resolution

<sup>۲</sup>Compressed sensing

<sup>۳</sup>Down sample

همزمان<sup>۱</sup> با الحاق دو لغت نامه یکی برای قطعه‌های تفکیک پذیری پایین و یکی برای تفکیک پذیری بالا می‌توان تشابهی از نمایش تُنک بین قطعه‌های تصاویر با تفکیک پذیری پایین و بالای متناظر با لغت نامه ارائه شده ایجاد کرد. بنابراین نمایش تُنک از قطعه‌های تفکیک پذیری پایین می‌تواند به همراه قطعه‌های تفکیک پذیری بالا بکار رود تا لغت نامه با فراتفکیک پذیری بالا را تولید کند. جفت لغت نامه یادگرفته شده نمایش فشرده‌تری از جفت قطعه‌ها نسبت به رهیافت های قبلی می‌باشد. این رهیافت بطور ساده در نمونه های ورودی زیاد (زوج قطعه تصاویر) هزینه محاسباتی را کاهش می‌دهد. مسئله‌ی مینم‌سازی تفکیک پذیری با نمایش تُنک، در اینجا به صورت رابطه ۱-۲ می‌باشد:

$$\min_{\alpha} \|\alpha\|_0 \quad \text{st} \quad \|FD_l\alpha - Fy\|_2 \leq \varepsilon \quad (2-1)$$

که در رابطه‌ی بالا بیانگر ضرایب تُنک و  $F$  بیانگر عملگر استخراج ویژگی و  $y$  قطعه‌های تصویر با تفکیک پذیری پایین و  $D_l$  معرف لغت نامه با تفکیک پذیری پایین است. و همچنین محاسبه‌ی تصویر با تفکیک پذیری بالا را طبق رابطه ۱-۳ داریم:

$$x \approx D_h\alpha \quad (3-1)$$

$D_h$  بیانگر لغت نامه با تفکیک پذیری بالا است.

تاثیر تُنک بودن هم در زمینه فراتفکیک پذیری و هم در زمینه بازسازی تصویر صورت<sup>۲</sup> نشان داده است این الگوریتم‌ها تصاویر با تفکیک پذیری بالا را تولید می‌کنند که در رقابت با روش های دیگر کیفیت بهتری دارند بعلاوه این ارائه‌ی مبتنی بر تُنک در برابر نویز هم مقاوم است بنابراین در ورودی های نویزی کاربرد دارد. در فصل ۳ فراتفکیک پذیری به تفصیل بیان خواهد شد.

## ۳-۱ ردیابی بصری<sup>۳</sup>

ردیابی بصری به دلیل اینکه کاربرد های زیادی در زمینه‌های مختلف مانند حمل و نقل عمومی، وسایل نقلیه عمومی، رباتیک، تعامل انسان و کامپیوتر و ... دارد، توجه زیادی را در شاخه بینایی ماشین به خود جلب کرده است. ردیابی مداوم در محیط های بدون محدودیت به دلیل تغییرات ظاهر جسم ناشی از تغییر حالت، تغییر مقیاس، تغییر شکل و پوشیدگی و غیره، مسئله ای چالش برانگیز است. بنابراین الگوریتم های زیادی برای حل این مسئله ارائه شده است. بعضی از این الگوریتم ها فقط در محیط های با زمینه استاتیک کار می‌کنند ولی بعضی دیگر خود را با تغییرات زمینه و شی تطبیق می‌دهند. زمانی که ظاهر شی با زمینه عوض شود، پوشیدگی جزئی رخ دهد و یا تغییر شکل پدید آید، بیشتر روش های ردیابی بصری موجود قادر

<sup>۱</sup>Jointly training

<sup>۲</sup> Special case of face hallucination

<sup>۳</sup>Visual Tracking

به ردیابی هدف نیستند. برای حل این مشکل می‌توان از روش‌های نمایش تُنک که قدرت بالایی در دسته‌بندی و بازسازی دارند استفاده کرد. فرض تُنک بودن یک سیگنال یک فرض کاملاً کاربردی در بسیاری از سیگنال‌ها است زیرا حتی اگر یک سیگنال در یک حوزه تُنک نباشد، معمولاً می‌توان آن را به یک حوزه دیگر برد که در آن حوزه تُنک باشد. مقالات چاپ شده در مورد استفاده از نمایش تُنک در ردیابی بصری تاکنون، مربوط به سال‌های اخیر بوده است که نشان دهنده‌ی جدید و مورد توجه بودن این روش است. با توجه به روش‌های پیشین ارائه شده در زمینه ردیابی با استفاده از نمایش تُنک و قدرت این روش‌ها مشخص می‌شود که اگر این روش‌ها به صورت مناسبی برای کار ردیابی مدل شوند می‌توانند از پس موانع آن به خوبی برآیند.

### ۱-۳-۱ ردیابی شی در ویدئو<sup>۱</sup>

در ردیابی شی در ویدئو، هدف سامانه ردیاب، پیدا کردن مکان شی در هر فریم ویدئو، با دانستن مکان شی در فریم اول است. یکی از گام‌های اصلی در تحلیل ویدئو، ردیابی شی است که در آن هدف پیدا کردن مکان شی در طول ویدئو است. در رهیافت‌های نمایش تُنک سعی شده است تا عملکرد ردیابی شی را در مواجهه با چالش‌هایی از قبیل انسداد، تغییر حالت، اجسام مشابه در اطراف جسم مورد نظر، در هم ریختگی محیط پیرامون شکل و تغییر نورپردازی بهبود دهد.

در [۸] روش نوین ارائه شده مبتنی بر آنالیز مولفه‌های اصلی وزندار افزایشی<sup>۲</sup> و نمایش تُنک می‌باشد. روند تکراری شامل گام قطعه‌بندی نرم<sup>۳</sup> و گام پروژرسانی توزیع پیش زمینه می‌باشد که به همراه آنالیز مولفه‌های اصلی وزندار افزایشی به کار می‌رود و در الگوی هدف تاثیرات کمتری از پیش زمینه پذیرفته است. برای دست یابی به مدل ظاهری هدف به صورت گسسته‌تر الگوی جزئی<sup>۴</sup> و الگوی پیش زمینه هر دو به لغت‌نامه اضافه می‌شوند. نتایج نشان می‌دهد که روش ارائه شده به همراه اطلاعاتی از پیش زمینه در برابر انسداد، تغییر شدت روشنایی و تنوع شکل ظاهری مقاوم است و از الگوریتم‌های مطرح اخیر هم کارا تر است.

مسئله‌ی بهینه‌سازی به صورت رابطه ۱-۴ می‌باشد

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|x - D\alpha\| \leq \varepsilon \quad (4-1)$$

که  $\alpha$  بیانگر ضرایب نمایش تُنک و  $D$  بیانگر لغت‌نامه است و با رویکرد نمایش تُنک جواب مناسبی به دست خواهد آمد.

<sup>۱</sup> Video Object Tracking

<sup>۲</sup> Incremental weighted PCA

<sup>۳</sup> Soft Segmentation

<sup>۴</sup> Trivial

## ۴-۱ طبقه‌بندی تصاویر<sup>۱</sup>

در این بخش استفاده از نمایش تَنک در کاربرد طبقه‌بندی تصاویر مورد بحث و بررسی قرار می‌گیرد. برای حل مسئله طبقه‌بندی تصویر رویکردهای مختلفی به کار گرفته شده است. در روش‌های معمول و قدیمی‌تر از هر تصویر یک بردار سراسری استخراج می‌شود. شیوه استخراج این بردار برای تمام تصاویر یکسان است و طبیعتاً بردار ویژگی به دست آمده برای تمام تصاویر طول یکسانی دارد. بردار ویژگی باید تا حد ممکن بیانگر اطلاعات موجود در تصویر باشد. ویژگی‌هایی که برای توصیف تصاویر استفاده می‌شود عمدتاً به سه دسته رنگ، بافت و شکل تقسیم بندی می‌شوند. بردار ویژگی استخراج شده توسط روش‌های یادگیری ماشین برای طبقه‌بندی تصاویر به کار گرفته می‌شود. چالشی که بعد از استخراج ویژگی از تصویر وجود دارد ساختن یک مدل بهینه برای طبقه‌بندی بر اساس این ویژگی‌ها است. چالش اصلی در این قسمت مدل کردن زیرفضای داده‌ها و انجام عمل طبقه‌بندی بر اساس مدل ارائه شده است. روش‌های نمایش تَنک در این زمینه قابل استفاده هستند. معمولاً کارایی این روش‌ها برای تصاویری مثل ارقام دست‌نویس و یا تشخیص بافت‌ها ارزیابی می‌شود. چرا که چالش اصلی طبقه‌بندی بعد از استخراج ویژگی است و فرض بر این است که ویژگی استخراج شده اطلاعات کافی برای طبقه‌بندی تصاویر را دارد. روش‌هایی که بهترین کارایی را ارائه کرده‌اند نیز بیشتر توجهشان معطوف به این قسمت بوده است. اما در این روش‌ها نیز بعد از استخراج ویژگی، بردار ویژگی به دست آمده تنها قسمتی از اطلاعات موجود در تصویر را نمایش می‌دهد. لذا به دنبال روشی کارا برای نمایش سیگنال‌ها هستیم. نمایش تَنک می‌تواند پاسخ مناسبی برای این مسئله باشد. چرا از نمایش تَنک استفاده می‌شود؟ تصاویر طبیعی که توسط دوربین‌های دیجیتال امروزی گرفته می‌شود معمولاً تفکیک پذیری در حدود ۱۰ میلیون پیکسل دارد. هر تصویر در یک بردار با ابعاد زیاد ذخیره می‌شود. برای مثال چنانچه تصویر  $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$  در نظر گرفته شود. هر تصویر در فضایی با  $n = 10^6$  بعد نمایش داده می‌شود ولی مجموعه تمام تصاویر طبیعی این فضا را پوشش نمی‌دهد.

اطلاعات موجود در نمایش تصویر به عنوان برداری از پیکسل‌ها دارای ساختارهای تکراری زیادی است که در برخی پژوهش‌ها به عنوان افزونگی یاد می‌شوند. این افزونگی دو علت عمده دارد. اول اینکه دنیای بیرونی که تصاویر از آن گرفته شده دارای ساختارهای تکراری است و دوماً سیستم‌های تصویر برداری با تفکیک پذیری بالایی کار می‌کنند که موجب می‌شود بسیاری اطلاعات حاوی افزونگی در داده مربوط به تصویر گنجانده شود. در واقع تفکیک پذیری بالای تصاویر دیجیتال باعث می‌شود در قسمت‌هایی از تصویر که جزئیات کمتر است اطلاعات اضافی داشته باشیم. برای مثال تصویر ۱-۱ که نمونه‌ای از یک تصویر طبیعی است. قسمت بالای تصویر که مربوط به آسمان است تقریباً در تمام نواحی یک ساختار دارد. اگر کل تصویر را به عنوان ویژگی برای طبقه‌بندی استفاده کنیم این اطلاعات تکراری سبب اشتباه و عملکرد ضعیف طبقه‌بندی خواهد شد. در مسئله طبقه‌بندی تصویر هر چه نمایش ما از داده‌ها به مفهوم نزدیک‌تر باشد طبقه‌بندی با دقت بیشتری انجام خواهد شد. لذا تلاش‌های بسیاری برای مدل کردن فضایی که تصاویر طبیعی در آن قرار می‌گیرند انجام شده است.

به عنوان مثال آنالیز مؤلفه‌های اصلی داده‌ها بر روی پایه‌هایی تصویر<sup>۲</sup> می‌شوند که بیشترین واریانس را در فضای تبدیل

<sup>۱</sup>Image Classification

<sup>۲</sup> Project



شکل ۱-۱: نمونه تصویر برای طبقه‌بندی

یافته دارد. ضعف عمده این روش ثابت بودن پایه‌های تبدیل است در مدل‌های پیچیده‌تر تعداد پایه‌ها از بعد فضا بیشتر است که باعث می‌شود روابط پیچیده‌تر و غیرخطی را بتوان مدل کرد. روش‌های زیادی برای به دست آوردن و یا تخمین فضای واقعی تصاویر طبیعی ارائه شده است. از آنجایی که به دنبال مدل کردن سیگنال‌ها در فضایی با بعد کمتر هستیم عموماً به این روش‌ها مدل‌های کم‌بعد سیگنال<sup>۱</sup> گفته می‌شود.

در بیشتر کاربردهای پردازش سیگنال پایه‌هایی که از آن برای نمایش در فضای جدید استفاده می‌شود مشخص است و نیازی به ارائه روش یا الگوریتم برای به دست آوردن پایه‌ها نیست. اما در مسئله طبقه‌بندی تصاویر مجموعه خاصی از تصاویر به ما داده شده است و برای نمایش  $\mathbf{t}_k$  این تصاویر باید پایه‌هایی را به دست آوریم که برای نمایش  $\mathbf{t}_k$  تصاویر ما مفید باشد.

## ۱-۴-۱ چالش‌های اصلی در طبقه‌بندی تصاویر با استفاده از نمایش $\mathbf{t}_k$

هر تصویر در مجموعه داده، توسط یک بردار با ابعاد بسیار بالا ذخیره می‌شود. اگر این بردار توسط یک لغت‌نامه به شکل  $\mathbf{t}_k$  نمایش داده شود، عملاً غیرممکن خواهد بود، چرا که اندازه لغت‌نامه بسیار بزرگ خواهد بود و اعمال فرض  $\mathbf{t}_k$  بودن نیز برای لغت‌نامه‌ای با این اندازه امکان‌پذیر نیست. روشی که برای حل این مشکل در اکثر مقالات در نظر گرفته شده [۹] استفاده از قطعه‌هایی<sup>۲</sup> از تصویر به جای کل تصویر است. با انجام این کار به صورت ضمنی فرض شده است که قطعه‌های تصویر از توزیع احتمالی خاصی پیروی می‌کنند و با در نظر گرفتن لغت‌نامه مناسب دارای نمایش  $\mathbf{t}_k$  هستند.

یکی دیگر از مسائل مهم در روش‌های یادگیری ماشین بر اساس نمایش  $\mathbf{t}_k$ ، نحوه استفاده از برچسب داده‌ها است. نمایش  $\mathbf{t}_k$  در کاربردهایی مانند حذف نویز یا فشرده‌سازی به سادگی و با کارایی مناسب به کار گرفته شده است [۱۰]. در این مسائل بازسازی سیگنال مورد توجه است. اما در مسئله طبقه‌بندی توان جداسازی بیشتر برای طبقه‌بند حائز اهمیت است. از این رو تنها مدل کردن سیگنال‌ها برای به دست آوردن فضای کم بعد سیگنال‌ها کافی نیست. روش‌هایی ارائه شده است

<sup>۱</sup> Low Dimensional Signal Models      <sup>۲</sup> Patch

که سعی می‌کنند نمایش جدید سیگنال‌ها خاصیت جداسازی بیشتری داشته باشند [۱۱]. با این وجود استفاده از برچسب‌ها در ساخت لغت‌نامه و نمایش تَنک سیگنال‌ها یکی از چالش‌های موجود است. علاوه بر دو مسئله ای که ذکر شد در تمام روش‌های ارائه شده در حوزه نمایش تَنک زمان اجرای الگوریتم یکی از نکات مهم است. روشی که ارائه می‌شود باید از لحاظ تئوری و در عمل دقت مناسبی داشته باشد و همچنین برای کاهش پیچیدگی زمانی راهکاری ارائه دهد [۱۲].

## ۲-۴-۱ طبقه‌بندی تصاویر ابرطیفی

از جمله روش‌های مبتنی بر نمایش تَنک در این حوزه [۱۳] طبقه‌بندی تصاویر ابرطیفی<sup>۱</sup> با استفاده از رهیافتی جدید برای یادگیری لغت‌نامه با استفاده از نمایش تَنک ساختاریافته مبتنی بر سنجش فشرده است. یک مرحله‌ی کارساز در این روش بخش‌بندی پیکسل‌های تصاویر ابرطیفی به تعدادی فضای همسایگی است که گروه‌های پیکسلی نامیده می‌شوند. این گروه‌های پیکسلی می‌توانند با سایزهای متفاوت مدل شوند. این ایده براساس به‌کارگیری همبستگی طیفی برای تصاویر سنجش از راه دور<sup>۲</sup>، بین پیکسل‌ها است و هدف، به‌دست آوردن لغت‌نامه برای هر پیکسل است. این لغت‌نامه ترکیب خطی از تعداد کمی از لغت‌نامه‌ی آموزش داده شده از داده‌های ابرطیفی است و می‌تواند به درستی تصاویر سنجش از راه دور را با ضرایب کمتر نمایش دهد. این پیکسل‌ها الگوی تَنک بودن را القاء می‌کنند و به طور ضمنی دارای همبستگی طیفی در گروه‌های پیکسلی همانی هستند. در ادامه ضرایب تَنک به وسیله ماشین بردار پشتیبان خطی<sup>۳</sup> برای طبقه‌بندی تصاویر ابرطیفی به‌کار می‌روند. نتایج نشان می‌دهد که روش ارائه شده نمایش بهتری برای تصاویر ابرطیفی و صحت کلی بالاتری دارد.

مسئله‌ی بهینه‌سازی در این روش به صورت رابطه‌ی ۱-۵ می‌باشد.

$$\arg \min_{\alpha} \frac{1}{\gamma} \|x - D\alpha\|_F^{\gamma} + \sum_{i=1}^g \gamma_{H_i} \|\alpha_{H_i}\|_{\gamma} \quad s.t. \quad \forall_i \|d_i\|_{\gamma} \leq 1 \quad (5-1)$$

که  $\alpha$  بیانگر ضرایب نمایش تَنک و  $D$  بیانگر لغت‌نامه و  $\gamma_{H_i}$  پارامتر تنظیم‌کننده است. طبق نتایج آزمایشات، الگوریتم یادگیری لغت‌نامه ساختاریافته به طور قابل ملاحظه‌ای صحت طبقه‌بندی را بهبود می‌بخشد. در عمل الگوریتم SDL<sup>۴</sup> توانایی بهبود صحت طبقه‌بندی از ۰.۷۶۷۸ (بدون استفاده از یادگیری لغت‌نامه ساختاریافته) تا ۰.۹۶۷۹ را دارد. از لحاظ زمان اجرا، زمان اجرا در SVM کمتر از زمان اجرا در الگوریتم SDL است و این امر بیانگر اینست که یادگیری لغت‌نامه ساختاریافته صحت طبقه‌بندی را با افزایش زمان اجرا بهبود می‌بخشد. در شکل ۲-۱ خروجی این روش نشان داده شده است.

<sup>۱</sup> Hyperspectral

<sup>۲</sup> Remote sensing images

<sup>۳</sup> Linear SVM

<sup>۴</sup> Structural Dictionary Learning



شکل ۱-۲: (آ) تصویر اصلی ، (ب) تصویر مرجع ، (پ) خروجی قطعه‌بندی با روش SVM و (ت) خروجی قطعه‌بندی با روش SDL

### ۱-۴-۳ تشخیص چهره<sup>۱</sup>

روشی که در [۱۴] ارائه شده است یکی از روش‌های موفق استفاده از نمایش تُنک در مسائل یادگیری ماشین مانند تشخیص چهره است. تصاویر چهره عموماً ابعاد بالایی دارند ولی از آنجا که تغییرات در این تصاویر اندک است می‌توان انتظار داشت این تصاویر روی زیرفضایی با بعد کم قرار گیرند. یکی از ساده‌ترین روش‌ها برای استخراج ویژگی و تخمین این زیر فضا روش چهره‌های ویژه<sup>۲</sup> است. در طبقه‌بندی با روش تُنک نیز فرض بر این است که تصاویر مربوط به یک دسته با ترکیب خطی تعداد کمی از داده‌های همان دسته قابل نمایش هستند. بنابراین در این روش لغت‌نامه متشکل از خود تصاویر است. در مسئله تشخیص چهره N تصویر چهره به عنوان ورودی داده شده است که هر کدام به یکی از K دسته تعلق دارند. اگر ابعاد هر کدام از تصاویر  $w * h = m$  باشد و  $n_i$  نشان دهنده تعداد تصاویر در دسته i باشد. ماتریس  $D_i$  حاوی کل تصاویر دسته i است  $D_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,n_i}\}$  و کل تصاویر در ماتریس  $D = [D_1, D_2, D_3, \dots, D_K]$  قرار گرفته است. در این نمایش هر ستون از ماتریس D مربوط به یکی از تصاویر در یکی از دسته‌ها است. نمایش یک تصویر آزمون مانند  $x$  به شکل تُنک به صورت  $x = D\alpha + z$  خواهد بود. در این نمایش  $z$  نشان‌دهنده نویز است که فرض می‌کنیم نویزی گاوسی است و انرژی آن محدود است ( $\|z\|_2 < \epsilon$ ). این فرض تضمین می‌کند که با خطای محدودی پاسخ تُنک وجود دارد. برای اینکه نمایش به دست آمده تُنک باشد مسئله را به شکل رابطه ۱-۶ تعریف می‌کنیم. برای حل این مسئله می‌توان از روش‌های معمول بهینه‌سازی استفاده کرد.

$$\arg \min_{\alpha} \|\alpha\|_1 \quad s.t. \quad x = D\alpha \quad (۶-۱)$$

که  $\alpha$  بیانگر ضرایب نمایش تُنک و D بیانگر لغت‌نامه است. نمایش به دست آمده از رابطه ۱-۶ معیار مناسبی برای تشخیص داده آزمون خواهد بود. چرا که اگر نمایش برای داده چهره موجود باشد این نمایش معادل یافتن زیرفضای کم‌بعدی است که با داده‌های آزمایشی ساخته می‌شود و بنابر فرض اولیه داده‌هایی که در ساخت این زیرفضا مشارکت می‌کنند با داده آزمون هم‌برچسب هستند. طبقه‌بندی‌های متفاوتی برای داده آزمون می‌توان در نظر گرفت. برای مثال اگر بزرگ‌ترین درایه  $\alpha$  را با

<sup>۱</sup>Image Recognition      <sup>۲</sup> Eigen Faces

$\hat{\alpha}$  نشان دهیم، ستون متناظر با این درایه در ماتریس  $D$  معیاری برای طبقه‌بندی خواهد بود و دسته مربوط به این ستون به عنوان دسته داده آزمون در نظر گرفته می‌شود. اما با این کار از قابلیت ساختن فضا توسط لغت‌نامه استفاده نشده است. به عبارت دیگر باید بررسی کنیم اتم‌های کدام دسته از لغت‌نامه در بازسازی سیگنال آزمون موفق‌تر عمل کرده است. همان طور که از عنوان این روش پیداست، قابلیت اصلی آن مقاوم بودن در برابر نویز است. هم چنین نسبت به روش‌های دیگر نمایش  $\hat{\alpha}$  که برای طبقه‌بندی استفاده شده اند زمان اجرای مناسب‌تری دارد، چرا که برای هر داده آزمون تنها یک بار کدگذاری  $\hat{\alpha}$  داریم. اما این روش، با این فرض کارایی قابل قبولی دارد که داده‌های آموزشی به تعداد زیاد در دسترس باشد. در غیر اینصورت ساخت لغت‌نامه و بازسازی سیگنال‌ها توسط داده‌های آموزشی امکان پذیر نخواهد بود. البته باید توجه کنیم که اگر تعداد داده‌های آموزشی خیلی زیاد باشد محاسبه نمایش  $\hat{\alpha}$  بسیار زمان‌گیر خواهد بود. در حالت کلی عملکرد این روش وابسته به تعداد داده‌های آموزشی است. با در نظر گرفتن نکاتی که گفته شد می‌توان نتیجه گرفت این روش برای داده‌هایی که در فضای کم بعد قرار می‌گیرند و تعداد آن‌ها نیز معقول است کاربرد دارد.

همچنین در [۱۵] روشی سریع و کارا برای شناسایی چهره مبتنی بر نمایش  $\hat{\alpha}$  ارائه می‌شود. با توجه به اینکه نمایش  $\hat{\alpha}$  مبتنی بر محاسبه نمایش  $\hat{\alpha}$  با بهینه سازی نرم  $l_1$  برای لغت‌نامه با ابعاد بالا دارای حجم محاسبات بالا است، هنگامی که تعداد ویژگی‌های چهره کم است و شرایط یکتایی نمایش  $\hat{\alpha}$  نیز به درستی رعایت نشده است یا مدل بکار رفته دقیق نباشد، خطر گیر افتادن در مینیمم‌های محلی خیلی زیاد است. از این رو در اینجا، ایده نمایش  $\hat{\alpha}$  و زنده‌ار برای محدود کردن فضای جستجو ارائه شده است. نتایج نشان می‌دهد که روش ارائه شده دقت بالایی در شناسایی چهره با حجم محاسباتی کمتر دارد. مسئله‌ی بهینه‌سازی در این روش به صورت رابطه‌ی ۷-۱ می‌باشد.

$$\arg \min_{\alpha} \sum w_i \alpha_i^* \quad s.t. \quad x = D\alpha \quad (7-1)$$

که  $\alpha$  بیانگر ضرایب نمایش  $\hat{\alpha}$  و  $D$  بیانگر لغت‌نامه و  $w_i$  ها وزن‌های کنترلی و اعدادی مثبت هستند.

## ۵-۱ تخمین حالت انسان<sup>۱</sup>

مسئله‌ی تخمین سه بعدی حالت انسان یکی از مهمترین و پیچیده‌ترین مسائلی است که در زمینه‌ی بینایی ماشین مطرح می‌باشد، نیازها و کاربردهایی که در تخمین حالت وجود دارد باعث شده است که در سال‌های اخیر توجه ویژه‌ای به این بخش از تحقیقات شود و خیلی از آزمایشگاه‌های معتبر در دانشگاه‌های سراسر جهان بخشی از تحقیقات خود را به این بخش از مسائل بینایی ماشین اختصاص داده‌اند. در تخمین حالت انسان هدف تخمین حالت دو یا سه بعدی حالت بدن انسان با استفاده از داده‌های در دسترس می‌باشد، این داده‌ها ممکن است تصویر، ویدئو یا داده‌های حاصل از حسگرهای نوری یا مکانیکی باشد. با توجه به اینکه مسئله تخمین حالت یکی از سخت‌ترین زمینه‌های بینایی ماشین می‌باشد کارهایی

<sup>۱</sup> Human Pose Estimation



که محققان در سال‌های اخیر انجام داده‌اند با در نظر گرفتن فرضیات خاصی بوده و هر کدام به حل موردی و جزئی مشکلات پرداخته‌اند و هنوز مشکلاتی همچون ابهام در عمق<sup>۱</sup> حالت تخمین زده شده، خود انسدادی<sup>۲</sup>، تاری، تغییرات زیاد حالت و ابعاد بالای فضای ورودی و خروجی، تغییر نور در محیط و پیچیده بودن پس زمینه هنوز حل نشده باقی مانده‌اند و اکثر تحقیقاتی که در این زمینه بوده است به صورت جزئی با در نظر گرفتن محدودیت‌هایی به حل این مشکلات پرداخته‌اند. در چند سال گذشته روش‌های نمایش تنگ‌کارایی خوبی در مسائل بینایی ماشین داشته‌اند. مهمترین نکته در تخمین حالت این است که ابتدا کاربرد خود را به درستی بشناسیم و به شناسایی محدودیت‌ها و مشکلات پیش رو پردازیم. مثلاً تغییرات زیاد حرکت در تخمین حالت در فوتبال از اهمیت فراوانی برخوردار است اما همین مسئله در تخمین حالت در فضایی که شخص تحرک کمی دارد بی‌اهمیت است [۱۶].

از جمله روش‌های به‌کار گرفته شده با نمایش تنگ در [۱۷] تشخیص حالت دست با استفاده از ویژگی‌های مبتنی بر نمایش تنگ می‌باشد. با اینکه امروزه طبقه‌بندی نمایش تنگ کاربرد بسیار گسترده‌ای در حوزه‌های بینایی ماشین از جمله تشخیص چهره دارد اما تحقیقات کمی در زمینه به‌کارگیری طبقه‌بندی نمایش تنگ در شناسایی حالت دست انجام شده است. در اینجا روشی ارائه شده است که برجستگی بر پایه‌ی ویژگی و محاسبه ضرایب تنگ را برای تشخیص حالت دست به کار می‌گیرد. به علاوه، تحقیقات گذشته نشان می‌دهد که طبقه‌بندی نمایش تنگ زمانی که ویژگی‌های ما غیرخطی است نتایج شناسایی خوبی ندارد. به همین منظور در این روش از تابع کرنل تقاطع هیستوگرام برای نگاشت ویژگی‌های اصلی به فضای ویژگی کرنل استفاده شده است و در فضای ویژگی کرنل از طبقه‌بندی نمایش تنگ بهره برده است. میانگین نرخ شناسایی در این روش ۹۸.۹۱ درصد می‌باشد، که نشانگر اثربخشی این روش است. مسئله‌ی بهینه‌سازی در این روش به صورت رابطه‌ی ۱-۸ می‌باشد.

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad x = D\alpha \quad (8-1)$$

که  $\alpha$  بیانگر ضرایب نمایش تنگ و  $D$  بیانگر لغت‌نامه است.

## ۶-۱ تصاویر پزشکی

سیستم‌های تشخیص پزشکی به منظور کمک به پزشکان در تشخیص‌های پزشکی ایجاد شده‌اند. یکی از مهمترین حوزه‌هایی که سیستم‌های تشخیص پزشکی در آن به‌کار گرفته شده است، عکسبرداری پزشکی است. ویژگی‌های خاص این تصاویر نیاز به وجود چند مرحله پیش پردازش از قبیل بهبود کیفیت تصویر و ناحیه بندی تصاویر را الزامی کرده است. همچنین استخراج ویژگی از این تصاویر جهت به‌کارگیری تکنیک‌های یادگیری ماشین تا حدودی متفاوت است. در [۱۸] روشی مبتنی بر نمایش تنگ ارائه شده است. در تو موگرافی با داده‌های محدود در تصاویر پزشکی، نمای اسکن

<sup>۱</sup> Depth Ambiguity

<sup>۲</sup> Self-occlusion

شده چه در نواحی داخلی چه در گوشه‌ها به صورت محدود و تُنگ نمونه‌برداری شده است. در این شرایط الگوریتم‌های استاندارد از جمله FBP<sup>۱</sup> در بازسازی نتایج خوبی ندارند. در این جا نشان داده شده است که سیاست نمایش تُنگ مبتنی بر یادگیری لغت‌نامه برای تکه‌های تصویر بهبود قابل توجهی در بازسازی در تصویرکردن گوشه‌ها دارد. در اینجا در گام ساخت لغت‌نامه از K-SVD و در گام حل مسئله‌ی بهینه‌سازی از روش OMP استفاده شده است.

## ۷-۱ ترکیب تصاویر با هدف بهبود کیفیت<sup>۲</sup>

در این زمینه با استفاده از تصاویر ثبت شده توسط شبکه دوربین‌ها و یا تصاویر یک دوربین که در زمان‌ها و مکان‌های مختلف از یک منظره ضبط شده اند، کیفیت تصویر را افزایش می‌دهند. این مسئله به دلیل گسترش کاربرد شبکه دوربین‌ها و افزایش میل کاربران به برخورداری از تصاویر با کیفیت‌تر در سال‌های اخیر مورد توجه بوده و از اهمیت ویژه‌ای برخوردار است. در [۱۹] یکی از مشکلات دوربین‌های تصویربرداری بیان می‌شود و این مشکل محدودیت عمق میدان<sup>۳</sup> در لنزهای نوری می‌باشد. از این رو به دست آوردن تصویری که دوربین برای تمامی اشیاء فوکوس داشته باشد امری دشوار است. یکی از راه‌های غلبه بر این مشکل استفاده از تکنیک‌های ترکیب تصاویر است به گونه‌ای که چندین تصویر با نقاط فوکوس متفاوت با هم ترکیب شوند تا تصویری حاصل شود که در آن روی تمامی اشیاء فوکوس شده باشد. استخراج کارای اطلاعات از تصویر اصلی و ترکیب منطقی آن‌ها برای تولید تصویر ترکیبی نهایی امری حیاتی است. مسئله‌ی بهینه‌سازی در این روش مشابه قبل به صورت رابطه‌ی ۱-۹ می‌باشد.

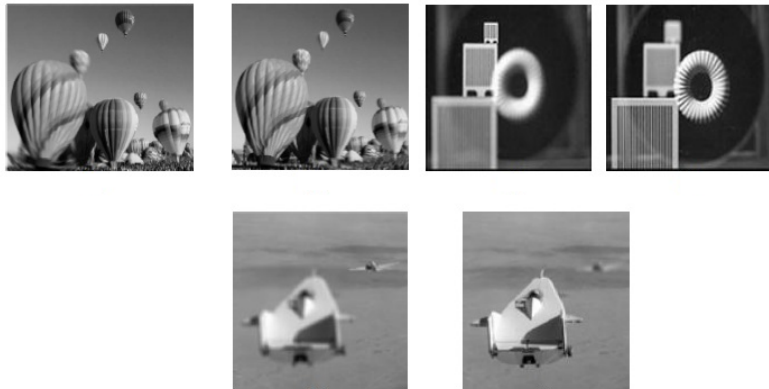
$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|x - D\alpha\| \leq \varepsilon \quad (9-1)$$

که  $\alpha$  بیانگر ضرایب نمایش تُنگ و  $D$  بیانگر لغت‌نامه است. گام اصلی برای این رهیافت، نمایش تُنگ تصاویر اولیه به صورت ضرایب تُنگ از لغت‌نامه فراکامل و ترکیب آن‌ها با الگوریتم OMP می‌باشد. روش نمایش تُنگ برای حل مشکل ترکیب تصاویر می‌تواند به طور همزمان با ترمیم تصویر به کار رود. تحت شرایط خطای تقریبی معین، با به کارگیری از الگوریتم OMP، ویژگی‌های تصویر از قبیل ساختارهای هندسی می‌تواند به خوبی بیان شوند و با یکدیگر ترکیب شوند. آزمایشات انجام شده بیان می‌دارد که این روش در عمل کارا است و در مقایسه با روش‌های پیشین چه از لحاظ مفهوم بصری و چه از لحاظ معیارهای ارزیابی حتی در نمونه‌ای نویزی عملکرد بهتری دارد. در شکل ۱-۳ و شکل ۱-۴ خروجی این روش بر روی تصاویر ورودی نشان داده شده است.

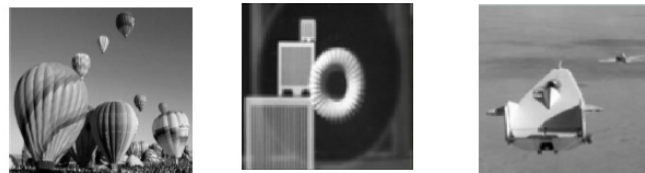
<sup>۱</sup> Filtered back-projection

<sup>۲</sup> Image Fusing

<sup>۳</sup> Depth-of-field



شکل ۱-۳: ورودی ها با فوکوس متفاوت



شکل ۱-۴: تصویر خروجی با کیفیت بالا ترکیب شده از تصاویر ورودی

## ۸-۱ مقایسه روش‌های بهینه‌سازی تُنک در زمینه‌های کاربردی

با بررسی زمینه‌های مختلف در پردازش تصویر و کارا بودن رهیافت نمایش تُنک در آن‌ها قدرتمند بودن این روش نشان داده شد گرچه که هنوز چالش‌هایی وجود دارد که با به‌کارگیری الگوریتم‌های کارا تر چه در یادگیری لغت‌نامه و چه در کدگذاری تُنک می‌توان بر آن‌ها غلبه کرد در جدول ۱-۱ مقایسه‌ای برای روش‌های ذکر شده نشان داده شده است.

جدول ۱-۱: مقایسه روش‌های بهینه‌سازی تُنک در زمینه‌های کاربردی

روش معرفی شده در حوزه	الگوریتم یادگیری لغت‌نامه	حل معادله مینیمم‌سازی
فرا تفکیک پذیری تصویر [۷]	L1QP	L1QP
ردیابی شی در ویدئو [۸]	K-SVD	BOMP
طبقه بندی تصاویر [۱۳]	Linear SVM classifier	M-FOCUSS
تشخیص چهره [۱۴]		NSWLO
تصاویر پزشکی [۱۸]	K-SVD	OMP
ترکیب تصاویر [۱۹]	OMP	OMP

در این فصل مسئله نمایش تُنک و کاربردهای آن در زمینه‌های مختلف و مقایسه‌ی روش‌های بهینه‌سازی بکارگرفته شده

در آن بررسی شد. همچنین به روش‌هایی از جمله OMP و K-SVD و L1QP اشاره شد که در فصل ۲ به تفصیل توضیح داده خواهد شد.

## فصل ۲

# روش‌های بهینه‌سازی تَنک

فرض تَنک بودن بیان می‌کند که این داده را می‌توان به شکل ترکیب خطی تعداد کمی از پایه‌ها که از قبل در نظر گرفته شده بیان کرد. تعداد پایه‌هایی که در نمایش  $x$  موثر هستند با  $k$  نشان داده می‌شود. این تعداد باید به شکل قابل توجهی از بعد فضای اصلی داده کوچکتر باشد. پایه‌هایی که برای نمایش داده‌ها استفاده می‌شود در یک ماتریس قرار داده می‌شود. این ماتریس لغت‌نامه نامیده می‌شود. عموماً این ماتریس کل فضای برداری مربوط به داده‌ها را پوشش می‌دهد. به عبارت دیگر ستون‌های این ماتریس برای ساخت کل یا قسمتی از داده‌هایی که از فضای برداری  $R^n$  هستند استفاده می‌شوند. هرکدام از ستون‌های ماتریس لغت‌نامه، که پایه‌های مورد نیاز برای مدل کردن داده هستند یک اتم نامیده می‌شود. اگر تعداد اتم‌های لغت‌نامه به اندازه بعد فضای برداری باشد و پایه‌ها کل فضا را پوشش دهند ماتریس لغت‌نامه کامل نامیده می‌شود. در این حالت هر داده یک نمایش یکتا توسط اتم‌های لغت‌نامه خواهد داشت. در مسئله‌ی بهینه‌سازی تَنک:

$$\arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad x = D\alpha \quad (1-2)$$

رابطه‌ی ۱-۲ نمایش تَنک سیگنال  $x$  را می‌دهد. در این عبارت  $\alpha$  برداری شامل نمایش تَنک سیگنال است. فرض کنیم  $D$  بیانگر لغت‌نامه باشد و سیگنال با  $x$  نمایش داده شود. طبق معادله‌ی ۲-۲ که در زیر آمده است

$$D\alpha = x \quad (2-2)$$

در واقع بردار  $x$  به صورت ترکیب خطی از ستون‌های لغت‌نامه‌ی  $D$  نوشته می‌شود. اگر  $D$  یک ماتریس مربعی با رتبه کامل باشد، در این صورت نمایش سیگنال  $x$  توسط ماتریس  $D$ ، یعنی بردار  $\alpha$  به سادگی با معکوس کردن ماتریس  $D$  قابل محاسبه است:

$$\alpha = D^{-1}x \quad (3-2)$$

تا قبل از مطرح شدن نمایش‌های تنک، لغت‌نامه‌های مورد استفاده تقریباً همه از این نوع ماتریس‌های مربعی با رتبه‌ی کامل بودند. در حالتی که تعداد ستون‌های لغت‌نامه‌ی  $D$  بیشتر از تعداد سطرهای آن باشد، بی‌نهایت  $\alpha$  وجود خواهد داشت که در معادله‌ی ۲-۲ صدق کند. در این حالت بسته به کاربرد مورد نظر روی  $\alpha$  فرض‌های اضافه‌ای باید گذاشت تا جواب یکتا شود. در ادامه برخی از روش‌های حل مسئله‌ی بهینه‌سازی تنک معرفی خواهد شد.

## ۱-۲ روش جست و جوی تطابقی

در روش جست و جوی تطابقی<sup>۱</sup> ابتدا ستون‌های مورد استفاده از ماتریس لغت‌نامه در ترکیب خطی آشکار شده، سپس ضرایب این ستون‌ها که همان مقادیر غیر صفر بردار  $s$  هستند با حل مسئله حداقل مربعات محاسبه می‌شوند. این روش ابتدا توسط ژانگ و ملت<sup>۲</sup> [۲۱] مطرح شد. در این روش در هر مرحله تنها ضریب یکی از اتم‌ها را مشخص می‌کند. در هر مرحله یک اتم لغت‌نامه که بیشترین شباهت (بزرگ‌ترین ضرب داخلی) را به داده آزمون دارد، بعنوان عضو فعال در ترکیب خطی در نظر گرفته شده ضریب مربوط به آن محاسبه می‌شود. در مرحله بعد، باقی‌مانده سیگنال آزمون و اتم نخست با بقیه اتم‌ها مقایسه شده و دوباره مشابه‌ترین اتم انتخاب می‌شود. یعنی تفاضل حاصل ضرب این تقریب ۱-تنک در لغت‌نامه از داده آزمون را به عنوان باقی‌مانده در نظر گرفته، مراحل فوق تکرار می‌شوند. در هر مرحله جمع تقریب‌های ۱-تنک به دست آمده با تقریب‌های قبلی به عنوان تقریب جدید در نظر گرفته می‌شود و این روند تا جایی ادامه می‌یابد که یا تعداد مراحل مشخصی طی شود و یا خطا از مقدار معینی کمتر شود.

فرض کنید  $b_{m \times 1} = A_{m \times n} x_{n \times 1}$  و  $x$  یک بردار  $k$  تنک است. در این صورت  $b$  را می‌توان بصورت ترکیب خطی  $k$  تا از ستون‌های  $A$  تلقی کرد و عمل بازسازی، معادل با یافتن این  $k$  ستون و ضرایب آن‌ها در ترکیب خطی است. در این روش ابتدا ستون‌های مورد استفاده از  $A$  در ترکیب خطی آشکار می‌شوند<sup>۳</sup> و سپس به کمک حل یک مسئله حداقل مربعات، ضرایب این ستون‌ها که همان مقادیر ناصفر بردار  $x$  هستند محاسبه می‌شوند. در هر مرحله از روش MP یکی از ستون‌های  $A$  به عنوان عضوی فعال در ترکیب خطی آشکار می‌شود، در این مرحله ضرب داخلی  $b$  با تمام ستون‌های  $A$  محاسبه می‌شود و ستونی که بیشترین اندازه را حاصل کند (که بیشترین شباهت را به  $b$  دارد) به عنوان عضو فعال شناسایی می‌شود و مقدار ضرب داخلی که به اندازه این ستون تقسیم شده را به عنوان ضریب آن لحاظ می‌کنیم. اکنون تقریبی ۱-تنک از بردار اصلی بدست آورده‌ایم که در این مرحله انتخاب بهینه به شمار می‌رود  $(\hat{x}_{n \times 1}^{(1)})$ . برای ادامه، با فرض اینکه محاسبات تا این لحظه صحیح است، بردار  $r_{m \times 1} = b_{m \times 1} - A_{m \times n} \hat{x}_{n \times 1}^{(1)}$  را به عنوان باقی‌مانده<sup>۴</sup> در نظر می‌گیریم و تمامی مراحل قبلی را طی می‌کنیم تا تخمین ۱-تنک جدیدی بدست آید. اکنون حاصل جمع این بردار ۱-تنک جدید و  $(\hat{x}_{n \times 1}^{(1)})$  را به عنوان تقریب جدید برای  $(\hat{x}_{n \times 1}^{(2)})$  تلقی می‌کنیم و این مراحل را مجدداً تا رسیدن به شرط پایانی (نرم باقی‌مانده کوچک یا تعداد مرحله معلوم و یا ترکیبی از هر دو) ادامه می‌دهیم. الگوریتم MP به دلیل اینکه در هر مرحله نیاز به یک جستجوی ساده دارد، بسیار سریع است، اما به دلیل حریصانه بودن، تضمینی وجود ندارد که پاسخ نهایی مشابه با پاسخ تنک باشد. بمنظور بهبود این

<sup>۱</sup>MP

<sup>۲</sup>Malat, zhang

<sup>۳</sup>Support Recovery

<sup>۴</sup>Residual

---

الگوریتم ۱-۲ جست و جوی تطابقی با ورودی‌های ماتریس  $A$  و بردار  $b$  و حد آستانه  $\varepsilon_0$ .

---

**Require:** We are given the matrix  $A$ , the vector  $b$ , and the error threshold  $\varepsilon_0$  for approximating the solution of  $(P0) : \min_x \|x\|_0$  subject to  $Ax = b$

**Ensure:** The proposed solution is  $x^k$  obtained after  $k$  iterations

// Initialize

1:  $k = 0$ ,

2: The initial solution  $x^0 = 0$ ,

3: The initial residual  $r^0 = b - Ax^0 = b$

4: The initial solution support  $S^0 = \text{Support}\{x^0\} = 0$

// Main Iteration

5: Increment  $k$  by 1 and perform the following steps:

//Sweep

6: Compute the errors  $\varepsilon(j) = \min_{z_j} \|a_j z_j - r^{k-1}\|_2^2$  for all  $j$  using the optimal choice

$$z_j^* = a_j^T r^{k-1} / \|a_j\|_2^2$$

//Update Support

7: Find a minimizer,  $j_0$  of  $\varepsilon(j) : \forall 1 \leq j \leq m, \varepsilon(j_0) \leq \varepsilon(j)$ , and update  $S^k = S^{k-1} \cup \{j_0\}$ .

//Update Provisional Solution

8: Set  $x^k = x^{k-1}$ , and update the entry  $x^k(j_0) = x^{k-1}(j_0) + z_j^*$

//Update Residual

9: Compute  $r^k = b - Ax^k = r^{k-1} - z_{j_0}^* a_{j_0}$

//Stopping Rule

10: If  $\|r^k\|_2 < \varepsilon_0$ , stop. Otherwise, apply another iteration.

---

الگوریتم با حفظ سرعت بالا، الگوریتم‌های دیگری پیشنهاد شده‌است که در ذیل به شرح و توضیح آن پرداخته شده‌است.

## ۲-۲ جست و جوی تطابقی متعامد

در روش OMP که تعمیم یافته روش قبل است در هر مرحله ضرایب ستون‌های فعال از ماتریس لغت‌نامه به صورت مستقل از نتایج مراحل قبل انتخاب می‌شوند و از نتایج قبلی تنها در یافتن مکان مولفه‌های غیرصفر استفاده می‌شود. یعنی در هر مرحله بعد از مشخص شدن اتم جدید، از تمام ضرایب قبلی صرف نظر شده و ترکیبی خطی از تمام اتم‌های انتخاب شده تا این مرحله، محاسبه می‌شود که کمترین خطا را برای بازسازی سیگنال آزمون داشته باشد. از جمله مهم‌ترین مزایای آن سرعت زیاد آن است.

الگوریتم ۲-۲ جست و جوی تطابقی متعامد با ورودی‌های ماتریس  $A$  و بردار  $b$  و حد آستانه  $\varepsilon_0$ .

**Require:** We are given the matrix  $A$ , the vector  $b$ , and the error threshold  $\varepsilon_0$  for approximating the solution of  $(P0) : \min_x \|x\|_0$  subject to  $Ax = b$

**Ensure:** The proposed solution is  $x^k$  obtained after  $k$  iterations

// Initialize

1:  $k = 0$ , and set

2: The initial solution  $x^0 = 0$ ,

3: The initial residual  $r^0 = b - Ax^0 = b$

4: The initial solution support  $S^0 = \text{Support}\{x^0\} = 0$

//Main Iteration

5: Increment  $k$  by 1 and perform the following steps:

//Sweep

6: Compute the errors  $\varepsilon(j) = \min_{z_j} \|a_j z_j - r^{k-1}\|_2^2$  for all  $j$  using the optimal choice

$$z_j^* = a_j^T r^{k-1} / \|a_j\|_2^2$$

//Update Support

7: Find a minimizer,  $j_0$  of  $\varepsilon(j) : \forall j \notin S^{k-1}, \varepsilon(j_0) \leq \varepsilon(j)$ , and update  $S^k = S^{k-1} \cup \{j_0\}$ .

//Update Provisional Solution

8: Compute  $x^k$ , the minimizer of  $\|Ax - b\|_2^2$  subject to  $\text{Support}\{x\} = S^k$

//Update Residual:

9: Compute  $r^k = b - Ax^k$

//Stopping Rule

10: If  $\|r^k\|_2 < \varepsilon_0$ , stop. Otherwise, apply another iteration.

## ۳-۲ روش KSVD

مبنای الگوریتم‌های آموزش لغت‌نامه تکرار دوگام نمایش تنگ و بروز کردن لغت‌نامه است و تفاوت این الگوریتم‌ها بیشتر در گام بروزکردن لغت‌نامه است. الگوریتم KSVD یکی از موفق‌ترین الگوریتم‌های آموزش لغت‌نامه بوده که نسبت به بقیه الگوریتم‌ها رابطه‌ای نزدیک‌تر با الگوریتم K-means دارد. KSVD در گام نمایش تنگ از الگوریتم OMP استفاده می‌کند. انتخاب این الگوریتم به دلیل سرعت بالای آن نسبت به بقیه‌ی الگوریتم‌های کدینگ تنگ است. اگر چه به دلیل ماهیت حریص آن ممکن است جواب‌های نادرستی داشته باشد. برای بروز کردن لغت‌نامه، KSVD اتم‌ها را یک به یک و بطور متوالی به روز می‌کند. در روش K-means هم بردارهای کد یک به یک بروز می‌شوند. برای بروز کردن اتم  $d_k$ ، بقیه اتم‌ها را ثابت می‌گیریم. در اینصورت داریم:

$$E = \|Y - DX\|_F^2 \quad (۴-۲)$$



که در آن منظور از  $X_T^i$  سطر  $i$  ام ماتریس  $X$  است. در نتیجه، برای بروز کردن  $d_k$  باید مسئله‌ی زیر را حل کنیم:

$$\min_d \left\| E_k - dx_T^k \right\|_F^2 \quad s.t. \quad \|d\|_2 = 1 \quad (5-2)$$

ایده‌ی KSVD اما فراتر از این است. KSVD علاوه بر هر اتم، سطر متناظر آن در ماتریس ضرایب را نیز بروز می‌کند و بعلاوه در بروز کردن هر اتم تنها سیگنال‌هایی به کار می‌روند که قبلاً (در گام اول) از این اتم در نمایش خود استفاده کرده‌اند. برای توضیح بیشتر، اولاً دقت کنید که اگر در محاسبه‌ی ماتریس  $E_k$ ، از مقادیر به روز شده‌ی  $d_i$  و  $x_T^i$  ها استفاده کنیم، سرعت همگرایی الگوریتم بیش‌تر خواهد شد. به عنوان مثال، بعد از به روز کردن  $d_1$  و  $x_T^1$ ، برای بروز کردن اتم دوم و سطر متناظر آن، از مقادیر بروز شده‌ی  $d_1$  و  $x_T^1$  در محاسبه‌ی  $E$  استفاده کنیم و به همین ترتیب برای بقیه اتم‌ها و سطر متناظرشان. مسئله‌ی به روز کردن همزمان اتم  $d_k$  و  $x_T^k$  به صورت زیر است:

$$\min_{d, x_T} \left\| E_k - dx_T \right\|_F^2 \quad (6-2)$$

مسئله‌ی فوق در واقع رتبه-1 ماتریس  $E_k$  است. بطور کلی، اگر تجزیه به مقدار تکین (SVD) ماتریس دلخواه  $A \in \mathbb{R}^{n \times m}$  بصورت  $A = U \Sigma V^T$  باشد. تقریب رتبه- $r$  این ماتریس که  $r \leq \min(n, m)$  بصورت  $A_r = \sum_{i=1}^r \sigma_i u_i v_i^T$  است. مشکلی که در اینجا بوجود می‌آید این است که طی این فرآیند به احتمال زیاد تمام درایه‌های بردار  $x_T^k$  پر خواهد شد. به همین دلیل تنها از سیگنال‌هایی استفاده می‌شود که در گام قبلی از اتم  $d_k$  در نمایش خود استفاده کرده‌اند (به شباهت این کار با بروز کردن بردارهای کد در K-means) به این ترتیب تنها درایه‌های غیر صفر بردار  $x_T^k$  به روز شده و بقیه‌ی درایه‌های آن صفر می‌مانند. اگر تعریف کنیم  $w_k = \{i : 1 \leq i \leq m, x_T^k(i) \neq 0\}$ ، در نهایت مسئله‌ی زیر را برای بروز کردن اتم  $d_k$  و درایه‌های غیر صفر سطر  $x_T^k$  داریم:

$$\min_{d, x_r} \left\| E_k^{w_k} - dx_r \right\|_F^2 \quad s.t. \quad \|d\|_2 = 1 \quad (7-2)$$

که در آن  $E_k^{w_k}$  شامل تنها ستون‌های متناظر با  $w_k$  از  $E_k$  و نیز برداری به طول  $|w_k|$  است. اگر تجزیه به مقدار تکین  $E_k^{w_k}$  بصورت  $E_k^{w_k} = U \Sigma V^T$  باشد داریم:

$$d_k \leftarrow u_1, \quad x_T^k(w_k) \leftarrow \sigma_1 v_1 \quad (8-2)$$

که  $\sigma_1$  بزرگترین مقدار تکین است. توجه کنید که چون نرم ستون‌های  $U$  واحد است، بنابراین نیازی به نرمالیزه کردن  $d_k$  نیست. همه‌ی اتم‌های لغت‌نامه به این ترتیب به روز می‌شوند. وجه تسمیه KSVD نیز مشابه K-means به خاطر این است که به تعداد اتم‌ها  $K$  ( $m$ )، SVD انجام می‌شود [۱].

## ۴-۲ روش L1QP

روش L1QP که در [۷] از این روش استفاده شده، در الگوریتم ۳-۲ نشان داده شده است. ابتدا برای هر قطعه تصویر ضرایب تنک  $\beta$  در گام ۲ محاسبه می شود. در هر تکرار ضرایب تنک محاسبه می شود و هزینه بررسی می شود تا مشخص شود که این ضرایب جدید اضافه شود و یا حذف شود و این ضرایب بروزرسانی می شود تا زمانی که به وضعیت همگرایی برسد [۲۲].

---

الگوریتم ۳-۲ الگوریتم L1QP با ورودی های لغت نامه  $M$  و بردار ویژگی  $p$

---

**Require:** Dictionary  $M = FD_l$ , feature vector  $p = Fy$  for the solution of  $\min \frac{1}{2}x'Ax + b'x + \lambda|x|$  and  $b = y$ ,  $A = D_l$ ,  $x = \beta$

**Ensure:** Sparse prior  $\beta$

- 1: Find the max value  $v_{\max}$  and its index  $i_{\max}$  in  $abs(p)$
  - 2: Compute  $\beta(i_{\max})$  :  
 $if(p(i_{\max}) > \lambda) \quad \beta(i_{\max}) = \lambda - p(i_{\max})$   
 $else if(p(i_{\max}) < -\lambda) \quad \beta(i_{\max}) = -\lambda - p(i_{\max})$
  - 3: Find the idx of non-zero value in  $\beta$   
 $get \ subM = M(idx, idx), \ subp = p(idx), \ sub\beta = \beta(idx)$   
 Compute  $vect$  by  $subp$  and  $sub\beta$ ,  $\beta_{new}$  by  $subM$  and  $vect$   
 Check if all elements of  $\beta_{new} \times sub\beta > 0$  :  
 $Yes, update \ \beta$  by  $\beta_{new}$  and go step 5.  
 $No, go \ step \ 4.$
  - 4: Compute the cost for  $\beta_{new} \cdot sub\beta < 0$  and update  $\beta$ . Then go back step 3.
  - 5: Compute  $v = M\beta + p$ , find  $v_{\max}$  and  $i_{\max}$  in  $abs(v)$ .
  - 6: Check if  $v_{\max} < \lambda$ , if no, return step 2.
- 

## ۵-۲ الگوریتم KSVD تنک

بر اساس روشی که در [۲۳] پیشنهاد شده است، ایده ی اصلی الگوریتم KSVD تنک پیوند این دو لغت نامه و استفاده از مزایای هر دو آنهاست. لذا این الگوریتم ابتدا لغت نامه پایه ای همچون  $\phi \in \mathbb{R}^{N \times L}$  از انواع لغت نامه های ثابت، با توجه به داده ها، انتخاب کرده و سپس برای آموزش، اتم های لغت نامه دیگری همچون  $A \in \mathbb{R}^{L \times L}$  را تغییر می دهد. لغت نامه نهایی در این الگوریتم از ضرب این دو لغت نامه بدست می آید

$$\phi \in \mathbb{R}^{N \times L} = \phi \times A \quad (۹-۲)$$

از آنجا که لغت‌نامه  $\phi$  خود می‌تواند تا حدی داده‌ها را به صورت تُتک نمایش دهد، ماتریس  $A$  اتم‌های تُتک خواهد داشت. این امر مزیت بزرگ دیگری برای این الگوریتم به ارمغان می‌آورد. و آن این است که در قسمت آموزش لغت‌نامه  $A$  با اتم‌های تُتک مواجه است که موجب می‌شود بار محاسباتی کمی را برای محاسبه‌ی آن‌ها نیاز داشته باشد. کلیات الگوریتم آموزش لغت‌نامه تُتک مشابه الگوریتم KSVD است. این مدل آموزش لغت‌نامه ضمن تطابق زیادی که با داده‌ها پیدا می‌کند، آن طور که در بخش نمایش تُتک مضاعف توضیح داده خواهد شد، این ظرفیت را دارد تا با استفاده از آن، ابعاد مسئله را تا حد زیادی کاهش دهد و در نتیجه بار محاسباتی را به میزان چشمگیری کم کند.

## ۶-۲ الگوریتم بروز کردن موازی اتم‌ها

در [۲۰] الگوریتم بروز کردن موازی اتم‌ها پیشنهاد شده است. این الگوریتم کارا می‌تواند جایگزینی برای KSVD باشد. همانطور که پیش‌تر بیان شد، KSVD اتم‌ها را یک به یک و بطور متوالی بروز می‌کند. در اینجا طبق روشی که ارائه شده است، برای سادگی، سطر متناظر با هر اتم در ماتریس ضرایب را نمایه‌ی  $1$  آن اتم می‌نامیم، چرا که این سطر نشان می‌دهد چه داده‌هایی در نمایش تُتک خود از آن اتم استفاده کرده‌اند. مشکل اصلی الگوریتم KSVD حجم محاسبات زیاد آن (خصوصاً در ابعاد بالا) به دلیل محاسبه‌ی SVD ماتریس است. به عبارت دقیق‌تر، به عنوان تقریبی از جواب مسئله‌ی ۱۰-۲:

$$\min \|E_r - dx_r\|_F^2 \quad s.t. \quad \|d\|_2 = 1 \quad (10-2)$$

مقادیر به‌روز شده‌ی درایه‌های غیرصفر نمایه‌ی آن بصورت  $x_r = d^T E_r$  KSVD بدست می‌آید. اگر چه با انجام تعداد تناوب بیشتر برای حل مسئله‌ی ۱۰-۲ جواب‌های دقیق‌تری بدست می‌آید، اما باز هم عملکرد متوسط این روش بهتر از استفاده از SVD نخواهد بود. الگوریتم پیشنهادی برای حل مسئله‌ی ۱۰-۲ همان روش بهینه‌سازی تناوبی است، اما با شیوه‌ای متفاوت. برای توضیح این روش، فرض کنید برای بروز کردن هر اتم به همراه درایه‌های غیرصفر نمایه‌ی آن، تعداد  $K$  تناوب انجام دهیم. بعد از انجام این  $K$  تناوب، ماتریس  $A_i = d_i x_i^T$  (البته تنها ستون‌های متناظر با درایه‌های غیرصفر سطر  $x_i^T$ ، بقیه‌ی ستون‌های آن صفر باقی می‌مانند) به عنوان تقریبی رتبه-۱ از  $E_i^T$  بروز می‌شود. پیشنهاد در اینجا به این ترتیب است که به جای آنکه هر ماتریس  $A_i$  را جداگانه با انجام  $K$  تناوب بروز کنیم، این کار را برای همه‌ی این ماتریس‌ها بصورت موازی با هم انجام دهیم. به عبارت بهتر، در هر یک از  $K$  تناوب، همه‌ی ماتریس‌های  $A_i$  را موازی یکدیگر بروز می‌کنیم. توجه این کار به این ترتیب است که با توجه به رابطه‌ی زیر:

$$E = Y - DX = Y - (A_1 + A_2 + \dots + A_m) \quad (11-2)$$

در روش معمولی، برای بروز کردن به عنوان مثال  $A_3$ ، دو ماتریس  $A_1$  و  $A_2$  قبلاً بطور کامل (یعنی با انجام  $K$  تناوب) بروز

<sup>1</sup>Profile

شده‌اند. در حالیکه ماتریس‌های  $A_1, \dots, A_m$  هنوز مقادیر مربوط به گام ۱ آموزش لغت‌نامه را دارند و بعبارت دیگر اصلاً بروز نشده‌اند. در روش پیشنهادی، این  $K$  تناوب بصورت موازی برای همه‌ی این ماتریس‌ها انجام می‌شود. در نتیجه، هنگام بروز کردن هر ماتریس، بقیه‌ی ماتریس‌ها همگی تا حدی بروز شده‌اند. به همین دلیل این الگوریتم، آموزش لغت‌نامه با بروز کردن موازی اتم‌ها، نامیده شده‌است.

## ۷-۲ نمایش تُنک مضاعف

در [۲۳] الگوریتمی ارائه شده‌است که با کمک گرفتن از یک تقریب، سرعت یافتن نمایش تُنک را آنچنان که در تجربه‌ی شبیه‌سازی نشان داده‌است تا حدود پنج برابر افزایش می‌دهد. آن طور که در این بخش مطرح می‌شود، چون مسئله به مسئله‌ای با ابعاد بسیار کوچک‌تر تبدیل شده‌است زمان مورد نیاز برای حل مسئله نیز بسیار کمتر خواهد بود. همانطور که پیش‌تر بیان شد، در روش‌های اصلی نمایش تُنک از ماتریس  $D$  استفاده نموده و داده‌های ورودی  $X$  بر اساس آن نمایش داده می‌شود. یعنی باید بوسیله‌ی الگوریتم مسئله‌ی بهینه‌سازی ۱۲-۲ حل شود:

$$\min_{\alpha} \|X - D\alpha\|_F \quad s.t. \quad \|\alpha\|_0 \leq c \quad (12-2)$$

که در آن  $\alpha$  نمایش تُنک  $X$  است. این مسئله بار محاسباتی فراوانی دارد و در عوض در بیشتر کاربردها از معادل آن در شرایط خاص، یعنی معادله‌ی ۱۳-۲ استفاده می‌شود:

$$\min_{\alpha} \|X - D\alpha\|_F \quad s.t. \quad \|\alpha\|_1 \leq \varepsilon \quad (13-2)$$

این مدل بر اساس استفاده از مدل آموزشی لغت‌نامه تُنک بنا شده‌است. در مدل لغت‌نامه تُنک، ضرب ماتریسی دو ماتریس پایه  $\phi$  و تُنک  $A$  در حکم لغت‌نامه هستند. حال با فرض اینکه ماتریس  $A$  با داده‌های آموزشی آموزش دیده باشد و قصد طبقه‌بندی داده‌های ماتریس  $Y$  که همچون داده‌های آموزشی  $X$  هر ستون آن مبین یک نمونه است را داشته باشیم، می‌توان ماتریس تُنک  $\hat{Y}$  را بر اساس نمایش تُنک  $Y$  بر روی ماتریس  $\phi$  داشته و از آنجایی که خود ماتریس تُنک  $\hat{Y}$  قابلیت آن را دارد که بر روی ماتریس تُنک  $A$  تُنک‌تر نمایش داده شود خواهیم داشت:

$$Y = \phi \hat{Y} \approx \phi A \alpha \quad (14-2)$$

و یا برای یک نمونه از نمونه‌های آزمایشی خواهیم داشت:

$$y \approx \phi \hat{y} \approx \phi A \gamma \quad (15-2)$$

همانطور که گفته شد این دو با استفاده از الگوریتم‌های نمایش تُتک با در نظر گرفتن خطای دلخواه تقریب زده می‌شوند. بنابراین اگر این خطا را در مرحله‌ی نمایش تُتک بر روی  $\phi$  خطای  $e_\phi \in \mathbb{R}^{N \times L}$  و در مرحله‌ی نمایش تُتک بر روی ماتریس  $A$  خطای  $e_D \in \mathbb{R}^{N \times L}$  را داشته باشیم در این صورت

$$y = \phi \hat{y} + e_\phi = \phi A \gamma + e_D \quad (16-2)$$

ولذا

$$\|\phi(\hat{y} - A\gamma)\|_2 \leq (S+1)c \quad (17-2)$$

که در آن  $S$  مبین تعداد اعداد غیر صفر  $\gamma$  و یا در واقع همان نرم صفر آن است. سپس مسئله‌ی اصلی می‌تواند از طریق مسئله‌ی بسیار کوچکتری حل شود. برای این منظور مسئله‌ی بهینه‌سازی ۱۸-۲ را در نظر بگیرید:

$$\min_{\gamma} \|\hat{y} - A\gamma\|_2 \quad s.t. \quad \|\gamma\|_0 \leq \varepsilon' \quad (18-2)$$

اگرچه ماتریس  $A$  ابعادی بسیار بزرگتر از ماتریس  $D = \phi A$  دارد و شاید در نظر اول مسئله‌ی بزرگتر بنظر برسد، اما باید توجه داشت که ماتریس  $A$  تُتک است و از طرفی چون  $\hat{y}$  خود نیز تُتک است، می‌توان از بسیاری از ضرب‌ها در محاسبه‌ی این مسئله صرف نظر کرد و مسئله‌ی بسیار کوچک‌تر داشت. راه کار به این شکل است که ابتدا مکان مقادیر غیر صفر در هر ستون  $\hat{Y}$ ، یعنی هر  $\hat{y}$  پیدا می‌شود. سپس باید در کل  $A$  تنها بدنبال ستون‌هایی بود که حداقل یکی از ردیف‌های متناظر ردیف‌های غیر صفر ستون  $\hat{y}$ ، در آن نیز غیر صفر است. بدیهی است از آنجایی که  $\hat{y}$  باید جمع جبری این ستون‌ها باشد، ستون‌هایی که هیچ کدام از ردیف‌های غیر صفر  $\hat{y}$ ، آن‌ها غیر صفر نیست، هیچ گونه تاثیری نمی‌توانند در نمایش  $\hat{y}$  داشته باشند و قطعاً ضریب مرتبط با این ستون از  $A$  در  $\gamma$  متناظر صفر است. اگر از ستون‌های پیدا شده که ستون‌های مرتبط نامیده می‌شود، یک ماتریس جدید با نام  $A_R$  ساخته شود و متناظر آن  $\gamma_R$  و  $\hat{y}_R$  خواهیم داشت که مجموعاً مسئله‌ی بهینه‌سازی کوچک ابعاد ۱۹-۲ را می‌سازند:

$$\min_{\gamma_R} \|\hat{y}_R - A_R \gamma_R\|_2 \quad s.t. \quad \|\gamma_R\|_0 \leq \varepsilon' \quad (19-2)$$

بعد از حل مسئله‌ی جدید و یافتن  $\gamma_R$ ،  $\gamma$  اصلی براحتی با نداشت ردیف‌های  $\gamma_R$  به ردیف‌های متناسب که از مرحله‌ی قبل مشخص هستند، بدست می‌آید. حال با داشتن  $\gamma_R$  که با توجه به آموزش  $A$  تفکیکی هم شده است، با یکی از روش‌های طبقه‌بندی، بعنوان مثال روش SVM می‌توان نمونه‌های آزمایشی  $Y$  را طبقه‌بندی کرد. در این فصل روش‌های بهینه‌سازی تُنک مورد بحث قرار گرفت که پایه‌ی حل مسئله فراتفکیک پذیری با نمایش تُنک می‌باشد، این مسئله در فصل ۳ به تفصیل بیان خواهد شد.

## فصل ۳

# فرا تفکیک پذیری با نمایش تنگ

### ۱-۳ مقدمه

چون پایه‌ی پایان‌نامه درباره‌ی مبحث فرا تفکیک پذیری است در این فصل به طور جزئی‌تر بیان خواهد شد. در امر فرا تفکیک پذیری به عنوان مسئله‌ی معکوس، احیای تصویر با فرا تفکیک پذیری بالا از روی ترکیب تصاویر با تفکیک پذیری پایین مبتنی بر فرضیات منطقی و یا دانش اولیه درباره‌ی مدل مشاهداتی است که تصویر با تفکیک پذیری بالا را به تصویر تفکیک پذیری پایین متناظرش نگاشت می‌کند. مسئله‌ی بازسازی تصویر در فرا تفکیک پذیری به طور کلی به دلیل تعداد ناکافی از تصاویر با تفکیک پذیری پایین و عملگرهای محوکنندگی نامعلوم مسئله‌ای بد تعریف<sup>۲</sup> است و جواب مسئله‌ی بهینه‌سازی یکتا نیست.

در نمایش‌های تنگ سعی می‌شود که اطلاعاتی را که زوج قطعات نمونه در خود دارند، در دو لغت‌نامه، یکی لغت‌نامه‌ای که برای نمایش فضای با تفکیک پذیری بالا است و دیگری لغت‌نامه‌ای که برای نمایش فضای با تفکیک پذیری پایین است اطلاعات قطعات نمونه تا حد امکان حفظ شود. در این صورت به جای کار با صدها هزار زوج قطعه، می‌توانیم فقط با چند هزار اتم لغت‌نامه کار کنیم که سرعت روش‌های فرا تفکیک پذیری را به صورت قابل ملاحظه‌ای افزایش می‌دهد و در عین حال کیفیت روش‌ها چندان تغییر نمی‌کند. اولین مقاله‌ای که از نمایش تنگ برای فرا تفکیک پذیری استفاده کرد [۷] است که توسط یانگ و همکاران نوشته شده است و از این روش بنام روش یانگ نام برده خواهد شد. علی‌رغم تمام این تلاش‌ها روشی که یانگ در سال ۲۰۱۰ ارائه کرده است هم‌چنان به طور میانگین بهتر عمل می‌کند و در دیگر روش‌ها بهبود قابل توجهی نسبت به روش یانگ حاصل نمی‌شود. به همین علت هم‌چنان روش یانگ به عنوان معیاری برای بهترین روش موجود در مقالات برای مقایسه‌ها استفاده می‌شود. به همین خاطر در ادامه روش یانگ به تفصیل بیان خواهد شد.

در روش یانگ دو لغت‌نامه‌ی  $D_l$  برای نمایش تنگ قطعه‌های با تفکیک پذیری پایین و  $D_h$  برای نمایش قطعه‌های با تفکیک پذیری بالا آموزش داده می‌شود. هدف نهایی این است که لغت‌نامه‌ها طوری آموزش داده شوند که نمایش قطعه‌های

<sup>۲</sup>ill-posed

با تفکیک پذیری پایین روی  $D_l$  با نمایش قطعه‌های متناظر با تفکیک پذیری بالا روی  $D_h$  یکی باشد. بدین ترتیب می‌توان از روی نمایش یک قطعه با تفکیک پذیری پایین، نمایش قطعه‌ی متناظر با تفکیک پذیری بالا را در لغت‌نامه‌ی  $D_h$  و در نتیجه خود قطعه‌ی با تفکیک پذیری بالا بدست آورد. برای آموزش لغت‌نامه از روش ارائه شده در [۲۰] استفاده می‌شود. برای این که یکسان بودن نمایش تُتک قطعات با تفکیک پذیری بالا و پایین تحمیل شود، دو لغت‌نامه همزمان آموزش داده می‌شوند.

بنابراین به طور خلاصه الگوریتم دو مرحله دارد. در مرحله‌ی اول از روی نمونه‌ها دو لغت‌نامه طوری آموزش داده می‌شوند که نمایش قطعه‌های با تفکیک پذیری بالا و پایین روی لغت‌نامه‌های  $D_l$  و  $D_h$  یکی باشند. در مرحله دوم نمایش تُتک هر قطعه از تصویر با تفکیک پذیری پایین روی  $D_l$  به دست می‌آید و با ضرب این نمایش در  $D_h$  قطعه‌ی با تفکیک پذیری بالا به دست می‌آید. به عبارت دیگر لغت‌نامه‌ی  $D_h$  مستقیماً برای پیدا کردن نمایش تُتک استفاده نمی‌شود. بلکه برای بازسازی قطعه‌های با تفکیک پذیری بالا از روی نمایش تُتک قطعه‌های با تفکیک پذیری پایین روی  $D_l$  استفاده می‌شود. از آنجایی که در نمایش‌های تُتک، بسته به میزان خطای نمایش، تعداد اتم‌های مورد نیاز برای نمایش یک قطعه به طور وقفی تعیین می‌شود و این تعداد برای هر قطعه می‌تواند با دیگری متفاوت باشد. از آنجا که همواره از تعداد ثابتی نمونه برای نمایش قطعه‌ی مورد نظر استفاده نمی‌شود در نتیجه مشکل بیش برآزش و زیربرآزش وجود ندارد. این مطلب باعث می‌شود که کیفیت این روش بهتر از روش‌های پیشین باشد به علاوه نمایش‌های تُتک همان طور که در [۲۴] آمده‌است، تا حدودی نسبت به نویز مقاوم هستند. این مطلب باعث می‌شود روش‌های فراتفکیک پذیری مبتنی بر نمایش‌های تُتک مقاوم تر از سایر روش‌ها نسبت به نویز باشند و آن را تا حدودی حذف کنند، در حالی که اکثر روش‌های دیگر به نویز حساسند. در این روش به جای کار مستقیم با قطعه‌های با تفکیک پذیری بالا و پایین، نمایش فشرده‌ای از جفت قطعه‌ها آموزش داده می‌شود که سرعت الگوریتم را افزایش می‌دهد.

در ادامه‌ی فصل ابتدا روش‌یابنگ به تفصیل مورد بررسی قرار خواهد گرفت و مواردی چون فراتفکیک پذیری تصاویر از منظر تُتک بودن، فراتفکیک پذیری برای تصاویر عمومی، القای محدودیت بازسازی سراسری، بهینه‌سازی سراسری، فراتفکیک پذیری برای تصاویر صورت، انواع آموزش لغت‌نامه و نمایش ویژگی برای قطعه‌های با تفکیک پذیری پایین بیان خواهد شد. در ادامه، معیار شباهت ساختاری [۲۵] به منظور بهبود روش‌یابنگ مورد بررسی قرار خواهد گرفت.

## ۲-۳ روش‌یابنگ

### ۱-۲-۳ فراتفکیک پذیری تصاویر از منظر تُتک بودن

<sup>۱</sup> در مسئله‌ی فراتفکیک پذیری تصویر منفرد: تصویر  $Y$  با تفکیک پذیری پایین داده می‌شود، و تصویر  $X$  با تفکیک پذیری بالاتر از همان صحنه احیا می‌شود. دو نوع محدودیت برای حل این مسئله در این روش مدل شده‌است:

<sup>۱</sup>Image SR From Sparsity



(۱) محدودیت بازسازی: این گام نیازمند این است که تصویر احیا شده  $X$  بایستی با ورودی  $Y$  سازگار باشد. تصویر  $Y$  نسخه  $Y$  محو شده و زیرنمونه برداری شده  $^1$  از تصویر با تفکیک پذیری بالای  $X$  می باشد.

$$Y = SHX \quad (1-3)$$

در اینجا  $H$  نمایانگر فیلتر محوکنندگی و  $S$  عملگر زیرنمونه برداری می باشد.

(۲) فرض پیشین  $^2$  تُنک بودن: در این گام فرض بر این است که قطعه های با تفکیک پذیری بالا می توانند به صورت تُنک در یک لغت نامه فراکامل مناسب ارائه شوند و نمایش  $^3$  آن ها بر اساس مشاهدات تفکیک پذیری پایین بازیابی شود. قطعه های  $x$  از تصویر  $X$  می توانند با یک ترکیب خطی  $^4$  تُنک از لغت نامه  $D_h$  نمایش داده شوند که این لغت نامه توسط قطعه های با تفکیک پذیری بالای نمونه گیری شده، آموزش داده شده است.

$$x = D_h \alpha \quad \alpha \in \mathbb{R}^K \quad \|\alpha\|_0 \ll K \quad (2-3)$$

نمایش  $^5$  تُنک  $\alpha$  توسط قطعه های  $y$  از تصاویر ورودی  $Y$  بر اساس لغت نامه  $D_l$  که با لغت نامه  $D_h$  مقید شده است بازیابی می شود

در اینجا رهیافت  $^6$  تُنک هم برای تصاویر عمومی و هم تصاویر صورت به کار می رود. برای تصاویر عمومی مسئله به دو گام تقسیم می شود. در گام اول، همانطور که در رابطه ۳-۲ اشاره شد، نمایش  $^7$  تُنک برای هر قطعه محلی بر اساس سازگاری فاصله ای (فضایی) بین همسایه ها پیدا می شود، سپس نتایج بدست آمده از این مرحله توسط محدودیت بازسازی مطابق رابطه ۳-۱ تنظیم و تصحیح می شود. در این روش مدل محلی فرض پیشین  $^8$  تُنک  $^3$  برای احیای جزئیات با فرکانس بالای از دست رفته به کار می رود و در ادامه مدل سراسری محدودیت بازسازی برای حذف اثرات تصنعی  $^4$  های ممکن بر اثر اجرای گام اول اعمال می شود تا تصویری با ثبات بیشتر و طبیعی تر بدست آید. تصاویر صورت با تصاویر عمومی متفاوت است زیرا تصاویر صورت دارای ساختاری منظم تر هستند و از اینرو محدودیت بازسازی در زیرفضای صورت موثرتر است. در تصاویر صورت دو گامی که پیش تر بیان شد برعکس اعمال می شود تا استفاده ی بهتری از ساختار سراسری صورت به عنوان یک منظم ساز داشته باشد. در ابتدا زیرفضای مناسبی برای تصاویر صورت پیدا می شود و محدودیت بازسازی برای ایجاد تصویری با تفکیک پذیری متوسط اعمال می شود سپس جزئیات محلی با اعمال فرض پیشین  $^9$  تُنک بودن بر روی قطعه های تصویر احیا می شود.

---

<sup>1</sup>downsampled    <sup>2</sup>sparsity prior    <sup>3</sup>sparse prior    <sup>4</sup>artifact

### ۲-۲-۳ فراتفکیک پذیری برای تصاویر عمومی

مدل محلی از نمایش تُتک: این الگوریتم سعی دارد تا قطعه تصویر با تفکیک پذیری بالا را برای هر قطعه تصویر با تفکیک پذیری پایین از روی ورودی حدس بزند. در این مدل دو لغت نامه  $D_l$  و  $D_h$  وجود دارد که بگونه‌ای آموزش داده شده‌اند که برای هر جفت قطعه با تفکیک پذیری بالا و پایین تصویر دارای نمایش تُتک یکسان باشند. سپس میانگین مقادیر برای هر قطعه از آن کم می‌شود در نتیجه لغت نامه بیشتر نمایانگر ساختار تصویر نسبت به شدت روشنایی مطلق تصویر است. در فرآیند بازیابی میانگین مقادیر برای قطعه‌های با تفکیک پذیری بالا از روی قطعه‌های با تفکیک پذیری پایین تخمین زده می‌شود. برای هر قطعه با تفکیک پذیری پایین  $y$  نمایش تُتک با توجه به  $D_l$  پیدا می‌شود  $D_h$  با ضرایب بدست آمده ترکیب می‌شود و قطعه‌های با تفکیک پذیری بالا را تولید می‌کند. مسئله‌ی پیدا کردن تُتک ترین نمایش از  $y$  بصورت زیر فرمول بندی می‌شود:

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon \quad (3-3)$$

که  $F$  بیانگر عملگر استخراج ویژگی (خطی) می‌باشد. نقش اصلی  $F$  در رابطه ۳-۳ این است که شرایطی را فراهم می‌کند که چقدر  $\alpha$  باید نزدیک باشد تا  $y$  را تخمین بزند از آنجا که مسئله‌ی بهینه‌سازی رابطه‌ی ۳-۳ مسئله‌ی NP-Hard است می‌توان از نرم  $L_1$  به جای نرم  $L_0$  مطابق رابطه‌ی زیر استفاده کرد. (این رابطه ریلکس شده‌ی رابطه‌ی قبل با نرم صفر است (لاسو))

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon \quad (4-3)$$

و به طور معادل با ضرایب لاگرانژ رابطه‌ی زیر را داریم:

$$\min_{\alpha} \|FD_l\alpha - Fy\|_2^2 + \lambda \|\alpha\|_1 \quad (5-3)$$

در رابطه‌ی بالا  $\lambda$  پارامتری برای تعیین میزان تُتک بودن جواب و صحت و درستی تخمین  $y$  می‌باشد. رابطه‌ی ۵-۳ سازگاری بین قطعه‌های مجاور را تضمین نمی‌کند به این منظور رابطه‌ی ۴-۳ به صورت زیر اصلاح می‌شود و رابطه‌ی  $D_h\alpha$  برای ایجاد سازگاری قطعه‌های مجاور به شرایط اضافه می‌شود. قطعه‌ها از چپ به راست و از بالا به پایین مورد بررسی قرار

می گیرند.

$$\min_{\alpha} \|\alpha\|_1 \quad s.t. \quad \|FD_l\alpha - Fy\|_2^2 \leq \varepsilon_1 \quad (6-3)$$

$$\|PD_h\alpha - w\|_2^2 \leq \varepsilon_2$$

در رابطه‌ی ۶-۳ ماتریس  $P$  ناحیه‌ی همپوشانی بین قطعه‌هدف کنونی و تصویر با تفکیک پذیری بالای بازسازی شده‌ی قبلی را استخراج می‌کند و  $w$  حاوی مقادیر تصویر با تفکیک پذیری بالای بازسازی شده‌ی قبلی در همپوشانی می‌باشد. مسئله‌ی بهینه‌سازی قبل به طور مشابه به صورت زیر فرمول نویسی می‌شود:

$$\min_{\alpha} \left\| \tilde{D}\alpha - \tilde{y} \right\|_2^2 + \lambda \|\alpha\|_1 \quad (7-3)$$

در اینجا  $\tilde{y} = \begin{bmatrix} Fy \\ \beta w \end{bmatrix}$  و  $\tilde{D} = \begin{bmatrix} FD_l \\ \beta PD_h \end{bmatrix}$  می‌باشد. پارامتر  $\beta$  میزان تعادل بین تخمین  $y$  و پیدا کردن قطعه‌با تفکیک پذیری بالایی که با همسایه هایش سازگار است را تعیین می‌کند. در بیشتر عملیات  $\beta = 1$  در نظر گرفته می‌شود. در ادامه قطعه‌با تفکیک پذیری بالا به صورت  $x = D_h\alpha^*$  محاسبه می‌شود که  $\alpha^*$  جواب بهینه رابطه ۷-۳ می‌باشد.

### ۳-۲-۳ القای محدودیت بازسازی سراسری

بایستی توجه شود که رابطه‌ی ۴-۳ و ۶-۳ کیفیت عینی بین قطعه‌های با تفکیک پذیری پایین  $y$  و  $D_l\alpha$  ایجاد نمی‌کند. به این دلیل و همچنین وجود نویز،  $X_0$  ای که در مرحله‌ی قبل بدست آمد نمی‌تواند شرایط بازسازی رابطه‌ی ۱-۳ را ارضا کند. برای حذف این ناسازگاری و اختلاف  $X_0$  را روی فضای جواب  $SHX = Y$  تصویر<sup>۱</sup> می‌کنیم. رابطه‌ی زیر را در نظر بگیرد:

$$X^* = \arg \min_x \|SHX - Y\|_2^2 + c \|X - X_0\|_2^2 \quad (8-3)$$

برای حل این مسئله‌ی بهینه‌سازی می‌توان از روش گرادیان کاهشی<sup>۲</sup> استفاده کرد. معادله‌ی بروزرسانی برای این متد تکراری به صورت زیر است:

$$X_{t+1} = X_t + \nu [H^T S^T (Y - SHX_t) + c(X - X_0)] \quad (9-3)$$

<sup>۱</sup> Project      <sup>۲</sup> gradient descent

که  $X_t$  تخمینی از تصویر با تفکیک پذیری بالا بعد از تکرار  $t$  ام و  $\nu$  اندازه‌ی گام در الگوریتم ۱-۳ می‌باشد.  $X^*$  جواب نهایی حل مسئله‌ی بهینه‌سازی قبل برای تخمین تصویر با تفکیک پذیری بالا است. این جواب تا حد ممکن به جواب اولیه‌ی  $X$  نزدیک است که  $X$  جواب بدست آمده از تئنگ بودن با توجه به محدودیت بازسازی است.

فرآیند فراتفکیک پذیری بصورت الگوریتم ۱-۳ خلاصه می‌شود:

---

### الگوریتم ۱-۳ فراتفکیک پذیری برای تصاویر طبیعی

---

**Require:** training dictionaries  $D_h$  and  $D_l$ , a low-resolution image  $Y$

**Ensure:** SR image  $X^*$

- 1: **for** each  $3 \times 3$  patch  $y$  of  $Y$ , taken starting from the upper-left corner with 1 pixel overlap in each direction, **do**
  - 2:   compute the mean pixel value  $m$  of patch  $y$
  - 3:   Solve the optimization problem with  $\tilde{D}$  and  $\tilde{y}$  defined in 8-3
  - 4:   Generate the high-resolution patch  $x = D_h \alpha^*$ . Put the patch  $x + m$  into a high-resolution image  $X_0$ .
  - 5: **end for**
  - 6: Using gradient descent, find the closest image to  $X_0$  which satisfies the reconstruction constraint defined in 9-3
- 

### ۴-۲-۳ شرح بهینه‌سازی سراسری

آنچه که در مبحث های قبل بیان شد الگوریتم ساده ی فراتفکیک پذیری بود که در بسیاری از کاربردهای مسائل معکوس در پردازش تصویر کاربرد دارد. رابطه‌ی زیر با اینکه خیلی تئنگ نیست اما محدودیت بازسازی را بهتر ارضا می‌کند و رابطه‌ی قبل تبدیل به رابطه‌ی بهینه‌سازی طولانی زیر می‌شود:

$$X^* = \arg \min_{X, \{\alpha_{i,j}\}} \{ \|SHX - Y\|_2^2 + \lambda \sum_{i,j} \|\alpha_{i,j}\| + \gamma \sum_{i,j} \|D_h \alpha_{i,j} - P_{i,j} X\|_2^2 + \tau \rho(X) \} \quad (10-3)$$

در اینجا  $\alpha_{i,j}$  نمایانگر ضرایب نمایش تئنگ برای قطعه  $X_{i,j}$  ام است و  $P_{i,j}$  یک ماتریس تصویر<sup>۱</sup> ای است که قطعه  $i,j$  ام از  $X$  را انتخاب می‌کند.  $\rho(X)$  تابع جریمه ای است که اطلاعات اولیه از تصویر با تفکیک پذیری بالا را می‌دهد.

الگوریتم ۱-۳ می‌تواند به عنوان تخمینی کارا از رابطه‌ی ۱۰-۳ تعبیر شود. گام نمایش تئنگ، ضرایب  $\alpha$  را بر اساس ترم دوم و سوم از رابطه‌ی ۱۰-۳ بدست می‌آورد. ترم  $\|\alpha_{i,j}\|$  به صورت  $\|\alpha\|_1$  ریلکس می‌شود. و ترم  $\|D_h \alpha_{i,j} - P_{i,j} X\|_2$  بوسیله ی ترم  $\|FD_l \alpha_{i,j} - Fy_{i,j}\|_2$  تخمین زده می‌شود.

---

<sup>۱</sup>projection matrix

بایستی توجه شود که اگر ضرایب تُنک  $\alpha$  ثابت در نظر گرفته شود، ترم سوم به عنوان جریمه تفاوت بین  $X$  و بازسازی شده‌ی آن با ضرایب تُنک را بر اساس رابطه  $\|X_0 - X\|_F^2 \approx \sum_{i,j} \|D_h \alpha_{i,j} - P_{ij} X\|_F^2$  بیان می‌دارد.

### ۵-۲-۳ فراتفکیک پذیری برای تصاویر صورت

تصاویر صورت برخلاف تصاویر عمومی از نظر ساختمان منظم‌تر هستند از اینرو بکارگیری آنها آسان‌تر است. در واقع برای تصاویر صورت ما می‌توانیم با ورودی‌های با تفکیک پذیری کمتری سروکار داشته باشیم. ایده‌ی اصلی به این صورت است که در ابتدا از اطلاعات اولیه‌ی چهره برای زوم کردن و ایجاد یک تصویر با تفکیک پذیری متوسط بهره می‌گیرد. در ادامه از فرض پیشین تُنک محلی<sup>۱</sup> برای احیای جزئیات استفاده می‌کند. اگر دقیق‌تر بیان  $a, n$  جواب از دو رویکرد بدست می‌آید (۱) مدل سراسری: استفاده از محدودیت بازسازی برای ایجاد تصویر با تفکیک پذیری متوسط، در حالیکه جواب فقط در زیرفضای صورت جست و جو می‌شود.

(۲) مدل محلی: استفاده از مدل محلی تُنک برای احیای جزئیات تصویر

### ۶-۲-۳ آموزش لغت‌نامه برای فراتفکیک پذیری

دو مسئله‌ی آموزش لغت‌نامه را نیز می‌توان برای ساده‌تر شدن به صورت زیر نوشت:

$$\min_{D, \{w_t\}_{t=1}^N} \sum_{t=1}^N (\|x_t - Dw_t\|_F^2 + \lambda \|w_t\|_1) \quad (11-3)$$

که در آن جملات نرم صفر با اضافه کردن هزینه برای تُنک نبودن نمایش سیگنال‌ها، نمایش‌ها را به سمت تُنک بودن سوق می‌دهند. اگر بردارهای  $x_i$  را کنار هم بچینیم تا ماتریس  $X$  به دست آید و ماتریس  $W$  ماتریسی باشد که از کنار هم چیدن بردارهای نمایش تُنک  $w_i$  به دست آمده‌است. این رابطه به فرم ماتریسی زیر در می‌آید:

$$\min_{D, W} \|X - DW\|_F^2 + \lambda \|W\|_1 \quad (12-3)$$

این رابطه را با استفاده از نرم  $l_1$  بجای نرم صفر به صورت زیر نوشت:

$$\min_{D, W} \|X - DW\|_F^2 + \lambda \|W\|_1 \quad (13-3)$$

حال چنانچه فرض شود به جای یک لغت‌نامه دو لغت‌نامه برای نمایش قطعات با تفکیک پذیری بالا و پایین آموزش داده شود. قطعه‌های  $\{x_{HR_i}\}_{i=1}^N$  با فراتفکیک پذیری بالا را در ماتریس  $X_{HR}$  کنار هم و قطعه‌های  $\{x_{LR_i}\}_{i=1}^N$  با تفکیک

<sup>۱</sup>prior sparsity local

پذیری پایین را در ماتریس  $X_{LR}$  کنار هم چیده می‌شود. مسئله‌ی آموزش دو لغت‌نامه برای نمایش این دو دسته سیگنال را می‌توان به صورت زیر نوشت:

$$\min_{D_h, W_1} \|X_{HR} - D_h W_1\|_F^2 + \lambda_1 \|W_1\|_1 \quad (14-3)$$

$$\min_{D_l, W_2} \|X_{LR} - D_l W_2\|_F^2 + \lambda_2 \|W_2\|_1 \quad (15-3)$$

حل جداگانه این دو مسئله معادل با حل مسئله‌ی زیر است:

$$\min_{D, W_1, W_2} \|X_{HR} - D_h W_1\|_F^2 + \lambda_1 \|W_1\|_1 + \|X_{LR} - D_l W_2\|_F^2 + \lambda_2 \|W_2\|_1 \quad (16-3)$$

از آنجایی که هدف آن است که ضرایب نمایش تنگ در هر دو لغت‌نامه یکی باشند،  $W_1 = W_2$  در نظر گرفته شود. در نتیجه رابطه‌ی زیر برای آموزش توام دو لغت‌نامه برای آنکه ضرایب نمایش تنگ در آن‌ها یکسان باشد بدست می‌آید:

$$\min_{D, W} \|X_{HR} - D_h W\|_F^2 + \alpha \|X_{LR} - D_l W\|_F^2 + \lambda \|W\|_1 \quad (17-3)$$

که در آن  $\alpha$  ضریبی است که برای مصالحه بین میزان خطای نمایش قطعات با تفکیک پذیری بالا  $X_{HR}$  و قطعات با تفکیک پذیری پایین  $X_{LR}$  به کار می‌رود.

برای استفاده از روش‌های آموزش لغت‌نامه باید شکل این مسئله‌ی آموزش توام بایستی به شکل مسئله‌ی آموزش یک لغت‌نامه در آورده شود. اگر بردار  $x_i$  برداری در نظر گرفته شود که از زیر هم گذاشتن بردارهای  $x_{LR_i}$  و  $x_{HR_i}$  به دست آمده است، فضایی ایجاد شود که بعد آن مجموع بعدهای دو فضای قطعات با تفکیک پذیری بالا و پایین است. برای آنکه در این فضا، هر دو زیرفضا اهمیت یکسانی داشته باشند، قطعه‌های با تفکیک پذیری بالا و پایین به جذر بعدشان تقسیم شده و سپس زیر هم چیده می‌شود تا فضای جدید به دست آید:

$$x_i = \begin{bmatrix} \frac{x_{LR_i}}{\sqrt{d}} \\ \frac{x_{HR_i}}{\sqrt{m}} \end{bmatrix} \quad i = 1, 2, \dots, N \quad (18-3)$$

که در آن  $d$  بعد فضای قطعه‌های با تفکیک پذیری پایین  $m$  بعد فضای قطعه‌های با تفکیک پذیری بالا است. به همین ترتیب اگر تعریف شود:

$$D = \begin{bmatrix} D_l \\ D_h \end{bmatrix} \quad (19-3)$$

به راحتی می توان دید که مسئله آموزش لغت نامه که در رابطه ۳-۱۷ برقرار بود، با قرار دادن  $\alpha = 1$  به مسئله زیر تبدیل می شود:

$$\min_{D, W} \|X - DW\|_F^2 + \lambda \|W\|_1 \quad (20-3)$$

که در آن  $X$  ماتریسی است که از کنار هم چیدن  $x$  ها به دست آمده است. که رابطه ۳-۲۰ دقیقاً همان رابطه آموزش یک لغت نامه است که در رابطه ۳-۱۳ بیان شد. بنابراین مسئله آموزش توام دو لغت نامه به سادگی به مسئله آموزش یک لغت نامه قابل تبدیل است.

پس از حل مسئله آموزش لغت نامه، حل مسئله فراتفکیک پذیری با استفاده از این لغت نامه ها بسیار ساده است. ابتدا تصویر با تفکیک پذیری پایین به قطعه هایی با مقداری هم پوشانی تقسیم می شود:  $\{y_{LR_i}\}_{i=1}^M$ . سپس نمایش تنگ هر قطعه را روی  $D_l$  به دست آورده می شود:  $y_{LR_i} = D_l w_i$ . سپس هر قطعه ی با تفکیک پذیری بالا را از روی این ضرایب و لغت نامه ی  $D_h$  از رابطه ی زیر به دست آورده می شود:

$$y_{HR_i} = D_h w_i \quad i = 1, 2, \dots, M \quad (21-3)$$

### ۷-۲-۳ آموزش لغت نامه

در بخش قبل درباره ی منظم سازی مسئله فراتفکیک پذیری با استفاده از فرض پیشین تنگ بودن که هر جفت قطعه با فراتفکیک پذیری بالا و پایین دارای نمایش تنگ یکسان متناظر با لغت نامه  $D_l$  و  $D_h$  هستند. اما از آنجا که لغت نامه بدست آمده بزرگ است و به زمان محاسباتی بالایی نیاز دارد در این بخش بیشتر تمرکز روی آموزش جفت لغت نامه فشرده تر است تا سرعت محاسبات بالا رود.

### ۸-۲-۳ آموزش لغت نامه منفرد

کدینگ تنگ مسئله ی یافتن نمایش تنگ برای سیگنال مطابق با لغت نامه فواکامل  $D$  است. این لغت نامه معمولاً با مجموعه نمونه های  $X = \{x_1, x_2, \dots, x_t\}$  آموزش داده می شود. بطور کلی آموزش لغت نامه فشرده امری دشوار است، بگونه ای

که تضمین کند نمایش تُنک از رابطه‌ی ۳-۳ می‌تواند از مینیم سازی نرم  $L_1$  در رابطه‌ی ۳-۴ بازسازی شود. رابطه‌ی اصلی ساخت لغت‌نامه در این روش مطابق فرمول زیر است:

$$D = \arg \min_{D,Z} \|X - DZ\|_1 + \lambda \|Z\|_1 \quad s.t. \quad \|D_i\|_1 \leq 1 \quad (22-3)$$

در اینجا  $\|Z\|_1$  خاصیت تُنک بودن را القا می‌کند و محدودیت نرم  $L_2$  روی ستون‌های  $D$  مبهم بودن بر اثر اسکیل را کم می‌کند. (در  $\|D_i\|_1 \leq 1$ ، از آنجا که لغت‌نامه نرمال است تضمین می‌کند که نرم آن کمتر از یک باشد. رابطه‌ی ۳-۲۲ محدب نیست اما اگر فقط روی  $D$  و یا روی  $Z$  باشد محدب می‌شود.

۱. ابتدا لغت‌نامه اولیه با ماتریس گوسی تصادفی و با هر ستون نرمال شده مقدار دهی می‌شود (همان لغت‌نامه اولیه  $B$  است در کد)

۲. با ثابت در نظر گرفتن  $D$ ، و رابطه‌ی زیر  $Z$  بروزسانی می‌شود

$$Z = \arg \min_Z \|X - DZ\|_1 + \lambda \|Z\|_1 \quad (23-3)$$

این رابطه با برنامه ریزی خطی حل می‌شود

۳. با ثابت در نظر گرفتن  $Z$ ، و رابطه‌ی زیر  $D$  بروزسانی می‌شود

$$D = \arg \min_D \|X - DZ\|_1 \quad s.t. \quad \|D_i\|_1 \leq 1 \quad (24-3)$$

۴. تکرار بین مرحله ۲ و ۳ تا زمانی که همگرا شود.

### الگوریتم ۲-۳ الگوریتم آموزش لغت‌نامه

**Require:** Initialize  $D$  with a Gaussian random matrix, with each column unit normalized

- 1: Fix  $D$ , update  $Z$  by  $Z = \arg \min_Z \|X - DZ\|_1 + \lambda \|Z\|_1$  which can be solved efficiently through linear programming.
- 2: Fix  $Z$ , update  $D$  by  $D = \arg \min_D \|X - DZ\|_1 \quad s.t. \quad \|D_i\|_1 \leq 1$  which is a Quadratically constraint Quadratic Programming that is ready to be solved in many optimization packages
- 3: Iterate between 1) and 2) until converge



### ۹-۲-۳ آموزش جفت لغت‌نامه

فرض بر این است که جفت قطعه‌های آموزشی  $P = \{X^h, Y^l\}$  که  $X^h = \{x_1, x_2, \dots, x_n\}$  مجموعه قطعه‌های نمونه با تفکیک پذیری بالا و  $Y^l = \{y_1, y_2, \dots, y_n\}$  قطعه‌های متناظر با تفکیک پذیری پایین می‌باشد. هدف در اینجا آموزش لغت‌نامه برای قطعه‌های با تفکیک پذیری بالا و پایین است. از اینرو نمایش تنگ برای قطعه‌های با تفکیک پذیری بالا و پایین مشابه است که مسئله‌ای دشوار است. فرمول بندی برای هر لغت‌نامه بصورت جدا به صورت زیر است:

$$D_h = \arg \min_{\{D_h, Z\}} \left\| X^h - D_h Z \right\|_F^2 + \lambda \|Z\|_1 \quad (25-3)$$

$$D_l = \arg \min_{\{D_l, Z\}} \left\| Y^l - D_l Z \right\|_F^2 + \lambda \|Z\|_1 \quad (26-3)$$

با ترکیب این دو تابع هدف و القای نمایش تنگ یکسان برای هر دو، رابطه‌ی زیر را داریم:

$$\min_{\{D_h, D_l, Z\}} \frac{1}{N} \left\| X^h - D_h Z \right\|_F^2 + \frac{1}{M} \left\| Y^l - D_l Z \right\|_F^2 + \lambda \left( \frac{1}{N} + \frac{1}{M} \right) \|Z\|_1 \quad (27-3)$$

که  $N$  و  $M$  بعد قطعه‌تصویر با تفکیک پذیری بالا و پایین هستند. ضرایب  $\frac{1}{N}$ ،  $\frac{1}{M}$  تعادل بین رابطه‌ی ۲۵-۳ و ۲۶-۳ برقرار می‌کند. رابطه ۲۷-۳ می‌تواند بصورت رابطه ۲۸-۳ نوشته شود.

$$\min_{\{D_h, D_l, Z\}} \left\| X_c - D_c Z \right\|_F^2 + \lambda + \left( \frac{1}{N} + \frac{1}{M} \right) \|Z\|_1 \quad (28-3)$$

بطور معادل داریم:

$$\min_{\{D_h, D_l, Z\}} \left\| X_c - D_c Z \right\|_F^2 + \hat{\lambda} \|Z\|_1 \quad (29-3)$$

بگونه‌ای که:

$$X_c = \begin{bmatrix} \frac{1}{\sqrt{N}} X^h \\ \frac{1}{\sqrt{M}} Y^l \end{bmatrix}, \quad D_c = \begin{bmatrix} \frac{1}{\sqrt{N}} D_h \\ \frac{1}{\sqrt{M}} D_l \end{bmatrix} \quad (30-3)$$

در اینجا شیوه ای مشابه لغت نامه منفرد برای آموزش جفت لغت نامه بکار می رود. از آنجا که در قطعه های با تفکیک پذیری پایین از ویژگی ها استفاده می کنیم  $D_l$  و  $D_h$  رابطه ی خطی با هم ندارند. لغت نامه حاصل نشان دهنده ی الگوی پایه قطعه تصاویر از جمله لبه های جهت دار به جای الگوی قطعه های افقی برای فشردگی می باشد.

### ۱۰-۲-۳ نمایش ویژگی برای قطعه های با تفکیک پذیری پایین

در رابطه ۳-۳ از  $F$  (تبدیل ویژگی) استفاده شد تا تضمین کند که ضرایب محاسبه شده بیشترین تطابق را برای سیگنال با تفکیک پذیری پایین دارد. از اینرو تخمین درستی برای بازسازی قطعه های با تفکیک پذیری بالا خواهیم داشت. ( $F$  نوعی فیلتر بالاگذر است). در این روش از مشتقات مرتبه اول و دوم به عنوان ویژگی استفاده شده است. برای فیلترهای یک بعدی بصورت زیر است:

$$\begin{aligned} f_1 &= [-1, 0, 1] \quad , \quad f_2 = f_1^T \\ f_3 &= [1, 0, -2, 0, 1] \quad , \quad f_4 = f_3^T \end{aligned} \quad (3-31)$$

این ۴ فیلتر برای هر قطعه با تفکیک پذیری پایین اعمال می شود (برای همین  $4 \times 25$  اندازه ۱۰۰ است در تفکیک پذیری پایین) که این ۴ قطعه با هم ترکیب شده و در نهایت به عنوان یک بردار به عنوان قطعه با تفکیک پذیری پایین بازگردانده می شود. لازم به ذکر است که این ۴ فیلتر روی هر قطعه اعمال نمی شود بلکه روی کل تصویر اعمال می شود. نگاهت آن فیلتر ساخته می شود و قطعه از آن انتخاب می شود. بنابراین این اعمال فیلتر اطلاعات همسایگی را در تصویر با تفکیک پذیری پایین کد می کند که برای ایجاد سازگاری بین قطعه های مجاور در تصویر نهایی سودمند است. در عمل اعمال فیلتر و استخراج ویژگی روی تصویر بزرگ شده نتایج بهتری داشته است از اینرو ابتدا تصویر را دوبرابر شده و سپس ویژگی ها را استخراج می کنیم. بدلیل طریقه ای که ویژگی ها استخراج می شود  $D_l$  و  $D_h$  رابطه خطی ندارند.

### ۱۱-۲-۳ تاثیر محدودیت سراسری

محدودیت بازسازی سراسری که در رابطه ی ۳-۸ القاشد برای تصحیح کردن مدل محلی است و اطمینان حاصل می کند که تصویر با تفکیک پذیری بالا با مشاهدات تفکیک پذیری پایین سازگار است. در عمل دیده می شود که فرض پیشین  $\hat{t}_{nk}$  بودن بسیار کارا است و بیشترین مشارکت را دارد.

در این فصل روش یانگ به تفصیل مورد بحث قرار گرفت و مشاهده شد که از معیار میانگین مربعات خطا استفاده می شود. در ادامه استفاده از معیار شباهت ساختاری بجای میانگین مربعات خطا بررسی خواهد شد.

## ۳-۳ بررسی معیار شباهت ساختاری بجای میانگین مربعات خطا

در روش بکار رفته در [۲۵] از معیار شباهت ساختاری تصاویر<sup>۱</sup> بجای معیار میانگین مربعات خطا<sup>۲</sup> در مسئله تکمیل تصویر<sup>۳</sup> با نمایش تُنگ استفاده شده است. بکارگیری این معیار در دو مرحله تخمین ضرایب نمایش تُنگ و برورسانی لغت نامه انجام شده است. بیشتر روش های قبلی موجود برای مسئله تکمیل تصویر مبتنی بر تقریب حداقل مربعات یعنی مینم کردن معیار میانگین مربعات خطا برای شدت روشنایی بوده است. معیار میانگین مربعات خطا که شناخته شده ترین متریک می باشد نمی تواند کیفیت بصری بالایی را فراهم کند بنابراین این معیار نمی تواند معیار مناسبی برای اندازه گیری کیفیت در مسئله تکمیل تصویر باشد. استفاده از شاخص<sup>۴</sup> شباهت ساختاری است که یکی از نمایانگرترین معیار اندازه گیری شباهت در بیشتر زمینه های پردازش تصویر می باشد. از آنجا که فرمول بندی آن برای تجزیه و تحلیل ساده و راحت است و برای ارزیابی تصویر می تواند بکار رود. بایستی توجه کرد از آنجا که در این رهیافت شاخص معیار شباهت ساختاری ماکسیمم می شود و این مسئله غیر محدب است، رابطه بصورت شبه محدب<sup>۵</sup> فرمول بندی می شود. بزرگترین تفاوت بین این روش و روش های قبلی در مرحله ساخت لغت نامه می باشد. نتایج نشان داده شده در این روش بیانگر این است که استفاده از معیار شباهت ساختاری نتایج بهتری دارد.

در روش بکارگرفته شده در [۲۶] از معیار شباهت ساختاری در مسئله فراتفکیک پذیری استفاده شده است. در این روش نشان داده شده است که نمایش تُنگ در مینم کردن نرم L2 به آسانی می تواند به ماکسیمم کردن معیار شباهت ساختاری تبدیل شود. فرض کنید رابطه مینیم سازی نمایش تُنگ محلی بصورت زیر باشد:

$$\hat{\alpha}_{ij} = \arg \min_{\alpha} \mu_{ij} \|\alpha\|_0 + \|D\alpha - F_{ij}X\|_p^2 \quad (32-3)$$

که در این رابطه  $X$  تصویر بازسازی شده خروجی است،  $F_{ij}$  ماتریسی است که بلاک  $ij$  ام را از تصویر استخراج می کند و  $D \in R^{n \times k}$  معرف لغت نامه و  $\alpha_{ij}$  و  $\mu_{ij}$  بیانگر میانگین بلاک  $ij$  از تصویر و بردار ضرایب تُنگ متناظر با بلاک  $ij$  از تصویر است. بطور معادل داریم:

$$\hat{\alpha}_{ij} = \arg \min_{\alpha} \|\alpha\|_0 \quad s.t. \quad \|D\alpha - F_{ij}X\| \leq T \quad (33-3)$$

با بکارگیری از معیار شباهت ساختاری رابطه قبل بصورت زیر تبدیل می شود:

$$\hat{\alpha}_{ij} = \arg \min_{\alpha} \mu_{ij} \|\alpha\|_0 + (1 - S(D\alpha, F_{ij}X)) \quad (34-3)$$

<sup>۱</sup>SSIM

<sup>۲</sup>MSE

<sup>۳</sup> Image Inpainting

<sup>۴</sup>index

<sup>۵</sup>quasi-convex

که  $S(., .)$  معرف اندازه گیری معیار شباهت ساختاری است و تعریف ریاضی این شاخص بصورت زیر است:

$$S(a, y) = \frac{\mu_a \mu_y + C_1}{\mu_a^2 + \mu_y^2 + C_1} \frac{\sigma_{a,y} + C_2}{\sigma_a^2 + \sigma_y^2 + C_2} \quad (35-3)$$

که در این رابطه  $\mu_a$  و  $\mu_y$  معرف میانگین  $a$  و  $y$  هستند. همچنین  $\sigma_a^2$  و  $\sigma_y^2$  بیانگر واریانس  $a$  و  $y$  هستند. و  $\sigma_{a,y}$  بیانگر کواریانس بین  $a$  و  $y$  است. ثابت های  $C_1, C_2$  هم ثابت های پایدارکننده و شمارشگری برای تاثیر اشباع در HSV است. بکارگیری معیار شباهت ساختاری در مقایسه با معیار میانگین مربعات خطا برای این مسائل بسیار مناسب است چرا که با بکارگیری از فضای رنگی HSV اجازه می دهد تا دستیابی به جزییات ساختاری در تصاویر نویزی بهبود یابد. همچنین حدآستانه ( پارامتر  $T$  در رابطه ی ۳-۳۳ ) برای توقف الگوریتم با این معیار، مشابه پارامتر  $T_{mse}$  است و با  $T_{sim}$  نمایش داده می شود که از رابطه ی زیر محاسبه می شود:

$$T_{sim} = 1 - \frac{T_{mse}}{\sigma_a^2 + \sigma_y^2 + C_2} \quad (36-3)$$

این رابطه نشان می دهد که تفاوت اصلی بین معیار شباهت ساختاری و میانگین مربعات خطا نرمال سازی مجزا است. فرض کنید مشابه روش یانگ ابتدا تصویر با تفکیک پذیری پایین از روی تصویر با تفکیک پذیری بالا طبق رابطه زیر تولید می شود:

$$Y = SHX \quad (37-3)$$

که بطور مشابه با مسئله ی فراتفکیک پذیری  $H$  ماتریس محوکنندگی و  $S$  ماتریس زیرنمونه برداری است. در اینجا از دو لغت نامه،  $D_l$  لغت نامه با فراتفکیک پذیری بالا و  $D_h$  لغت نامه با تفکیک پذیری پایین استفاده می شود. چنانچه  $y$  نمایانگر تصویر با تفکیک پذیری پایین و  $x$  نمایانگر تصویر با تفکیک پذیری بالا باشد روابط زیر برقرار است.

$$y = D_l \alpha \quad (38-3)$$

$$x = D_h \alpha \quad (39-3)$$

هر قطعه از تصویر با تفکیک پذیری پایین نیاز است که در ابتدا بزرگ شود و سپس با الگوریتم ۳-۳<sup>۱</sup> استخراج و بصورت

<sup>۱</sup> SSIM-inspired OMP

### الگوریتم ۳-۳ الگوریتم معیار شباهت ساختاری مبتنی بر OMP

**Require:** Initialize  $D = \{\}$  set of selected atoms,  $S_{opt} = 0$ ,  $r = Y$  while  $S_{opt} < T_{ssim}$

- 1: Add the next best atom in  $L_2$  sense to  $D$
- 2: Find the optimal  $L_2$ -based coefficient(s) using  $\sum_{j=1}^k c_j(D_i, D_j) = (y, D_i)$ ,  $1 \leq i \leq k$
- 3: Find the optimal SSIM-based coefficient(s) using
 
$$\alpha_i = \beta c_i$$
 and
 
$$\beta = \frac{-C_2 + \sqrt{C_2^2 + 4(B-A)(\sigma_y^2 + C_2)}}{2(B-A)}$$
- 4: Update the residual  $r$
- 5: Find SSIM-based approximation  $a$
- 6: Calculate  $S_{opt} = S(a, y)$

$$A = \frac{1}{n-1} \sum_{i=1}^k \sum_{j=1}^k c_i c_j \langle D_i, D_j \rangle, \quad B = \frac{1}{n-1} \sum_{j=1}^k c_j \langle y, D_j \rangle$$

که  $A = \frac{1}{n-1} \sum_{i=1}^k \sum_{j=1}^k c_i c_j \langle D_i, D_j \rangle$  ،  $B = \frac{1}{n-1} \sum_{j=1}^k c_j \langle y, D_j \rangle$  می‌باشد.

پس از محاسبه‌ی ضرایب تُنک ، تصویر با تفکیک پذیری بالا از رابطه‌ی ۳-۳۹ بدست می‌آید. روش ارائه شده برای مسئله‌ی فراتفکیک پذیری در الگوریتم ۳-۴ خلاصه شده است.

### الگوریتم ۴-۳ الگوریتم فراتفکیک پذیری مبتنی بر شباهت ساختاری

- 1: Dictionary Training Phase: trained high and low resolution dictionaries  $D_l$  and  $D_h$
- 2: Reconstruction Phase:
  - Sparse coding stage: use SSIM-optimal OMP to compute the representation vectors  $\alpha_{ij}$  (for all the patches of low resolution image)
  - High resolution patches reconstruction: Reconstruct high resolution patches by  $D_h \alpha_{ij}$
- 3: Global Reconstruction: merge high-resolution patches by averaging over the overlapped region to create the high resolution image

برای آموزش لغت‌نامه از KSVD استفاده می‌شود و از آنجا که قطعه‌های  $5 \times 5$  در نظر گرفته شده است مشابه روش یانگ اندازه لغت‌نامه با تفکیک پذیری بالا  $25 \times 1024$  و پایین  $100 \times 1024$  می‌باشد. به طور مشابه با روش یانگ [۷] ، برای جلوگیری از مصنوعات<sup>۲</sup> تصویر در لبه‌ها ، یک پیکسل هم‌پوشانی در نظر گرفته می‌شود. همچنین SSIM-OMP تعداد اتم برای هر قطعه را بصورت وقتی با توجه به اهمیت اندازه‌ی معیار شباهت ساختاری مشخص می‌کند.

<sup>۱</sup>SSIM-inspired image super resolution      <sup>۲</sup>Artifact

با توجه به جدول ۱-۳ میزان بهبود مسئله فراتفکیک پذیری در مقایسه با روش یانگ برای تصویر لنا نشان داده شده است.

جدول ۱-۳: بررسی معیار SSIM بجای MSE در فراتفکیک پذیری [۲۵]

معیار شباهت ساختاری	PSNR(db)	
۰.۸۸۸	۳۳.۴	روش یانگ
۰.۹۱۲	۳۳.۹	روش ارائه شده مبتنی بر شباهت ساختاری

در این فصل مسئله‌ی فراتفکیک پذیری با نمایش تُنک و روش یانگ به تفصیل مورد بحث قرار گرفت. همچنین روش فراتفکیک پذیری مبتنی بر معیار SSIM توضیح داده شد. در فصل ۱ آزمایشاتی بر روی روش یانگ با بررسی تغییر پارامترها و ویژگی‌ها به منظور بهبود تفکیک پذیری تصویر انجام خواهد شد.

## فصل ۴

# نتایج آزمایشات

در این بخش از پایان نامه از دیگر روش ها برای استخراج ویژگی و ساخت لغت نامه به منظور یافتن نتیجه‌ی بهتر بهره برده شده است. از طرفی با تغییر پارامترها تاثیرشان بررسی می شود و در ادامه با استفاده از پیاده سازی نمایش تُنگ، در طبقه بندی بکار برده می شود.

### ۱-۴ بکارگیری فیلتر موجک بجای فیلتر گرادیان افقی و عمودی در ساخت لغت نامه و استخراج ویژگی

چنانچه در الگوریتم یانگ بجای ۴ فیلتر گرادیان افقی و عمودی

$$hf1 = [-1, 0, 1]$$

$$vf1 = [-1, 0, 1]'$$

$$hf2 = [1, 0, -2, 0, 1]$$

$$vf2 = [1, 0, -2, 0, 1]'$$

از فیلتر موجک بصورت زیر استفاده شود:

```
load db8
w=db8
[hf1,vf1,hf2,vf2]=orthfilt(w)
```

۱  
۲  
۳

در این حالت، تصویر احیا شده با تفکیک پذیری بالا دارای  $PSNR = 33.96(db)$  است درحالیکه در حالت قبل این مقدار  $35.03$  بود. در نتیجه با این که فیلتر موجک تصویر با تفکیک پذیری پایین را بهبود می بخشد اما بهبود آن نسبت به فیلتر گرادیان کمتر است. در اینجا تنها فیلتر موجک جایگزین فیلتر گرادیان شده است، به عنوان کار آینده می توان هر دو

نوع فیلتر کنار هم بکار برده شود و نتیجه‌ی آن بررسی شود.

جدول ۴-۱: بکارگیری فیلتر موجک بجای فیلتر گرادیان افقی و عمودی در ساخت لغت نامه و استخراج ویژگی

PSNR(dB)	
۳۵.۰۳	فیلتر گرادیان
۳۳.۹۶	فیلتر موجک

## ۲-۴ مشکل روش فراتفکیک پذیری با نمایش تنگ

طبق رابطه ۳-۱۷ برای حل این مسئله بر اساس روش یانگ [۷] با زیر هم گذاشتن قطعه‌های نمونه‌ی با تفکیک پذیری بالا و پایین، بردارهایی به دست آورده می‌شود که بعد آن‌ها مجموع بعد فضاهای قطعات با تفکیک پذیری بالا و پایین است. سپس با استفاده از یک روش آموزش لغت‌نامه، لغت‌نامه ای برای نمایش تنگ این سیگنال‌ها بدست می‌آید. با فرض اینکه در این فرآیند لغت‌نامه‌ی بهینه به دست آید، لغت‌نامه ای خواهد بود که برای سیگنال‌های با بعد افزوده بهینه شده‌است، نه برای هر یک از فضاهای با بعد بالا پایین. برای روشن تر شدن این موضوع بردار  $x_i = [x_{HR_i}, x_{LR_i}]^T$  را در نظر گرفته که از زیر هم گذاشتن قطعه‌ی با تفکیک پذیری بالا و پایین مانند روش بکار رفته در [۷] به دست آمده‌است. فرض کنید نمایش تنگ این سیگنال  $x_{LR_i}$  روی لغت‌نامه‌ی  $D_l$  خواهد بود (زیرا این سیگنال زیر هم گذاشته شده‌است). اما تضمینی وجود ندارد که  $x_{LR_i}$  نمایش تنگ تری روی این لغت‌نامه نداشته باشد. ایرادی که به روش بکار رفته در [۷] وارد است، این است که آموزش توام لغت‌نامه برای فضای افزوده‌ی قطعات با تفکیک پذیری بالا و پایین، برای این فضا بهینه است و برای تک تک این فضاها بهینه نیست. [۶]

همچنین مشکلی که در استفاده از فیلترهای عمودی و افقی بکار گرفته شده در این روش وارد است این است که برای بدست آوردن مقدار آن‌ها در یک نقطه، به مقادیر نقاط مجاور نیز نیاز داریم. به همین علت است که در لبه‌های تصویر ویژگی‌ها به درستی به دست نمی‌آیند. بنابراین با این روش نمی‌توان با استفاده از فراتفکیک پذیری قطعه‌های با تفکیک پذیری بالای متناظر با قطعه‌های با تفکیک پذیری پایین در لبه‌ها را بدست آورد.

## ۳-۴ تغییر پارامترها

در این بخش به منظور دستیابی به تفکیک‌پذیری بالاتر، تاثیر تغییر پارامترهایی از قبیل تعداد نمونه‌ها، اندازه قطعه‌های تصویر، میزان همپوشانی و تعداد اتم‌های لغت‌نامه مورد بررسی قرار خواهد گرفت.



### ۱-۳-۴ تاثیر میزان همپوشانی در افزایش تفکیک پذیری

واضح است که هر چه میزان همپوشانی قطعات مجاور بیشتر باشد، تعداد تخمین هایی که از هر پیکسل در تصویر با تفکیک پذیری بالا خواهیم داشت بیشتر خواهد بود و در نتیجه کیفیت تصویر نهایی بالاتر و خطای آن کمتر خواهد بود. با کاهش میزان همپوشانی در لبه های قطعه ها اثرات نامطلوبی مشاهده می شود که هم به کیفیت بصری تصاویر و هم خطای آن ها تاثیر می گذارد که در تصاویر خروجی زیر نتایج قابل مشاهده است. از آنجا که قطعات ۵x۵ هستند میزان همپوشانی را از یک تا چهار تغییر داده شده است. مطابق با جدول ۴-۲ به ازای همپوشانی ۴ که مقدار پیش فرض در روش یانگ [۷] است و تعداد اتم های ۱۰۲۴ برای لغت نامه، میزان PSNR برای تصویر بازسازی شده نهایی ۳۵.۰۳۶۸۰۴ می باشد که با کاهش میزان همپوشانی مطابق با جدول زیر این مقدار کاهش می یابد. همچنین هر چه میزان همپوشانی کمتر باشد سرعت اجرای کد بیشتر است.

جدول ۴-۲: مقایسه میزان همپوشانی بر روی تصویر لنا (تعداد اتم های لغت نامه ۱۰۲۴ و اندازه قطعات ۵ و تعداد قطعات ۱۰۰۰)

میزان همپوشانی	PSNR(dB)
۱	۳۴.۰۱۶۵۹۵
۲	۳۴.۵۱۹۳۶۱
۳	۳۴.۸۴۳۸۳۸
۴	۳۵.۰۳۶۸۰۴

جدول ۴-۳: مقایسه زمان اجرا برای تصویر لنا (تعداد اتم های لغت نامه ۱۰۲۴ و اندازه قطعات ۵ و تعداد قطعات ۱۰۰۰)

میزان همپوشانی	زمان اجرا
۱	۳۰.۶۹۹۶۰۳
۲	۴۸.۷۵۶۸۷۵
۳	۸۶.۶۳۰۹۳۳
۴	۳۲۰.۴۱۶۹۹۷

### ۲-۳-۴ اثر تعداد اتم های لغت نامه در افزایش تفکیک پذیری

از نظر شهودی هر چه تعداد اتم های لغت نامه بیشتر باشد، قادر است زوج قطعه های نمونه را بیشتر نمایش دهد و در نتیجه اطلاعات بیشتری از زوج قطعه ها را در خود نگه دارد. در حالتی که تعداد اتم ها برابر با تعداد نمونه ها گرفته شود، لغت نامه به خود نمونه ها تبدیل می شود. برای لغت نامه تعداد ۶۴، ۵۱۲، ۲۵۶ و ۱۰۲۴ در نظر گرفته می شود. همان طور که از نتایج مشخص است با افزایش تعداد اتم ها همواره خطا کاهش می یابد و این روند تا وقتی که لغت نامه با خود نمونه ها یکی شود

همچنان ادامه دارد. در نتایج مشاهده می‌شود که اثرات نامطلوب در بازسازی تصویر با افزایش تعداد اتم‌های لغت‌نامه از بین می‌روند. همچنین زمان اجرای الگوریتم فرائتفکیک پذیری تقریباً بصورت خطی با اندازه‌ی لغت‌نامه تغییر می‌کند. بنابراین در عمل بین کیفیت بالاتر و زمان اجرای کمتر باید مصالحه‌ای برقرار کرد. که همان طور که در روش یانگ [۷] گفته شده است تعداد اتم‌های ۱۰۲۴ تعداد مطلوبی می‌باشد که نه در آن اثرات نامطلوب ناشی از کم بودن اتم‌ها به سادگی قابل مشاهده است و نه از نظر زمان الگوریتم خیلی کند خواهد بود. در جدول ۴-۴ میزان PSNR برای تصویر بازسازی شده نهایی به ازای تعداد اتم‌های لغت‌نامه نشان داده شده است. همچنین در شکل ۴-۱ تصویر خروجی لنا با تفکیک پذیری بالا را برای تعداد اتم‌های مختلف نشان داده شده است.



شکل ۴-۱: از چپ به راست تعداد اتم‌ها ۲۰۴۸، ۱۰۲۴، ۵۱۲، ۲۵۶ برای لغت‌نامه

جدول ۴-۴: مقایسه تعداد اتم‌های لغت‌نامه بر روی تصویر لنا (اندازه قطعات ۵ و هم‌پوشانی ۴ و تعداد قطعات ۱۰۰۰)

تعداد اتم‌های لغت‌نامه	PSNR(dB)
۶۴	۳۴.۳۳۴۴۸۲
۲۵۶	۳۴.۳۵۶۰۰۲
۵۱۲	۳۴.۴۵۸۲۶۲
۱۰۲۴	۳۵.۰۳۶۸۰۴

جدول ۴-۵: مقایسه زمان اجرا بر روی تصویر لنا (اندازه قطعات ۵ و هم‌پوشانی ۴ و تعداد قطعات ۱۰۰۰)

تعداد اتم‌های لغت‌نامه	زمان اجرا
۶۴	۶۸.۴۵۶۰۹۷
۲۵۶	۱۱۰.۲۷۸۶۳۷
۵۱۲	۱۸۵.۶۴۱۱۲۴
۱۰۲۴	۳۲۰.۴۱۶۹۹۷

### ۴-۳-۳ اثر اندازه قطعه‌های تصویر در افزایش تفکیک پذیری

بررسی اثر اندازه‌ی قطعه‌های تصویر بر کیفیت نهایی فراتفکیک پذیری بسیار مشکل است. اندازه‌ی قطعه‌ها باید به اندازه‌ای بزرگ باشد که اطلاعات معنی داری را در خود جای داده باشد که بتوان از آن قطعه‌های با تفکیک پذیری بالا را بدست آورد. از سوی دیگر اگر اندازه‌ی قطعه‌ها بیش از حد بزرگ باشد، به تعداد نمونه‌های بسیار زیادی برای یادگیری لغت‌نامه احتیاج است، تا بتواند هر گونه قطعه‌ی ممکن را در بر داشته باشد. این محدودیت که با بزرگ شدن اندازه‌ی قطعه، شکل‌های ممکن درون قطعه به سرعت زیاد می‌شوند باعث می‌شود که عملاً مقذور نیست اندازه از حدی بزرگتر شود. در کاربردهای فراتفکیک پذیری معمولاً اندازه‌ی قطعه‌ها بین  $3 \times 3$  تا  $5 \times 5$  گرفته می‌شود. قطعه‌های کوچک تر اطلاعات کافی را در خود ندارند و برای قطعه‌های بزرگتر به نمونه‌های بسیار زیادی نیاز است.

به علاوه با افزایش اندازه‌ی قطعه‌ها، الگوریتم فراتفکیک پذیری بسته به اینکه از چه روشی برای یافتن تنک ترین نمایش قطعه‌ها استفاده شود، کند می‌شود. همان طور که قبلاً گفته شد کند بودن روش‌های فراتفکیک پذیری یکی از عمده‌ترین مشکل‌های آن‌هاست که کاربرد آن‌ها در عمل محدود کرده‌است. به همین جهت هر تغییری که این روش‌ها را کندتر کند نامطلوب است و استفاده عملی از آن‌ها را محدودتر می‌کند.

از آنجا که با بزرگ شدن اندازه‌ی قطعه‌ها، عوامل بسیار زیاد دیگری نیز در کیفیت روش فراتفکیک پذیری وارد می‌شوند، نمی‌توانیم اندازه‌ی قطعه‌ها را به عنوان یک متغیر در نظر گرفته و اثر آن را به سادگی بر روی روش فراتفکیک پذیری به دست آوریم. همین نکته که با تغییر اندازه به یک دسته دیگر از زوج قطعه‌های نمونه احتیاج داریم، به تنهایی برای مشکل بودن این قضیه کافی است. شبیه سازی‌ها نشان می‌دهد که کیفیت روش فراتفکیک پذیری بسیار به نمونه وابسته است. اگر نمونه‌ها به اندازه کافی زیاد نباشند و یا خوب انتخاب نشده باشند کیفیت روش شدیداً افت می‌کند. این دو هر دو مسائلی هستند که هم‌چنان بی‌پاسخ مانده‌اند: چه تعداد نمونه برای آموزش لغت‌نامه کافی است؟ و دیگر اینکه نمونه‌ها را چگونه انتخاب شود تا نماینده‌ی بهتری از کل نمونه‌ها باشند؟ از آنجا که در اینجا نمونه‌ها به صورت تصادفی از تعدادی تصویر نمونه انتخاب می‌شوند، نتایج برای پارامترهای کاملاً یکسان و فقط نمونه‌های متفاوت کاملاً متفاوت است. به همین علت است که مقایسه بین دو روش که با نمونه‌های متفاوت آموزش دیده‌اند چندان عادلانه نیست، مگر اینکه هر کدام چندین بار با نمونه‌های متفاوت آموزش داده شوند و بهترین حالت برای هر کدام انتخاب شود که با توجه به اینکه آموزش لغت‌نامه بسیار وقت گیر است، کاری بسیار سخت به شمار می‌رود. لازم به ذکر است که هرگونه افزایش اندازه‌ی قطعه‌ها مستلزم افزایش تعداد نمونه‌ها و همچنین افزایش تعداد اتم‌های لغت‌نامه است، در غیر اینصورت ممکن است نه تنها بهبودی در کیفیت حاصل نشود بلکه خطای فراتفکیک پذیری افزایش یابد [۶]. در جدول ۴-۶ برای اندازه قطعات مقایسه‌ای انجام شده‌است. (سرعت اجرای کد در هر دو حالت  $3 \times 3$  و  $7 \times 7$  بیشتر از  $5 \times 5$  است و در  $3 \times 3$  کمتر از ۲ ثانیه است)

جدول ۴-۶: مقایسه اندازه قطعات لغت نامه بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۲۵۶ و هم‌پوشانی ۴ و تعداد قطعات ۱۰۰۰)

اندازه قطعات	PSNR(dB)
۳×۳	۳۳.۹۱۹۸۵۱
۵×۵	۳۴.۳۵۶۰۰۲
۷×۷	۳۴.۰۵۹۸۷۵

جدول ۴-۷: مقایسه زمان اجرا بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۲۵۶ و هم‌پوشانی ۴ و تعداد قطعات ۱۰۰۰)

اندازه قطعات	زمان اجرا
۳×۳	۱.۶۵۳۱۴۸
۵×۵	۱۱۰.۲۷۸۶۳۷
۷×۷	۲۰.۳۷۵۸۲۷

#### ۴-۳-۴ اثر تعداد نمونه‌ها برای ساخت لغت‌نامه

انتظار می‌رود که هر چه تعداد نمونه‌ها بیشتر باشد، لغت‌نامه‌ی ساخته‌شده غنی‌تر و در نتیجه کیفیت تصویر نهایی بالاتر و خطای آن کمتر باشد. در جدول ۴-۸ برای تعداد قطعات مقایسه‌ای انجام شده است (تعداد اتم‌های لغت‌نامه ۵۱۲ و هم‌پوشانی ۴ در نظر گرفته شده است و تصویر ورودی (dB) ۳۲.۷۹۴۶۷۸ است). همچنین هر چه تعداد نمونه‌ها کمتر باشد سرعت اجرای کد بیشتر است. انتخاب بهترین تعداد نمونه بحثی است که می‌تواند برای کارهای آینده مطرح شود.

جدول ۴-۸: مقایسه تعداد نمونه قطعات بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۵۱۲ و هم‌پوشانی ۴)

تعداد قطعات	PSNR(dB)
۱۰۰۰	۳۴.۳۸۵۱۶۱
۵۰۰۰	۳۴.۵۲۸۳۸۶
۱۰۰۰۰	۳۴.۵۵۷۴۹۶

#### ۴-۴ ساخت لغت‌نامه از روی تصاویر صورت

پیش‌تر بیان شد که یکی از کاربردهای روش یانگ برای بهبود تفکیک پذیری تصاویر صورت است به این منظور برای ساخت لغت‌نامه، پایگاه داده تصاویر صورت بکار گرفته شده است. اگر لغت‌نامه را از روی تصاویر شکل ۴-۲ این تصاویر با نام olivettiFaces می‌باشد که شامل ۴۰۰ تصویر خاکستری با اندازه ۶۴×۶۴ از تصاویر صورت افراد مختلفی می‌باشد. این پایگاه داده از آدرس<sup>۱</sup> قابل دسترسی است حال اگر لغت‌نامه بر اساس این پایگاه داده ساخته و کد روی تصویر صورت اعمال

<sup>۱</sup><http://cs.nyu.edu/roweis/data/olivettifaces.mat>

جدول ۴-۹: مقایسه زمان اجرا بر روی تصویر لنا (تعداد اتم‌های لغت‌نامه ۵۱۲ و هم‌پوشانی ۴)

تعداد قطعات	زمان اجرا
۱۰۰۰	۱۸۵.۶۴۱۱۲۴
۵۰۰۰	۲۱۱.۵۹۷۹۹۳
۱۰۰۰۰	۳۱۶.۸۱۷۰۵۳

شود، نتیجه بصورت شکل ۴-۳ خواهد بود. تصویر ورودی در شکل ۴-۳ (آ) نشان داده شده است.



شکل ۴-۲: داده‌های ورودی برای لغت‌نامه



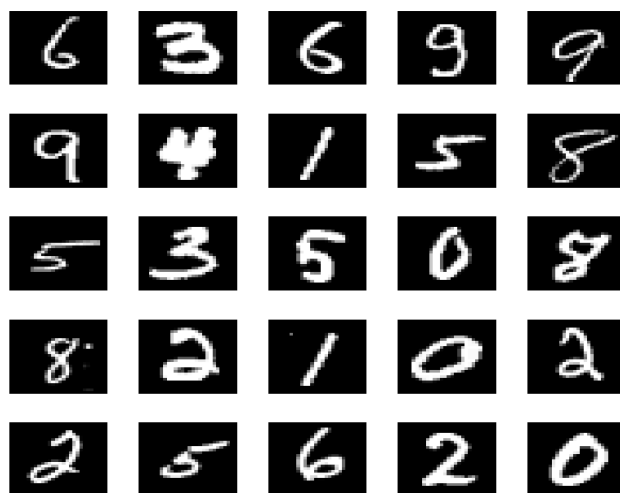
(آ) تصویر ورودی با وضوح پایین (ب) تصویر خروجی با وضوح بالا

شکل ۴-۳: ورودی و خروجی فراتفکیک‌پذیری با لغت‌نامه تصاویر صورت

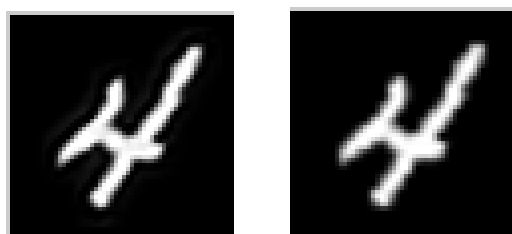
## ۵-۴ ساخت لغت‌نامه بر اساس پایگاه داده اعداد mnistAll و اعمال

### فراقتیک پذیري با استفاده از اين لغت‌نامه

چنانچه آموزش لغت‌نامه بر اساس پایگاه داده اعداد mnistAll انجام شود که این پایگاه شامل ارقام دست‌نویس ۰ تا ۹ بصورت انگلیسی می‌باشد که تصویر هر رقم آن بصورت یک تصویر خاکستری  $28 \times 28$  پیکسل ذخیره شده است. مجموعه آموزش آن شامل ۶۰۰۰۰ تصویر و مجموعه آزمایش آن حاوی ۱۰۰۰۰ تصویر می‌باشد. این پایگاه داده از آدرس<sup>۱</sup> قابل دسترس است. اگر لغت‌نامه را از روی تصاویر شکل ۴-۴ ساخته و کد را روی تصویر صورت اعمال کنیم، نتیجه بصورت شکل ۵-۴(ب) خواهد بود. تصویر ورودی در شکل ۴-۴(ا) نشان داده شده است.



شکل ۴-۴: داده‌های ورودی برای لغت‌نامه



(ا) تصویر ورودی با وضوح پایین  
(ب) تصویر خروجی با وضوح بالا

شکل ۵-۴: ورودی و خروجی فراقتیک پذیري با لغت‌نامه تصاویر اعداد

<sup>۱</sup><http://yann.lecun.com/exdb/mnist/>

در اینجا عملکرد به این صورت است که لغت نامه بگونه‌ای ساخته شود که هر ستون  $D_l$  یک کاراکتر با وضوح پایین و هر ستون  $D_h$  یک کاراکتر با وضوح بالا باشد. در ادامه کد روی یک کاراکتر تست که در آموزش نیست اعمال شده سپس روی تصویر لنا هم کد اعمال می‌شود. نتیجه نشان می‌دهد که ساخت لغت نامه به این روش کارا نیست و نتایج خوبی حاصل نمی‌شود.

## ۶-۴ تغییر روش حل مسئله‌ی بهینه‌سازی تُنک

چنانچه در فصل ۱ بیان شد، روش‌هایی از قبیل OMP و KSVD هم برای مرحله‌ی آموزش لغت نامه و هم مرحله‌ی نمایش تُنک قابل استفاده است. از اینرو در این بخش به منظور افزایش تفکیک پذیری به جایگزینی این روش‌ها بجای روش L1QP که در روش یانگ بکار برده شده است، پرداخته می‌شود.

### ۱-۶-۴ جایگزینی روش OMP بجای روش L1QP

چنانچه در روش بهینه‌سازی یانگ بجای روش L1QP از روش OMP استفاده شود سرعت اجرای برنامه برای لغت نامه ای با سایز ۶۴ کاهش می‌یابد. نتیجه بصورت شکل ۶-۴(ب) خواهد بود. تصویر ورودی در شکل ۶-۴(آ) نشان داده شده است.



(ب) تصویر خروجی با وضوح بالا



(آ) تصویر ورودی با وضوح پایین

شکل ۶-۴: ورودی و خروجی تفکیک پذیری با روش OMP (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

جدول ۱۰-۴: مقایسه روش برای آموزش لغت نامه و عمل فراتفکیک پذیری بر روی تصویر لنا (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

PSNR(dB)	
۳۳.۳۱۵۹۹۶	روش OMP
۳۴.۳۳۴۴۸۲	روش L1QP

جدول ۴-۱۱: مقایسه زمان اجرا بر روی تصویر لنا (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

زمان اجرا	
۱۳۶۱.۱۱۹۷۵۲	روش OMP
۶۸.۴۵۶۰۹۷	روش L1QP

#### ۲-۶-۴ جایگزینی روش KSVD بجای L1QP

چنانچه در پیاده سازی در ساخت لغت نامه بجای روش L1QP از روش KSVD استفاده شود و از ضرایب تنک بدست آمده در مرحله ساخت لغت نامه در مرحله ی احیای تصویر با تفکیک پذیری (مرحله بازسازی تصویر) استفاده شود و همچنین لغت نامه بصورت بکه باشد (غیر توام) نتیجه بر اساس PSNR برای تصویر ورودی لنا با (dB) 32.794678 برابر با (dB) 33.919877 می باشد. با این روش سرعت اجرای برنامه بسیار بالاتر است. نتیجه بصورت شکل ۴-۷ (ب) خواهد بود. تصویر ورودی در شکل ۴-۷ (آ) نشان داده شده است.



(ب) تصویر خروجی با وضوح بالا

(آ) تصویر ورودی با وضوح پایین

شکل ۴-۷: ورودی و خروجی تفکیک پذیری با روش KSVD (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

جدول ۴-۱۲: مقایسه روش برای آموزش لغت نامه و عمل فراتفکیک پذیری بر روی تصویر لنا (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

PSNR(dB)	
۳۳.۹۱۹۸۷۷	روش KSVD
۳۴.۳۳۴۴۸۲	روش L1QP



جدول ۴-۱۳: مقایسه زمان اجرا بر روی تصویر لنا (لغت نامه با اندازه ۶۴ اتم و اندازه قطعات ۵)

زمان اجرا	
۱.۵۲۹۰۶۷	روش KSVD
۶۸.۴۵۶۰۹۷	روش L1QP

## ۷-۴ نتیجه گیری

با جایگزینی روش های دیگر برای ساخت لغت نامه و حل مسئله ی بهینه سازی بجای روش های بکارگرفته شده در روش یانگ نشان داده شد گرچه می توان سرعت را بهبود بخشید اما در امر بهبود تفکیک پذیری همچنان روش یانگ بهتر عمل می کند و این بیانگر اهمیت این روش است که همچنان روش پایه ای بکار گرفته شده در بیشتر روش های امروزی است گرچه که اشکالاتی هم بر این روش وارد است که اشاره شد. همچنین با تغییر پارامترها و بررسی تاثیر آنها نشان داده شد که برای انتخاب بهترین مقدار در هر پارامتر باید اولویت هدف ما مشخص شود که آیا بهبود تفکیک پذیری و یا سرعت اجرا، در بهترین حالت برای ما حائز اهمیت است. که در عمل مقادیری بهترین نتیجه را دارد که بین دو هدف ما تعادل برقرار کند.

# فهرست منابع

- [۱] صادقی، مصطفی. نمایش تنک و کاربرد آن در نويززدایی تصاویر. پایان‌نامه کارشناسی ارشد، دانشکده برق، دانشگاه صنعتی شریف، ۱۳۹۱.
- [2] Vetterli, Martin and Kovacevic, Jelena. *Wavelets and subband coding*. no. LCAV-BOOK-1995-001. Prentice-hall, 1995.
- [3] Elad, M., Figueiredo, M.A.T., and Ma, Yi. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, 98(6):972–982, June 2010.
- [4] Candy, James V. *Signal Processing: Model Based Approach*. McGraw-Hill, Inc., New York, NY, USA, 1986.
- [۵] جنیدی، محسن، بخشی گلستانی، حسین، و صادقی، مصطفی. نويز زدایی تصاویر به کمک خوشه‌بندی متعادل در حوزه نمایش تنک. در بیست و یکمین کنفرانس مهندسی برق ایران، مشهد، ایران، ۱۳۹۲. دانشگاه فردوسی مشهد.
- [۶] صحرایی، مجتبی. کاربردهای نمایش تنک در افزایش تفکیک پذیری تصاویر. پایان‌نامه کارشناسی ارشد، دانشکده برق، دانشگاه صنعتی شریف، ۱۳۹۱.
- [7] Yang, Jianchao, Wright, J., Huang, T.S., and Ma, Yi. Image super-resolution via sparse representation. *Image Processing, IEEE Transactions on*, 19(11):2861–2873, Nov 2010.
- [8] Xing, Xiaofen, Qiu, Fuhao, Xu, Xiangmin, Qing, Chunmei, and Wu, Yinrong. Robust object tracking based on sparse representation and incremental weighted pca. *Multimedia Tools and Applications*, pp. 1–19, 2016.
- [9] Mairal, J., Bach, F., Ponce, J., Sapiro, G., and Zisserman, A. Discriminative learned dictionaries for local image analysis. in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, June 2008.
- [10] Mairal, J., Elad, M., and Sapiro, G. Sparse representation for color image restoration. *Image Processing, IEEE Transactions on*, 17(1):53–69, Jan 2008.
- [11] Mairal, Julien, Ponce, Jean, Sapiro, Guillermo, Zisserman, Andrew, and Bach, Francis R. Supervised dictionary learning. in Koller, D., Schuurmans, D., Bengio, Y., and Bottou, L., eds., *Advances in Neural Information Processing Systems 21*, pp. 1033–1040. Curran Associates, Inc., 2009.

- [۱۲] حقیری، سیاوش. کاربردهای نمایش تنک در افزایش تفکیک پذیری تصاویر. پایان‌نامه کارشناسی ارشد، دانشکده کامپوتر، دانشگاه صنعتی شریف، ۱۳۹۱.
- [13] Qin, Zhen-tao, Yang, Wu-nian, Wu, Xiao-pin, and Yang, Ru. *Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment: Challenges, Processes and Strategies*, chap. Hyperspectral Image Classification Using a New Dictionary Learning Approach with Structured Sparse Representation, pp. 719–722. Springer International Publishing, Cham, 2016.
- [14] Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., and Ma, Yi. Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2):210–227, Feb 2009.
- [۱۵] غفاری، ابوزر، اسماعیلی، ایمان، و فاطمی زاده، عمادالدین. شناسایی چهره با استفاده از ترکیب طبقه بندهای نمایش تنک و نزدیک ترین زیرفضا. در اولین کنفرانس بازشناسی الگو و پردازش تصویر ایران، بیرجند، ایران، ۱۳۹۱. دانشگاه بیرجند.
- [۱۶] ذوالفقاری، محمدرضا. تخمین حالت انسان با استفاده از نمایش تنک. پایان‌نامه کارشناسی ارشد، دانشکده کامپوتر، دانشگاه صنعتی شریف، ۱۳۹۱.
- [17] Yang, Wenji, Kong, Lingfu, and Wang, Mingyan. Hand gesture recognition using saliency and histogram intersection kernel based sparse representation. *Multimedia Tools and Applications*, pp. 1–14, 2015.
- [18] Liao, H.Y. and Sapiro, G. Sparse representations for limited data tomography. in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pp. 1375–1378, May 2008.
- [19] RIT, Hassan. Multifocus image fusion using sparse representation by adaptive feature matching. *International Journal*, 3(8), 2013.
- [۲۰] صادقی، مصطفی. نمایش تنک و کاربرد آن در نوین زدایی تصاویر. پایان‌نامه کارشناسی ارشد، دانشکده برق، دانشگاه صنعتی شریف، ۱۳۹۱.
- [21] Mallat, Stéphane G and Zhang, Zhifeng. Matching pursuits with time-frequency dictionaries. *Signal Processing, IEEE Transactions on*, 41(12):3397–3415, 1993.
- [22] Chu, WC-C, Chao, H-C, and Yang, SJ-H. *Intelligent Systems and Applications: Proceedings of the International Computer Symposium (ICS) Held at Taichung, Taiwan, December 12–14, 2014*, vol. 274. IOS Press, 2015.
- [۲۳] جلالی، آرش. نهان‌کاوی تصاویر بر مبنای نمایش تنک. پایان‌نامه کارشناسی ارشد، دانشکده برق، دانشگاه صنعتی شریف، ۱۳۹۴.
- [24] Elad, Michael and Aharon, Michal. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image processing*, 15(12):3736–3745, 2006.

- [25] Ogawa, Takahiro and Haseyama, Miki. Image inpainting based on sparse representations with a perceptual metric. *EURASIP Journal on Advances in Signal Processing*, 2013(1):1–26, 2013.
- [26] Rehman, Abdul, Rostami, Mohammad, Wang, Zhou, Brunet, Dominique, and Vrscaj, Edward R. Ssim-inspired image restoration using sparse representation. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–12, 2012.

# پیوست آ

## نمونه برنامه‌ی های نوشته شده

نمونه برنامه‌ی های نوشته شده‌ی زیر مبتنی بر روش یانگ است و آزمایشات با اعمال تغییراتی انجام گرفته است.

برنامه آ-۱: برنامه‌ی فراتفکیک پذیری با نمایش تئک [۷]

```
clear all; clc;
% read test image
im_l = imread('Data/Testing/input.bmp');
% set parameters
lambda = 0.2; % sparsity regularization
overlap = 4; % the more overlap the better (patch size 5x5)
up_scale = 2; % scaling factor, depending on the trained dictionary
maxIter = 20; % if 0, do not use backprojection
% load dictionary
load('Dictionary/D_64_0.15_5.mat');
% change color space, work on illuminance only
im_l_ycbcr = rgb2ycbcr(im_l);
im_l_y = im_l_ycbcr(:, :, 1);
im_l_cb = im_l_ycbcr(:, :, 2);
im_l_cr = im_l_ycbcr(:, :, 3);
% image super-resolution based on sparse representation
[im_h_y] = ScSR(im_l_y, 2, Dh, Dl, lambda, overlap);
[im_h_y] = backprojection(im_h_y, im_l_y, maxIter);
% upscale the chrominance simply by "bicubic"
[nrow, ncol] = size(im_h_y);
im_h_cb = imresize(im_l_cb, [nrow, ncol], 'bicubic');
im_h_cr = imresize(im_l_cr, [nrow, ncol], 'bicubic');
```

```

im_h_ycbr = zeros([nrow, ncol, 3]);          ۳۱
im_h_ycbr(:, :, 1) = im_h_y;                ۳۲
im_h_ycbr(:, :, 2) = im_h_cb;               ۳۳
im_h_ycbr(:, :, 3) = im_h_cr;               ۳۴
im_h = ycbcr2rgb(uint8(im_h_ycbr));         ۳۵
                                              ۳۶
% bicubic interpolation for reference          ۳۷
im_b = imresize(im_l, [nrow, ncol], 'bicubic');  ۳۸
                                              ۳۹
% read ground truth image                    ۴۰
im = imread('Data/Testing/gnd.bmp');         ۴۱
                                              ۴۲
% compute PSNR for the illuminance channel    ۴۳
bb_rmse = compute_rmse(im, im_b);            ۴۴
sp_rmse = compute_rmse(im, im_h);            ۴۵
                                              ۴۶
bb_psnr = 20*log10(255/bb_rmse);             ۴۷
sp_psnr = 20*log10(255/sp_rmse);             ۴۸
                                              ۴۹
fprintf('PSNR for Bicubic Interpolation: %f dB\n', bb_psnr);  ۵۰
fprintf('PSNR for Sparse Representation Recovery: %f dB\n', sp_psnr);  ۵۱
                                              ۵۲
% show the images                            ۵۳
figure, imshow(im_h);                        ۵۴
title('Sparse Recovery');                    ۵۵
figure, imshow(im_b);                        ۵۶
title('Bicubic Interpolation');              ۵۷

```

### برنامه آ-۲: برنامه‌ی ساخت لغت‌نامه

```

clear all; clc; close all;                  ۱
addpath(genpath('RegularizedSC'));          ۲
                                              ۳
TR_IMG_PATH = 'Data/Training';              ۴
                                              ۵
dict_size   = 1024;                          % dictionary size      ۶
lambda      = 0.15;                          % sparsity regularization  ۷
patch_size  = 5;                             % image patch size        ۸
nSmp        = 1000;                          % number of patches to sample ۹
upscale     = 2;                             % upscaling factor        ۱۰
                                              ۱۱
% randomly sample image patches              ۱۲
[Xh, Xl] = rnd_smp_patch(TR_IMG_PATH, '*.bmp', patch_size, nSmp,  ۱۳
    upscale);                                ۱۴
                                              ۱۵
% prune patches with small variances, threshold chosen based on the  ۱۶
% training data                              ۱۷
[Xh, Xl] = patch_pruning(Xh, Xl, 10);       ۱۸
                                              ۱۹
% joint sparse coding                        ۲۰
[Dh, Dl] = train_coupled_dict(Xh, Xl, dict_size, lambda, upscale);  ۲۱
dict_path = ['Dictionary/D_' num2str(dict_size) '_' num2str(lambda)  ۲۲
    num2str(patch_size) '.mat' ];
save(dict_path, 'Dh', 'Dl');                 ۲۳

```

برنامه آ-۳: برنامه L1QP

```

function [x]=L1QP_FeatureSign_yang(lambda,A,b)      ۱
                                                    ۲
A = double(A);                                    ۳
b = double(b);                                    ۴
                                                    ۵
EPS = 1e-9;                                       ۶
x=zeros(size(A, 1), 1);                          %coeff      ۷
                                                    ۸
grad=A*sparse(x)+b;                               ۹
[ma mi]=max(abs(grad).*(x==0));                  ۱۰
                                                    ۱۱
while true,                                       ۱۲
                                                    ۱۳
    if grad(mi)>lambda+EPS,                       ۱۴
        x(mi)=(lambda-grad(mi))/A(mi,mi);        ۱۵
    elseif grad(mi)<-lambda-EPS,                 ۱۶
        x(mi)=(-lambda-grad(mi))/A(mi,mi);      ۱۷
    else                                          ۱۸
        if all(x==0)                             ۱۹
            break;                                ۲۰
        end                                       ۲۱
    end                                          ۲۲
end                                              ۲۳
                                                    ۲۴
while true,                                       ۲۵
    a=x~=0; %active set                          ۲۶
    Aa=A(a,a);                                    ۲۷
    ba=b(a);                                       ۲۸
    xa=x(a);                                       ۲۹
                                                    ۳۰
    %new b based on unchanged sign                ۳۱
    vect = -lambda*sign(xa)-ba;                   ۳۲
    x_new= Aa\vect;                                ۳۳
    idx = find(x_new);                             ۳۴
    o_new=(vect(idx)/2 + ba(idx))*x_new(idx) + lambda*sum(abs(x_new( ۳۵
        idx)));
                                                    ۳۶
    %cost based on changing sign                  ۳۷
    s=find(xa.*x_new<=0);                         ۳۸
    if isempty(s)                                  ۳۹
        x(a)=x_new;                                ۴۰
        loss=o_new;                                ۴۱
        break;                                     ۴۲
    end                                          ۴۳
    x_min=x_new;                                   ۴۴
    o_min=o_new;                                   ۴۵
    d=x_new-xa;                                    ۴۶
    t=d./xa;                                       ۴۷
    for zd=s',                                     ۴۸
        x_s=xa-d/t(zd);                             ۴۹
        x_s(zd)=0; %make sure it's zero           ۵۰
    %      o_s=L1QP_loss(net,Aa,ba,x_s);          ۵۱
    idx = find(x_s);                               ۵۲
    o_s = (Aa(idx, idx)*x_s(idx)/2 + ba(idx))*x_s(idx)+lambda*sum( ۵۳

```

abs(x_s(idx));	
if o_s<o_min,	۵۴
x_min=x_s;	۵۵
o_min=o_s;	۵۶
end	۵۷
end	۵۸
	۵۹
x(a)=x_min;	۶۰
loss=o_min;	۶۱
end	۶۲
	۶۳
grad=A*sparse(x)+b;	۶۴
	۶۵
[ma mi]=max(abs(grad).*(x==0));	۶۶
if ma <= lambda+EPS,	۶۷
break;	۶۸
end	۶۹
end	۷۰

برنامه آ-۴: برنامه‌ی ساخت استخراج ویژگی

function [lImFea] = extr_lIm_fea( lIm )	۱
	۲
[nrow, ncol] = size(lIm);	۳
	۴
lImFea = zeros([nrow, ncol, 4]);	۵
	۶
<i>% first order gradient filters</i>	۷
hf1 = [-1,0,1];	۸
vf1 = [-1,0,1]';	۹
	۱۰
lImFea(:, :, 1) = conv2(lIm, hf1, 'same');	۱۱
lImFea(:, :, 2) = conv2(lIm, vf1, 'same');	۱۲
	۱۳
<i>% second order gradient filters</i>	۱۴
hf2 = [1,0,-2,0,1];	۱۵
vf2 = [1,0,-2,0,1]';	۱۶
	۱۷
lImFea(:, :, 3) = conv2(lIm, hf2, 'same');	۱۸
lImFea(:, :, 4) = conv2(lIm, vf2, 'same');	۱۹

برنامه آ-۵: برنامه‌ی ScSR

function [hIm] = ScSR(lIm, up_scale, Dh, Dl, lambda, overlap)	۱
	۲
<i>% normalize the dictionary</i>	۳
norm_Dl = sqrt(sum(Dl.^2, 1));	۴
Dl = Dl./repmat(norm_Dl, size(Dl, 1), 1);	۵
	۶
patch_size = sqrt(size(Dh, 1));	۷
	۸
<i>% bicubic interpolation of the low-resolution image</i>	۹
mIm = single(imresize(lIm, up_scale, 'bicubic'));	۱۰
	۱۱



```

hIm = zeros(size(mIm));
cntMat = zeros(size(mIm));

[h, w] = size(mIm);

% extract low-resolution image features
lImfea = extr_lIm_fea(mIm);

% patch indexes for sparse recovery (avoid boundary)
gridx = 3:patch_size - overlap : w-patch_size-2;
gridx = [gridx, w-patch_size-2];
gridy = 3:patch_size - overlap : h-patch_size-2;
gridy = [gridy, h-patch_size-2];

A = D1'*D1;
cnt = 0;

% loop to recover each low-resolution patch
for ii = 1:length(gridx),
    for jj = 1:length(gridy),

        cnt = cnt+1;
        xx = gridx(ii);
        yy = gridy(jj);

        mPatch = mIm(yy:yy+patch_size-1, xx:xx+patch_size-1);
        mMean = mean(mPatch(:));
        mPatch = mPatch(:) - mMean;
        mNorm = sqrt(sum(mPatch.^2));

        mPatchFea = lImfea(yy:yy+patch_size-1, xx:xx+patch_size-1, :);
        mPatchFea = mPatchFea(:);
        mfNorm = sqrt(sum(mPatchFea.^2));

        if mfNorm > 1,      %???
            y = mPatchFea./mfNorm;
        else
            y = mPatchFea;
        end

        b = -D1'*y;

        % sparse recovery
        w = L1QP_FeatureSign_yang(lambda, A, b);

        % generate the high resolution patch and scale the contrast
        hPatch = Dh*w;
        hPatch = lin_scale(hPatch, mNorm);

        hPatch = reshape(hPatch, [patch_size, patch_size]);
        hPatch = hPatch + mMean;

        hIm(yy:yy+patch_size-1, xx:xx+patch_size-1) = hIm(yy:yy+
            patch_size-1, xx:xx+patch_size-1) + hPatch;
        cntMat(yy:yy+patch_size-1, xx:xx+patch_size-1) = cntMat(yy:yy+
            patch_size-1, xx:xx+patch_size-1) + 1;

```

```

        end
end
% fill in the empty with bicubic interpolation
idx = (cntMat < 1);
hIm(idx) = mIm(idx);

cntMat(idx) = 1;
hIm = hIm./cntMat;
hIm = uint8(hIm);

```

### برنامه آ-۶: برنامه‌ی آموزش لغت‌نامه

```

function [Dh, Dl] = train_coupled_dict(Xh, Xl, dict_size, lambda,
    upscale)

addpath(genpath('RegularizedSC'));

hDim = size(Xh, 1);
lDim = size(Xl, 1);

% should pre-normalize Xh and Xl !
hNorm = sqrt(sum(Xh.^2));
lNorm = sqrt(sum(Xl.^2));
Idx = find( hNorm & lNorm );

Xh = Xh(:, Idx);
Xl = Xl(:, Idx);

Xh = Xh./repmat(sqrt(sum(Xh.^2)), size(Xh, 1), 1);
Xl = Xl./repmat(sqrt(sum(Xl.^2)), size(Xl, 1), 1);

% joint learning of the dictionary
X = [sqrt(hDim)*Xh; sqrt(lDim)*Xl];
Xnorm = sqrt(sum(X.^2, 1));

clear Xh Xl;

X = X(:, Xnorm > 1e-5);
X = X./repmat(sqrt(sum(X.^2, 1)), hDim+lDim, 1);

idx = randperm(size(X, 2));

% dictionary training
[D] = reg_sparse_coding(X, dict_size, [], 0, lambda, 40);

Dh = D(1:hDim, :);
Dl = D(hDim+1:end, :);

% normalize the dictionary
% Dh = Dh./repmat(sqrt(sum(Dh.^2, 1)), hDim, 1);
% Dl = Dl./repmat(sqrt(sum(Dl.^2, 1)), lDim, 1);

patch_size = sqrt(size(Dh, 1));

```

```
dict_path = ['Dictionary/D_' num2str(dict_size) '_' num2str(lambda) '_f'
            num2str(patch_size) '_s' num2str(upscale) '.mat' ];
save(dict_path, 'Dh', 'Dl');
```

```
function [HP, LP] = sample_patches(im, patch_size, patch_num, upscale)
if size(im, 3) == 3,
    hIm = rgb2gray(im);
else
    hIm = im;
end

% generate low resolution counter parts
lIm = imresize(hIm, 1/upscale, 'bicubic');
lIm = imresize(lIm, size(hIm), 'bicubic');
[nrow, ncol] = size(hIm);

x = randperm(nrow-2*patch_size-1) + patch_size;
y = randperm(ncol-2*patch_size-1) + patch_size;

[X,Y] = meshgrid(x,y);

xrow = X(:);
ycol = Y(:);

if patch_num < length(xrow),
    xrow = xrow(1:patch_num);
    ycol = ycol(1:patch_num);
end

patch_num = length(xrow);

hIm = double(hIm);
lIm = double(lIm);

H = zeros(patch_size^2, length(xrow));
L = zeros(4*patch_size^2, length(xrow));

% compute the first and second order gradients
hf1 = [-1,0,1];
vf1 = [-1,0,1]';

lImG11 = conv2(lIm, hf1, 'same');
lImG12 = conv2(lIm, vf1, 'same');

hf2 = [1,0,-2,0,1];
vf2 = [1,0,-2,0,1]';

lImG21 = conv2(lIm,hf2,'same');
lImG22 = conv2(lIm,vf2,'same');

for ii = 1:patch_num,
    row = xrow(ii);
    col = ycol(ii);
```

```

Hpatch = hIm(row:row+patch_size-1,col:col+patch_size-1);      02
                                                                    03
Lpatch1 = lImG11(row:row+patch_size-1,col:col+patch_size-1);  04
Lpatch2 = lImG12(row:row+patch_size-1,col:col+patch_size-1);  05
Lpatch3 = lImG21(row:row+patch_size-1,col:col+patch_size-1);  06
Lpatch4 = lImG22(row:row+patch_size-1,col:col+patch_size-1);  07
                                                                    08
Lpatch = [Lpatch1(:),Lpatch2(:),Lpatch3(:),Lpatch4(:)];      09
Lpatch = Lpatch(:);                                           10
                                                                    11
HP(:,ii) = Hpatch(:)-mean(Hpatch(:));                          12
LP(:,ii) = Lpatch;                                           13
end                                                            14

```

```

function [Xh, Xl] = rnd_smp_patch(img_path, type, patch_size, num_patch
, upscale)                                                    2
                                                                    3
img_dir = dir(fullfile(img_path, type));                       4
                                                                    5
Xh = [];                                                       6
Xl = [];                                                       7
                                                                    8
img_num = length(img_dir);                                    9
nper_img = zeros(1, img_num);                                 10
                                                                    11
for ii = 1:length(img_dir),                                   12
    im = imread(fullfile(img_path, img_dir(ii).name));        13
    nper_img(ii) = prod(size(im));                             14
end                                                            15
                                                                    16
nper_img = floor(nper_img*num_patch/sum(nper_img));          17
                                                                    18
for ii = 1:img_num,                                         19
    patch_num = nper_img(ii);                                  20
    im = imread(fullfile(img_path, img_dir(ii).name));        21
    [H, L] = sample_patches(im, patch_size, patch_num, upscale); 22
    Xh = [Xh, H];                                             23
    Xl = [Xl, L];                                             24
end                                                            25
                                                                    26
patch_path = ['Training/rnd_patches_' num2str(patch_size) '_' num2str(num_patch)
'_s' num2str(upscale) '.mat'];                                27
save(patch_path, 'Xh', 'Xl');

```

# Hakim Sabzevari University

An Outline of MSc. Thesis



Surname:Asaran	Name:Alale	Student No.:9313137031
Supervisor: Dr. Mahmood Amintoosi		
Advisor: Dr. Mahdi Zaferanie		
Faculty of Mathematics and Computer Science		
Program: Decision Science and Knowledge Engineering		
Title of thesis: Super Resolution Via Sparse Representation		
Keywords: Sparse Representation, Dictionary Learning, Super Resolution, Feature Extraction		
<p>Abstract: In the last decades, sparse signal processing has been considered significantly as a replacement for complete classical transformation. Based on the sparse representation, we aim to represent a signal by finding the least number of basic signals among a huge number of them. Every basic signal is known as an atom and a set of atoms is called a dictionary. In the recent years, by using the practical algorithms, it has been shown that the most sparse representation of a signal in an over-complete dictionary under a specific constraint is unique. In the first chapter different usages of this subject has been discussed such as the following: Image denoising, super-resolution, image classification, face recognition, image fusion to improve the resolution. There are two important problems in sparse representation. First problem is to find an appropriate over-complete dictionary for a class of specific data. In other words, the over-complete dictionary should be able to give an adequate sparse representation for every signal of the class as it is shown in the following algorithms: MOD and K-SVD. The second problem is having an effective algorithm to find the most sparse signal representation (i.e. sparse coding) such as MP and OMP. This thesis is based on Yang's approach, presented in details in chapter 3. To find an alternative better solution, other approaches for feature extraction and dictionary learning are tested and the effect of changing the parameters is processed.</p>		



**Hakim Sabzevari University**  
**Faculty of Mathematics and Computer Science**

**A Thesis Submitted in Partial Fulfilment of the Requirement for the  
Degree of Master of Science in Decision Science and Knowledge  
Engineering**

# **Super Resolution Via Sparse Representation**

**Supervisor:**  
**Dr. Mahmood Amintoosi**

**Advisor:**  
**Dr. Mahdi Zaferanie**

**By:**  
**Alale Asaran**

**Winter 2016**