

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

درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی

استاد راهنما

دکتر محمود امین طوسی

استاد مشاور

دکتر علیرضا قدسی

پژوهشگر

فرزاد زندی

تیر ۱۳۹۷



دانشگاه گیلان

باسمه تعالی

فرم ارزشیابی و صورتجلسه دفاع از پایان نامه کارشناسی ارشد

فرم ۱۱۳-ت

جلسه دفاع از پایان نامه آقای / خانم فرزاد زندی دانشجوی رشته علوم تصمیم و مهندسی دانش به شماره دانشجویی ۹۵۱۳۱۳۷۰۴۸ با عنوان:

درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی

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 فرزاد زندی 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، 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 ۱۹ 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 داوران

ردیف	نام و نام خانوادگی	سمت	مرتبۀ علمی	محل کار	امضا
۱	دکتر محمود امین طوسی	استاد راهنما	استادیار	دانشگاه سبزوار	حکیم
۲	دکتر علیرضا قدسی	استاد مشاور	استادیار	دانشگاه سبزوار	حکیم
۳	دکتر ...	استاد داور	استادیار	دانشگاه سبزوار	حکیم
۴	دکتر ...	نماینده تحصیلات تکمیلی	استادیار	دانشگاه سبزوار	حکیم

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 با راه و رسم انسانی، آیین پرهیزگاری، شرافت و اصول اخلاقی برخاسته از ادیان بزرگ الهی، به ویژه دین مبین اسلام، مباینت 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: فرزند زندی

تاریخ و امضا:

تأییدی صحت و اصالت نتایج

باسمه تعالی

اینجانب فرزاد زندی به شماره دانشجویی ۹۵۱۳۱۳۷۰۴۸ 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 و ...) 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 اینجانب خواهد بود و دانشگاه هیچ گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی : فرزاد زندی

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.

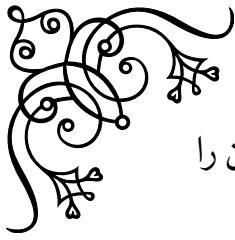
بهره برداری از این پایان نامه با اخذ مجوز از استاد راهنما، بلامانع است.

بهره برداری از این پایان نامه تا تاریخ ممنوع است.

استاد راهنما : دکتر محمود امین طوسی

تاریخ و امضا :

تقدیم به :



خدایی که آفرید جهان را، انسان را،
عقل را، علم را، عشق را و به کسانی که عشقشان را
در وجودم دمید.
و مهربان ترین فرشتگان؛
پدر و مادر عزیزم
و بهترین همراهانم؛
خواهرانم خانم دکتر زهرا زندی و خانم دکتر نازنین زندی



سپاس خدای را که سخنوران در ستودن او بمانند و شمارندگان شمردن نعمت های او ندانند.
بسی شایسته است از استاد فرهیخته و فرزانه جناب آقای دکتر محمود امین طوسی که با کرامتی چون خورشید،
سرزمین دل را روشنی بخشیدند و گلشن سرای علم و دانش را با راهنمایی های کار ساز و سازنده بارور ساختند؛
تقدیر و تشکر نمایم.
و از پدر و مادر عزیز، دلسوز و مهربانم که آرامش روحی و آسایش فکری فراهم نمودند تا با حمایت های همه جانبه
در محیطی مطلوب، مراتب تحصیلی و نیز پایان نامه درسی را به نحو احسن به اتمام برسانم؛ سپاسگزاری نمایم.
همچنین، از بهترین همراهان زندگییم خواهرانم به پاس عاطفه سرشار و گرمای امید بخش وجودشان، که بهترین
پشتیبان من بودند؛ سپاسگزاری می نمایم.
و در نهایت، همه آنانی که در راه کسب دانش راهنمایم بودند و نفس خیرشان و دعای روح پرورشان بدرقه راهم
بود.
امروز هستی ام به امید شماسست و فردا کلید باغ بهشتم رضای شما. ره آوردی گران سنگ تر از این ارزان نداشتم
تا به خاک پایتان نثار کنم، باشد که حاصل تلاشم، نسیم گونه غبار خستگیان را بزدايد و این خردترین، بخشی از
زحماتتان را سپاس گوید.

فرزاد زندی

تیر ۱۳۹۷

فهرست مطالب

ج	فهرست جداول
د	فهرست تصاویر
۱	چکیده
۲	پیش‌گفتار
۴	فصل ۱: مقدمه، معرفی و کارهای مرتبط
۵	۱-۱ تعاریف
۸	۲-۱ معرفی درهم تنیدگی تصویر
۱۱	۳-۱ کارهای مرتبط
۱۳	۴-۱ فرم بسته برای درهم تنیدگی تصویر
۱۳	۱-۴-۱ تصاویر خاکستری
۱۵	۲-۴-۱ تصاویر رنگی
۱۸	۵-۱ چندین نزدیکترین همسایه برای درهم تنیدگی تصویر
۱۸	۱-۵-۱ درهم تنیدگی مبتنی بر نزدیکترین همسایگی
۱۹	۱-۱-۵-۱ محاسبه ماتریس شباهت A با استفاده از نزدیکترین همسایگی
۱۹	۲-۱-۵-۱ بردار ویژگی X با مختصات مکانی
۱۹	۳-۱-۵-۱ حل فرم بسته با پیاده سازی سریع
۲۱	۴-۱-۵-۱ جمع آوری ویژگی
۲۳	فصل ۲: درهم تنیدگی تصویر با استفاده از مجموعه نمونه های وسیع
۲۴	۱-۲ بدست آوردن مجموعه نمونه های وسیع
۲۵	۲-۲ انتخاب نمونه های کاندید
۲۶	۳-۲ مشکل هم پوشانی توزیع های رنگ

۲۶	انتخاب بهترین جفت نمونه کاندید (F, B)	۴-۲
۲۸	پیش پردازش و پس پردازش	۵-۲
۳۰	فصل ۳: درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی	
۳۱	جریان اطلاعات ترکیب رنگ	۱-۳
۳۳	جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص	۲-۳
۳۶	پیش پردازش نقشه سه گانه	۳-۳
۳۷	جریان اطلاعات درون ناحیه نامشخص	۴-۳
۳۸	جریان اطلاعات محلی	۵-۳
۳۹	سیستم خطی و مینیمم سازی انرژی	۶-۳
۴۱	تنظیم برای روشهای برش بر پایه نمونه برداری	۷-۳
۴۳	نکات پیاده سازی	۸-۳
۴۷	فصل ۴: نتایج پیاده سازی	
۴۷	بحث و مقایسه	۱-۴
۴۹	طبقه بندی مبتنی بر درهم تنیدگی	۲-۴
۵۰	نتیجه گیری	۳-۴
۵۲	فهرست منابع	
	پیوست آ: کدهای مربوط به پیاده سازی روش درهم تنیدگی تصویر مبتنی بر جریان اطلاعات	
۵۵	بین پیکسلی در متلب برگرفته از سایت www.alphamatting.com	
	پیوست ب: کدهای مربوط به پیاده سازی طبقه بندی با استفاده از روش درهم تنیدگی جریان	
۷۱	اطلاعات بین پیکسلی در متلب	
۷۶	واژه نامه فارسی به انگلیسی	
۷۹	واژه نامه انگلیسی به فارسی	

فهرست جداول

۴۸	۱-۴	مقایسه روش جریان اطلاعات با چند روش قوی درهم تنیدگی توسط ۸ نمونه
۵۱	۲-۴	دقت برای ۳۰ نمونه تست
۵۱	۳-۴	زمان اجرا برای ۳۰ نمونه تست
۵۱	۴-۴	دقت برای ۶۰ نمونه تست
۵۱	۵-۴	زمان اجرا برای ۶۰ نمونه تست

فهرست تصاویر

- ۱-۱ نمونه عمل درهم‌تنیدگی تصویر ۸
- ۲-۱ خروجی حاصل از اعمال چند روش درهم‌تنیدگی بر روی تصویری هموار با جواب مشخص . ۱۰
- ۳-۱ نمونه اعمال شده چندین فرم از جریان اطلاعات ۱۱
- ۴-۱ مقایسه ماتریس شباهت غیر محلی A در درهم‌تنیدگی KNN و درهم‌تنیدگی غیر محلی . . ۲۰
- ۱-۲ تقسیم بندی ناحیه پیش زمینه در روش درهم‌تنیدگی مجموعه نمونه های وسیع ۲۴
- ۲-۲ مشکل همپوشانی نمونه ها ۲۵
- ۳-۲ مقایسه چند روش نمونه برداری ۲۷
- ۴-۲ مقایسه چند روش نمونه برداری و نمایش نمونه های از دست رفته ۲۸
- ۵-۲ مقایسه روش درهم‌تنیدگی مجموعه نمونه های وسیع با چند روش مشهور ۲۹
- ۱-۳ جریان ترکیب رنگ به تنهایی ۳۳
- ۲-۳ جریان اطلاعات \mathcal{K} -to- \mathcal{U} و ضریب اعتماد ۳۴
- ۴-۳ جریان اطلاعات \mathcal{K} -to- \mathcal{U} برای اشیاء شفاف ۳۶
- ۵-۳ نمایش عملکرد پیش پردازش ۳۷
- ۶-۳ پیش پردازش نقشه سه گانه ۳۸
- ۹-۳ مقایسه تنظیم سازی روش ما و روش گاستال و الیوپرا ۴۲
- ۱-۴ استفاده از گام تنظیم سازی برای ۳ روش بر پایه نمونه برداری ۴۸
- ۲-۴ مقایسه روش جریان اطلاعات با روش های فرم بسته و نزدیکترین همسایگی ۴۹
- ۳-۴ مقایسه چندین روش درهم‌تنیدگی با معیارهای درهم‌تنیدگی آلفا ۵۰



دانشگاه گیلان

فرم چکیده پایان نامه دوره تحصیلات تکمیلی

مدیریت تحصیلات تکمیلی

نام خانوادگی دانشجو: زندی	نام: فرزاد	ش. دانشجویی: ۹۵۱۳۱۳۷۰۴۸
استاد راهنما: دکتر محمود امین طوسی		
استاد مشاور: دکتر علیرضا قدسی		
دانشکده ریاضی و علوم کامپیوتر	رشته: علوم تصمیم و مهندسی دانش	
مقطع: کارشناسی ارشد	تاریخ دفاع: تیر ۱۳۹۷	تعداد صفحات: ۸۲
عنوان پایان نامه: درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی		
کلید واژه‌ها: درهم تنیدگی، جریان اطلاعات، ناحیه مشخص و نامشخص، نزدیکترین همسایگی		
<p>چکیده: استخراج اشیاء پیش زمینه از تصاویر نقش مهمی در ویرایش تصویر و ویدئو بازی می‌کند. جداسازی دقیق پیش زمینه و پس زمینه یک مسأله مهم طبقه بندی یادگیری نیمه نظارتی است که مرتبط با موضوع پردازش تصویر می‌باشد. یادگیری نیمه نظارتی یعنی براساس محدودیت های ورودی کاربر مسأله مدلسازی می‌شود. یکی از روشهای نسبتاً جدید در این حوزه، روش های موسوم به درهم تنیدگی تصویر یا Image Matting می‌باشد که تحقیقات متعددی را به خود معطوف نموده است.</p> <p>در این تحقیق چندین روش درهم تنیدگی تصویر معرفی می‌شوند. یکی از روشهای کارا درهم تنیدگی تصویر، روش درهم تنیدگی تصویر مبتنی بر اطلاعات بین پیکسلی میباشد که هدف اصلی این پایان نامه است. همچنین با استفاده از شباهتهایی که بین پیکسلها برقرار و تعریف میشوند جریانهای مختلف اطلاعات را توضیح داده و با ترکیب نهایی آنها این روش شکل می‌گیرد. از آنجا که در روش های متعدد درهم تنیدگی ضعف هایی دیده شده است، روش فوق الذکر می‌تواند با تنظیم پارامترهای معادلاتش به عنوان یک روش پس پردازش برای بهبود نقاط ضعف ها در آمده و استفاده شود.</p> <p>در نهایت از روش جریان اطلاعات بین پیکسلی به عنوان یک کار جدید برای طبقه بندی چندین مجموعه داده غیر تصویری با نمونه های آموزشی متفاوت و ویژگی های مختلف استفاده شده و نتیجه آن با روش ماشین بردار پشتیبان (SVM) و روش چند همسایگی (KNN) مقایسه شده که در بعضی از مجموعه داده ها نتایج بهتری را نیز کسب کرده است.</p>		

پیش‌گفتار

استخراج اشیاء پیش‌زمینه از تصاویر نقش مهمی در ویرایش تصویر و ویدئو بازی می‌کند که به طور گسترده بیش از ۲۰ سال است که مورد مطالعه قرار گرفته است. یکی از مرسوم‌ترین و قدیمی‌ترین روش‌های مونتاز فیلم، استفاده از زمینه ثابت یا تکنیک پرده آبی است که اشیاء موردنظر را از یک تصویر برداشته و در یک زمینه دیگر قرار می‌دهند. روش‌های بسیاری برای درهم‌تنیدگی تصویر با رویکردهای مختلفی عرضه شده‌اند. در این پایان‌نامه روش قوی درهم‌تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی [۱] که از بهترین و کم‌خطاترین روش‌های درهم‌تنیدگی تصویر است توضیح، تحلیل و تفسیر می‌شود. همچنین این روش با روش‌های دیگر درهم‌تنیدگی مقایسه شده است.

این نوشتار شامل ۴ فصل است؛

در فصل ۱ تعاریف مورد نیاز اولیه ذکر گردیده و سپس چندین روش مرسوم در قسمت‌های مختلف این نوشتار به آنها ارجاع داده شده بیان شده‌اند.

در بیان روش اصلی مورد بحث (جریان اطلاعات بین پیکسلی) از روش مجموعه نمونه‌های وسیع [۲] استفاده شده که فصل ۲ به بیان تفصیلی آن اختصاص داده شده است.

فصل ۳ روش درهم‌تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی [۱] را به طور کامل توضیح می‌دهد. در این فصل فرم‌های مختلف جریان اطلاعات در بخش‌های جداگانه تعریف، توضیح و بررسی می‌گردند و جریان‌های مختلف توضیح داده شده به صورت معادلاتی واحد برای درهم‌تنیدگی ارائه می‌شوند. در انتهای فصل به تشریح نحوه تنظیم این روش برای استفاده در روش‌های بر پایه نمونه برداری برای بهبود کارایی آنها پرداخته می‌شود.

در فصل ۴ به بحث، نتیجه‌گیری و مقایسه روش اصلی مورد بحث با چند روش انجام شده در قبل پرداخته و نتایج آورده شده‌اند. به عنوان یک کار جدید این روش بر روی داده‌های غیر تصویری اعمال شده است و نتایج آن با روش مشهور طبقه‌بندی مورد مقایسه قرار گرفته است که در بعضی از مجموعه داده‌ها نتایج بهتری کسب گردیده است.

مقالات اصلی مورد استفاده در این پایان نامه به قرار زیر است ؛

1. Aksoy, Yagız, Aydın, Tunç Ozan, Pollefeys, Marc, and Zürich, ETH. Designing effective inter-pixel information flow for natural image matting. in *Computer Vision and Pattern Recognition (CVPR)*, 2017.
2. Levin, Anat, Lischinski, Dani, and Weiss, Yair. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):228–242, 2008.
3. Chen, Qifeng, Li, Dingzeyu, and Tang, Chi-Keung. Knn matting. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2175–2188, 2013.
4. Shahrian, Ehsan, Rajan, Deepu, Price, Brian, and Cohen, Scott. Improving image matting using comprehensive sampling sets. in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 636–643. IEEE, 2013.

فصل ۱

مقدمه، معرفی و کارهای مرتبط

در یادگیری ماشینی و آمار، طبقه‌بندی^۲، مسأله شناسایی تعلق یک مشاهده جدید به یکی از دسته‌های از قبل مشخص شده، است. در اصطلاح یادگیری ماشینی، طبقه‌بندی نوعی یادگیری با نظارت^۳ است که مجموعه داده‌هایی برای آموزش موجود است. بخشی از این داده‌های آموزشی شامل مشاهداتی است که عضویت در دسته‌هایشان معلوم می‌باشد. طبقه‌بندی یکی از مهمترین مباحث پایه‌ای در زمینه‌های آمار^۴، یادگیری ماشینی^۵، شناسایی الگو^۶ و داده‌کاوی^۷ است.

در الگوریتم‌های طبقه‌بندی مجموعه داده اولیه به دو مجموعه داده با عنوان مجموعه داده‌های آموزشی^۸ و مجموعه داده‌های آزمایشی^۹ تقسیم می‌شود، با استفاده از مجموعه داده‌های آموزشی مدل ساخته می‌شود و از مجموعه داده آزمایشی برای اعتبار سنجی و محاسبه دقت مدل ساخته شده، استفاده می‌شود. در الگوریتم‌های طبقه‌بندی ویژگی طبقه مربوط به هر رکورد مشخص است بنابراین جزء الگوریتم‌های با نظارت محسوب می‌شوند. الگوریتم‌های با نظارت شامل دو مرحله با عنوان آموزش (یادگیری) و مرحله ارزیابی (تست) هستند [۳].

یادگیری نیمه‌نظارتی^{۱۰} دسته‌ای از روش‌های یادگیری ماشینی است که در آن از داده‌های بدون برچسب و داده‌های برچسب‌دار به صورت همزمان برای بهبود دقت یادگیری استفاده می‌شود.

روش درهم‌تنیدگی^{۱۱} یک مسأله قطعه‌بندی نیمه‌نظارتی است که یک تصویر را به دو لایه پیش‌زمینه و پس‌زمینه تجزیه می‌کند. اگر میزان شدت روشنایی یک پیکسل را I ، میزان مشارکت پیکسل پیش‌زمینه در تصویر اصلی را F و میزان مشارکت پیکسل پس‌زمینه در تصویر اصلی را B بنامیم؛ آن‌گاه هدف، بدست آوردن میزان α به نحوی است که ترکیب خطی I مطابق معادله زیر بدست آید:

$$I = \alpha F + (1 - \alpha)B \quad (1-1)$$

^۲Classification ^۳Supervised ^۴Statics ^۵Machine Learning ^۶Pattern Recognition

^۷Data Mining ^۸Train Dataset ^۹Test Dataset ^{۱۰}Semi Supervised ^{۱۱}Image

Matting

با توجه به داشتن تنها یک تصویر، سه مقدار F, B, α مقادیری نامعلوم هستند و باید این سه مقدار در هر پیکسل تعیین شوند. در تصاویر خاکستری مدل رنگ یک بعدی است و بنابراین برای هر پیکسل باید یک مقدار بدست آید که طبق معادله ۱-۱، ۳ مقدار نامعلوم وجود دارد. این در صورتی است که تنها یک معادله موجود می باشد. این مسأله در مدل رنگ سه بعدی (RGB) ^۱ به این صورت است که برای هر پیکسل باید سه مقدار را بدست آید؛ که با توجه به اینکه برای هر پیکسل سه لایه موجود می باشد هر پیکسل دارای ۷ مقدار نامعلوم است، در صورتیکه تنها سه معادله موجود می باشد. در هر دو مدل ذکر شده، در سیستم معادلات، بی نهایت جواب موجود می باشد.

۱-۱ تعاریف

پیکسل ^۲: کوچکترین جزء هر تصویر پیکسل است. پیکسل مخفف picture element به معنای المان تصویر است. یک تصویر متشکل از تعداد زیادی پیکسل است که در کنار یکدیگر قرار گرفته و هرکدام از این نقاط دارای مکان و مقدار مشخصی هستند.

برای نمایش یک تصویر $M \times N$ از یک ماتریس که M سطر و N ستون دارد، استفاده می شود. مقدار هر عنصر ماتریس نشان دهنده شدت روشنایی تصویر در آن نقطه است. هر عنصر می تواند مقداری بین ۰ و ۲۵۵ داشته باشد. مقدار صفر نشان دهنده رنگ سیاه و ۲۵۵ نشان دهنده رنگ سفید می باشد.

یک تصویر RGB متشکل از سه ماتریس است که هر یک از آن ها مقادیر قرمز (Red) و سبز (Green) و آبی (Blue) تصویر رنگی را نگه می دارند. یک پیکسل زمانی مقدار خاکستری خواهد داشت که مؤلفه های R,G,B آن مقدار یکسانی داشته باشند.

معمولاً از لغت خاکستری ^۳ در ارتباط با شدت روشنایی ^۴ تصاویر تک رنگ استفاده می شود. تصاویر رنگی از ترکیب تصاویر تک رنگ تشکیل می شوند. به عنوان مثال در سیستم رنگی RGB، تصویر رنگی از سه تصویر تک رنگ (قرمز، سبز و آبی) تشکیل می شود.

فضای رنگ ^۵: یک فضای رنگ، سازماندهی خاصی از رنگ ها در یک نرم افزار یا ابزار است. گستره فضاهای رنگ با یکدیگر متفاوت است. به طوری که ممکن است یک فضای رنگ، بر یک فضای دیگر کاملاً محیط باشد؛ یا ممکن است اشتراک داشته باشند. یک فضای رنگ در تعامل با ابزارهای فیزیکی دیگر و پروفایل رنگی امکان باز تولید و باز نمایش صحیح رنگها را فراهم می کند [۴].

فضای رنگ RGB: رایج ترین فضای رنگ مورد استفاده در گرافیک رایانه ای و نمایشگرها، فضای رنگ RGB است که از حروف ابتدای رنگ های قرمز Red، سبز Green و آبی Blue گرفته شده است. مبنای مدل RGB، ترکیب تجمعی ^۶ است به این معنی که از ترکیب مقادیر مختلف سه رنگ قرمز، سبز و آبی یک رنگ دلخواه به وجود

^۱Red,Green,Blue ^۲Pixel ^۳Gray ^۴Intensity ^۵Color Space ^۶Additive Combination

می آید. این استاندارد در سال ۱۹۳۱ میلادی از سوی موسسه CIE معرفی شد [۴].
 فام^۱: احساس فرد از شباهت بین دو ناحیه رنگی توسط این متغیر تعیین می شود. در واقع، فام مشخص کننده رنگ غالبی است که چشم بیننده دریافت می کند [۴].
 اشباع^۲: میزان پررنگی یک ناحیه نسبت به درخشندگی آن یا به عبارت دیگر خلوص نسبی فام را تعیین می کند [۴].

فضای رنگ HSV^۳: فضای رنگ RGB برای کاربردهایی مانند نمایش تصویر و چاپ تصاویر رنگی بسیار مناسب است اما در آن مستقیماً معیاری برای سنجش سه مولفه اشباع، فام و روشنایی وجود ندارد، در حالیکه بهترین شیوه توصیف رنگ از دید انسان بر اساس همین سه مولفه است. از این رو فضای رنگ HSV که شبیه ترین فضا به سیستم بینایی انسان است بوجود می آید و در کاربردهای پردازش تصاویر رنگی نیز مورد استفاده قرار می گیرد. مولفه های این فضای رنگ بر اساس طول موج فام (H)، درجه اشباع رنگ (S) و درجه روشنایی یا میزان شدت رنگ (V) تعریف می شود [۴].

فاصله اقلیدسی^۴: در ریاضیات، فاصله اقلیدسی، فاصله معمولی بین دو نقطه است. در ادبیات قدیمی تر به عنوان مترادف فاصله فیثاغورث اشاره شده است. طبق این نوع از شاخص اندازه گیری، فاصله بین دو نقطه p, q در فضای \mathbb{R}^n به صورت زیر محاسبه می شود:

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

فاصله باتاچاریا^۵: در آمار فاصله باتاچاریا، فاصله شباهت بین دو توزیع احتمال تعریف می شود. ضریب باتاچاریا، اندازه ای از مقدار هم پوشانی بین دو نمونه آماری یا دو جمعیت است. فاصله باتاچاریای بین دو کلاس با توزیع نرمال به صورت زیر محاسبه می شود:

$$D_B(p, q) = \frac{1}{4} \ln \left(\frac{1}{4} \left(\frac{\sigma_p^2}{\sigma_q^2} + \frac{\sigma_p^2}{\sigma_q^2} + 2 \right) \right) + \frac{1}{4} \left(\frac{(\mu_p - \mu_q)^2}{\sigma_p^2 + \sigma_q^2} \right)$$

$D_B(p, q)$ فاصله باتاچاریای بین توزیع های p یا کلاس های p و q است. σ_p^2 واریانس توزیع p ، μ_p میانگین توزیع p و p و q دو توزیع متفاوت هستند.

پیکسل همسایه^۶: یک پیکسل، همسایه پیکسل دیگری تلقی می شود اگر از لحاظ فاصله (مختصات فضایی) در یک فاصله مشخص قرار داشته باشد یا از لحاظ رنگ به آن شبیه باشد.

آستانه گیری^۷: هنگامی که قرار است از یک تصویر رنگی، تصویر سیاه و سفید ساخته شود باید تک تک پیکسل ها با مقدار مشخصی مقایسه شده اگر بزرگ تر از آن مقدار باشند به جای آنها رنگ سفید و اگر کوچک تر باشند به

^۱Hue ^۲Saturation ^۳Hue, Saturation, Value ^۴Euclidean Distance ^۵Bhattacharyya Distance
^۶Neighbor Pixel ^۷Threshold

جای آنها رنگ سیاه قرار می‌گیرد، آن مقدار مشخص آستانه می‌باشد.

هیستوگرام تصویر^۱: هیستوگرام در واقع همان نمودار میله‌ای افزونی هر رنگ می‌باشد. از هیستوگرام برای آستانه‌گیری استفاده می‌شود [۴].

نقشه سه گانه^۲: در مسائل درهم‌تیدگی تصویر، کاربر هر تصویر را به سه بخش تقسیم می‌کند:

\mathcal{F} : پس زمینه

\mathcal{B} : پیش زمینه

\mathcal{U} : بدون برچسب (نامشخص)

این تقسیم بندی در شکل ۱-۱ نشان داده شده است. رنگ سفید پیش زمینه \mathcal{F} ($\alpha = 1$)، رنگ سیاه پس زمینه \mathcal{B} ($\alpha = 0$) و رنگ خاکستری ناحیه نامشخص \mathcal{U} ($\alpha = ?$) را نشان می‌دهد.

اعتبارسنجی متقابل^۳: در این نوع اعتبارسنجی، داده‌ها به K زیرمجموعه افراز می‌شوند. از این K زیرمجموعه هر بار یکی برای اعتبارسنجی و $K-1$ تای دیگر برای مدل سازی به کار می‌روند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یک بار برای مدل سازی و یک بار برای اعتبارسنجی به کار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به عنوان تخمین نهایی برگزیده می‌شود.

مسائل بد طرح و خوش طرح: در اوایل قرن پیش، هادامارد^۴ اولین کسی بود که بحث دستگاه‌های معادلات خوش طرح^۵ (خوش وضع) و بد طرح^۶ (بد وضع) را مطرح نمود. او دستگاه معادلاتی را که در آن‌ها سه شرط وجود، یکتایی و پایداری جواب تأمین باشند، را مسائل خوش طرح و مسائلی که در آن‌ها یکی از سه شرط نقض گردیده باشد، را مسائل بد طرح نام گذاری کردند.

درک صحیح معیار درهم‌تیدگی: در دسترس بودن معیارهای کمی آنلاین برای مقایسه نتایج روش‌های مختلف درهم‌تیدگی تصویر اهمیت زیادی در پیشرفت زمینه‌های مربوطه داشته است. سه فاکتور برای یک سیستم موفق سنجش معیار وجود دارد:

۱- مجموعه تستی چالش برانگیز از زمینه‌های واقعی با کیفیت بالا.

۲- یک مخزن ارزیابی آنلاین که به صورت پویا با نتایج جدید به روز شده باشد.

۳- چند معیار مقایسه مناسب.

مرجع [۲۹] تلاش کرده است تا سه معیار فوق را برآورده کند. در سایت www.alphamatting.com مجموعه داده‌هایی شامل تصاویر چالش برانگیز و نقشه‌های سه گانه آنها با کیفیت بالا و پایین موجود می‌باشند و روش‌های درهم‌تیدگی تا کنون معرفی شده از لحاظ معیارهای مجموع قدر مطلق تفاوتها (SAD) و میانگین مربعات خطا (MSE) با هم مقایسه شده‌اند.

^۱Image Histogram ^۲Trimap ^۳Cross Validation ^۴Hadamard ^۵Well-posed

^۶Ill-posed



شکل ۱-۱: نمونه عمل درهم‌تنیدگی تصویر

۲-۱ معرفی درهم‌تنیدگی تصویر

استخراج کدری (مقدار α) اطلاعات از اشیاء پیش زمینه یک تصویر به عنوان درهم‌تنیدگی تصویر شناخته میشود. درهم‌تنیدگی تصویر در دهه اخیر علاقه مندان زیادی از جامعه محققین به خود اختصاص داده و این روزها به عنوان یکی از مسایل کلاسیک تحقیق مطرح می‌شود.

در شکل ۱-۱ (آ) تصویر یک گل‌دان نشان داده شده است. هدف از درهم‌تنیدگی تصویر در این شکل جدا کردن گل‌دان و برگ‌های گل از پس زمینه می‌باشد. برای این منظور نیاز به یک نقشه سه گانه مطابق با شکل ۱-۱ (ب) است. بعد از اعمال درهم‌تنیدگی تصویر توسط یکی از روش‌های موجود تصویر ۱-۱ (ج) حاصل می‌شود. از نظر ریاضی درهم‌تنیدگی تصویر نیازمند بیان رنگ‌های پیکسل در نواحی تغییر و انتقال از پیش زمینه به پس زمینه در ترکیب محدب رنگ‌های پیش زمینه و پس زمینه آنها است. وزن یا کدری رنگ پیش زمینه مقدار آلفای آن پیکسل معرفی می‌شود. از آنجا که کدری رنگ‌های پیکسل‌های پیش زمینه و پس زمینه مشخص نیستند، مشخص کردن مقادیر کدری پیکسلها مسأله‌ای بد طرح است. برای کم کردن سختی این مسأله عموماً یک نقشه سه گانه به تصویر اصلی اضافه می‌شود. نقشه سه گانه تقسیم بندی ناهموار تصویر ورودی به پیش زمینه، پس زمینه و نواحی با کدری نامشخص است.

مسأله درهم‌تنیدگی، برآورد \mathcal{F}, B, α در مناطق بدون برچسب است. اگرچه درهم‌تنیدگی یک مسأله بد وضع است ولی همبستگی قوی بین پیکسل‌های نزدیک به هم می‌تواند آهرمی برای کاهش مشکل بد وضعی باشد. به طور معمول فرض بر این است، پیکسل‌هایی که دارای رنگ‌های مشابه هستند اغلب پارامترهای درهم‌تنیدگی (مانند \mathcal{F}, B, α) مشابهی دارند. این همبستگی در بسیاری از مسائل حوزه پردازش تصویر مانند حذف نویز، تفکیک پذیری، رنگ بندی، تقسیم بندی و ... استفاده شده است [۵].

هنگامی که دو پیکسل همسایه به هم شبیه باشند مقادیر پارامترهای \mathcal{F} و B و I آن‌ها به هم نزدیک است در نتیجه از یک پیکسل شناخته شده می‌توان میزان α یک پیکسل نامشخص مشابه را با استفاده از روش‌های متعددی تخمین زد. از آنجا که پایه اصلی بیشتر روش‌های درهم‌تنیدگی تصویر تخمین پارامترهای مجهول هر پیکسل از روی نمونه‌های پیکسل‌های مشابه آن است، از این روش‌ها زیر مطرح می‌شوند:

چگونه باید پیکسل همسایه پیدا شود؟ به عبارتی؛ در چه فاصله ای پیکسل همسایه پیش زمینه و پس زمینه به درستی قابل شناسایی است؟ به چه صورت پیکسل های همسایه جمع آوری می شوند؟ چگونه می توان B و F مربوط به هر پیکسل را به درستی برآورد کرد؟

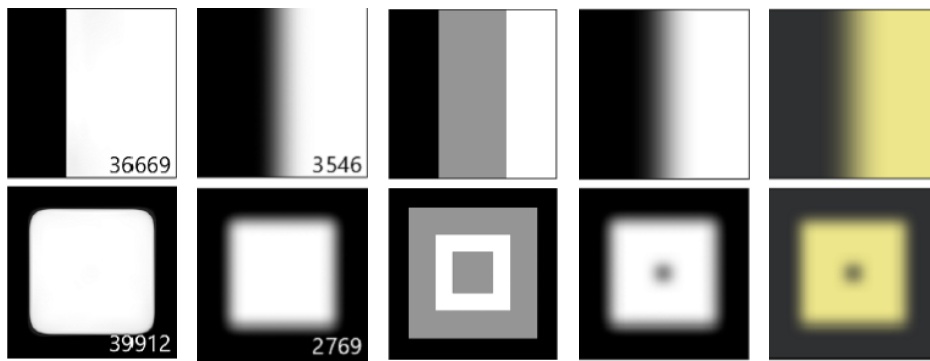
روش های متفاوتی برای پاسخ به این سوالات وجود دارد که روش مورد بررسی در این نوشتار (جریان اطلاعات بین پیکسلی) یکی از آنها می باشد. این پایان نامه یک الگوریتم جدید درهم تیدگی تصویر بر پایه شباهت را تشریح می کند که دقیقاً بر اساس شباهت های پیکسل به پیکسل که به طور مؤثر قادر است از اطلاعات موجود در تصویر و نقشه سه گانه استفاده نماید تعریف شده است. در روشی که در فصل ۳ تشریح می شود، جریانهای اطلاعاتی که توضیح داده خواهند شد از نواحی مشخص به نامشخص و همچنین داخل خود ناحیه نامشخص بوسیله چندین تعریف شباهت پیکسل تعریف شده، کنترل می شوند. همچنین از میان فرمهای دیگر جریان اطلاعات (که در ۳-۱ تا ۳-۵ توضیح داده می شود)، جریان ترکیب رنگ^۱ معرفی می شود که به طور مؤثر رابطه بین پیکسلهای مختلف را محاسبه می کند.

از چالش های روشهای مختلف درهم تیدگی تصویر می توان به ضعف هایی از قبیل عدم پوشش نواحی دور از ناحیه مشخص و نواحی دارای سوراخ اشاره کرد که روش جریان اطلاعات در مقابل آنها مقاوم است و در بخش های آتی این موضوع با مثال بیان می شود. ارزیابی انجام شده بر روی این روش نسبت به روش های دیگر (حتی بروزترین آنها) نشان دهنده برتری کارایی آن است.

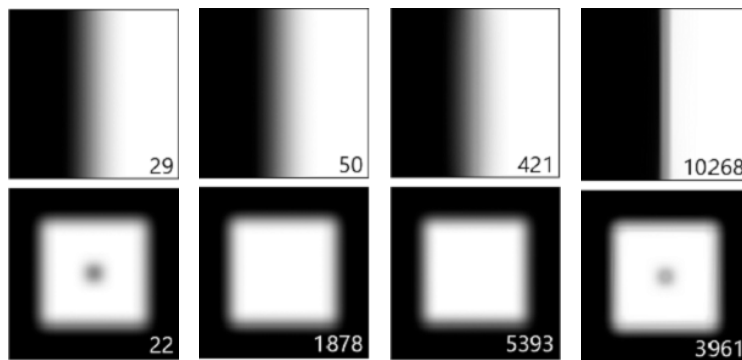
بسیاری از روش های درهم تیدگی تصویر از شباهتهای پیکسل برای گسترش مقادیر آلفا از نواحی مشخص به نواحی نامشخص استفاده کرده اند [۷، ۸، ۹]. آنها یک معادله ریاضی که می تواند در فرم بسته حل شود و عموماً درهم تیدگی ثابتی را ارایه می دهد تهیه کرده اند. روشهای ذکر شده فوق برای بکار بردن مؤثر نواحی با اندازه گرادیان های بزرگ و فضاهای نواحی منفصل نامشخص حتی در حالات ساده نشان داده شده در شکل ۱-۲ شکست خورده اند. علت این است که فقط استفاده از شباهت پیکسل، نمی تواند به خوبی ساختارهای پیچیده که عموماً در تصاویر دیده می شوند را نشان دهد. به منظور کم نمودن کاستی های فوق، به یک طراحی دقیق از اینکه چگونه مقادیر آلفا باید درون تصویر گسترش پیدا کنند نیاز است. به این گسترش مقادیر آلفا جریان اطلاعات اطلاق می شود.

شکل ۱-۲ (آ) دو تصویر 500×500 پیکسل را که برای داشتن نواحی ای با تغییرات نرم به صورت مات در آمده اند نشان داده است. شکل ۱-۲ (ب) جواب مرجع می باشد که برای مقایسه با تصویر حاصل از عمل درهم تیدگی با روش های مختلف بکار می رود. تصویر ۱-۲ (ج) نقشه سه گانه تهیه شده را نشان می دهد. در تصاویر ۱-۲ (د) تا ۱-۲ (ط) به ترتیب نتایج اعمال روش های فرم بسته [۷]، چند همسایگی با فضای رنگ HSV [۸]، چند همسایگی با فضای رنگ RGB [۸]، حفظ خمینه [۹] و جریان اطلاعات فقط با استفاده از جریان ترکیب رنگ و جریان اطلاعات [۱] نشان داده شده است. عددهای درون تصاویر حاکی از مجموع قدر مطلق تفاوت بین آلفای تخمین زده شده و آلفای واقعی است. بهبود کارایی حتی وقتی که فقط از جریان ترکیب رنگ استفاده شده

^۱Color Mixture Flow



(آ) تصویر اصلی (ب) زمینه واقعی (ج) نقشه سه گانه (د) فرم بسته [۷] (ه) HSV [۸]



(و) RGB [۸] (ز) M.P. [۹] (ح) CMF [۱] (ط) جریان اطلاعات [۱]

شکل ۱-۲: خروجی حاصل از اعمال چند روش درهم تنیدگی بر روی تصویری هموار با جواب مشخص

دیده می شود (بخش ۳-۱).

ایده اصلی مرجع [۱]، یک استراتژی جدید کنترل جریان اطلاعات هم از نواحی مشخص نقشه سه گانه به ناحیه نامشخص و هم داخل خود ناحیه نامشخص است. این استراتژی از طریق استفاده از تعاریف گوناگون شباهت فرموله می شود. انواع جریانهای اطلاعات که در این روش استفاده می شود و در فصل ۳ مفصل بیان خواهند شد به شرح ذیل است:

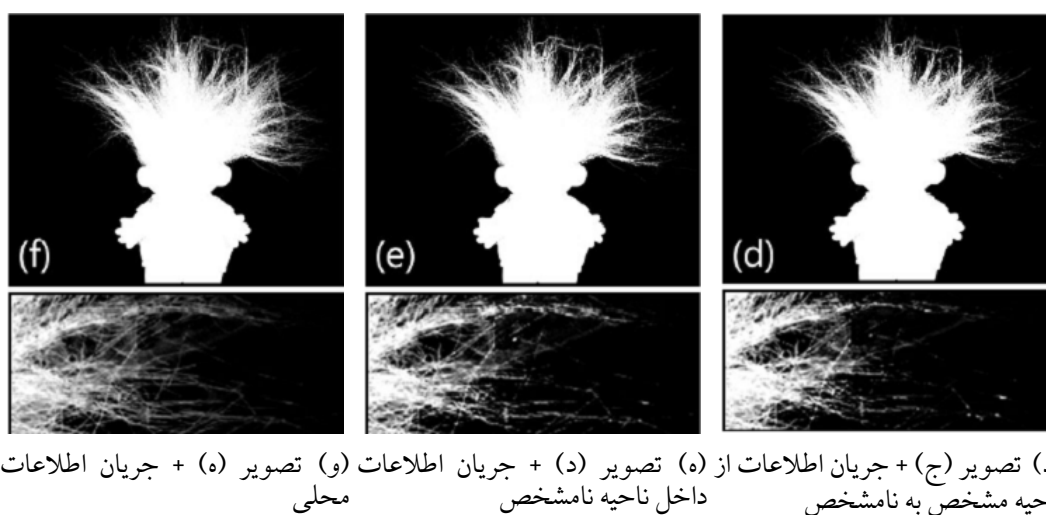
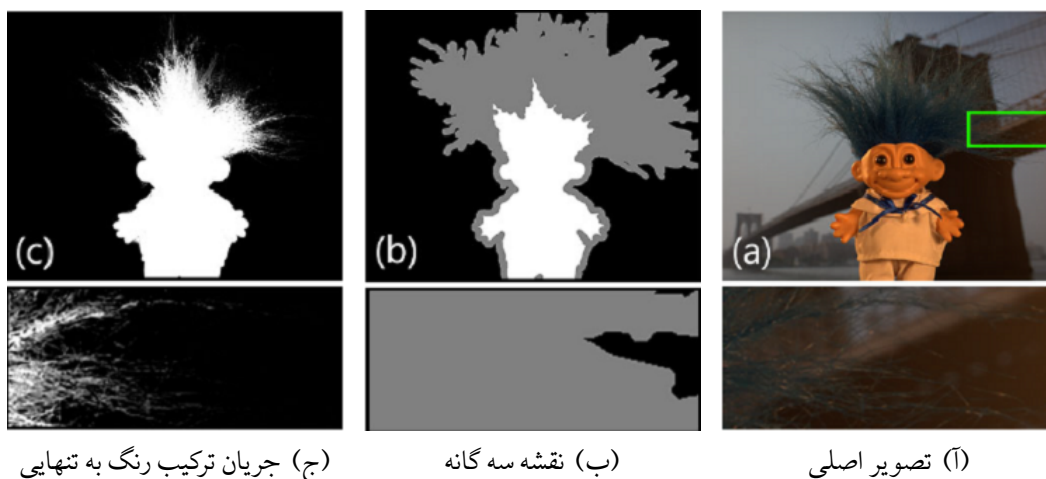
۱- جریان اطلاعات ترکیب رنگ (بخش ۳-۱)

۲- جریان اطلاعات از ناحیه مشخص به نامشخص (بخش ۳-۲)

۳- جریان اطلاعات درون ناحیه نامشخص (بخش ۳-۴)

۴- جریان اطلاعات محلی (بخش ۳-۵)

تاثیر استفاده از فرم های جریان اطلاعات در شکل ۱-۳ نشان داده شده است. در تصاویر ۱-۳ (ج) تا ۱-۳ (و) نتیجه اعمال هر جریان به فرم قبلی نشان داده شده است که می توان بهبود کیفیت در تصاویر بزرگ شده را مشاهده کرد.



شکل ۱-۳: نمونه اعمال شده چندین فرم از جریان اطلاعات

۳-۱ کارهای مرتبط

عموم روشهای درهم تنیدگی در نوشتارها عمدتاً به روشهای بر اساس نمونه برداری یا بر اساس شباهت تفکیک می شوند. در این بخش به طور خلاصه روشهایی مرور می شوند که با روش درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی بیشترین ارتباط را دارند و خواننده را به سوی بررسی جامع برای اطلاعات بیشتر هدایت میکنند.

روشهای بر اساس نمونه برداری [۲، ۲۳، ۲۴، ۲۵] عموماً به دنبال جمع آوری چندین نمونه از نواحی پیش زمینه و پس زمینه مشخص شده در نقشه سه گانه و انتخاب بهترین جفت با توجه به معیارهای تعریف شده برای نمایش یک پیکسل نامشخص به عنوان ترکیبی از پیش زمینه و پس زمینه هستند. تجربه حاصل برای عموم کدهای درهم تنیدگی موجود، حاکی از آن است که روشهای پیاده سازی بر پایه نمونه برداری زمان اجرای بالایی دارند. یکی از روش های تخمین مقدار آلفا، نمونه برداری از برخی رنگهای پیش زمینه و پس زمینه برای تخمین مقدار

آلفای هر پیکسل در ناحیه نامشخص می باشد. روش های موجود درهم تنیدگی مبتنی بر نمونه برداری اغلب فقط نمونه هایی نزدیک به پیکسل های ناحیه نامشخص را جمع آوری می کنند که اگر نمونه های تقریبی خوبی بدست نیاید نتیجه جالب توجهی نخواهند داشت. مرجع [۱۷] به جمع آوری نمونه ها در کل تصویر می پردازد. در این روش یک تابع هزینه ساده اما تأثیر گذار برای برخورد با ابهام های احتمالی در پروسه انتخاب نمونه ها تعریف شده است. برای رسیدگی به پیچیدگی محاسبات به دست آمده به علت وجود تعداد زیاد نمونه ها، مسأله تشابه در نمونه برداری مطرح شده است. مسأله تشابه، تأثیر بدست آمده از تعمیم الگوریتم تصادفی طراحی شده در کارهای قبلی برای تطابق تکه ای است.

مرجع [۲۷] نیز مانند دیگر روش های مبتنی بر نمونه برداری، برای پیکسل های نامشخص نمونه هایی از پیش زمینه و پس زمینه جمع آوری می کند اما با اهمیت بیشتری اعتماد بر این نمونه ها را آنالیز می کند. در این روش فقط نمونه های با اطمینان بالا برای مشارکت در تخمین تابع هدف انتخاب می شوند. همچنین تابع هزینه معرفی شده شامل یک اصطلاح همسایگی برای تخمین های هموار است.

روشهای بر پایه شباهت، اصولاً از ماتریس های تشابه پیکسل که بر شباهت رنگ یا نزدیکی مکانی و گسترش مقادیر آلفا از نواحی با مقدار آلفای مشخص تکیه دارند، استفاده می کنند. تعاریف شباهت محلی^۱ در شباهت درهم تنیدگی برجسته^۲ [۷]، بر اساس یک تکه محلی اطراف موقعیت یک پیکسل برای تعیین مقدار جریان اطلاعات محلی و گسترش مقادیر آلفا بر اساس آن، عمل می کنند. همچنین درهم تنیدگی بر پایه شباهت به طور گسترده به عنوان یک گام پس پردازش در روشهای بر پایه نمونه برداری [۲، ۲۳، ۲۵] که بوسیله گاستال و اویویرا^۳ در [۲۴] مطرح شده است پذیرفته شده اند. این روشها از شباهت غیر محلی با استفاده از تشابه رنگ و نزدیکی مکانی، برای تعیین مقادیر آلفای پیکسلهای مختلف که باید با یکدیگر ارتباط داشته باشند استفاده می کنند.

مرجع [۱۲] سعی در کاهش مقدار برچسب گذاری های کاربر در تصویر ورودی برای انجام عمل درهم تنیدگی دارد. ایده اصلی این روش استفاده از اصل غیر محلی برای حذف نویز تصاویر معرفی شده می باشد که می تواند در بدست آوردن تخمین آلفا به طور موفق اعمال شود. بنابراین به طور چشم گیری تعداد پیکسل هایی که کاربر به صورت دستی برچسب گذاری می کند را کاهش می دهد. همچنین نشان داده است که چگونه از ایجاد ورودی های غیر ضروری اجتناب می شود و یک روش خوشه بندی پیکسل های تصویر برای برچسب گذاری توسط کاربر توسعه داده شده است. این روش نتایج خوبی استخراج می کند.

درهم تنیدگی چند همسایگی [۸] چندین همسایه برای هر پیکسل در ناحیه نامشخص تعیین و آنها را وادار به داشتن مقادیر آلفای مشابه، نسبت به فاصله آنها در فضای ویژگی می کند. الگوریتم مورد اشاره در مرجع [۹] نیز یک مجموعه همسایه برای هر پیکسل تعیین می کند، اما هر پیکسل را در ترکیب خطی همسایه هایش در فضای ویژگی آنها نشان می دهد.

چن و همکاران^۴ در [۲۶] روشی ترکیبی از روش قدرتمند درهم تنیدگی بر پایه نمونه برداری [۲۷] به عنوان

^۱Local Affinity

^۲Prominently the Matting Affinity

^۳Gastal and Oliveira

^۴Chen

یک نقطه شروع استفاده کرده و نتایج آن را از طریق تکنیکی بر پایه گراف که ترکیب یک شباهت غیر محلی [۹] و روش درهم تنیدگی بر پایه شباهت است تصحیح کرده و مطرح کرده اند. چو و همکاران در [۲۸] نتایج درهم تنیدگی فرم بسته [۷] و درهم تنیدگی چند همسایگی [۸] و همچنین روش جامع بر پایه نمونه برداری را بوسیله بازخورد در یک شبکه عصبی پیچیده ترکیب کرده اند.

در ادامه دو روش مشهور در حوزه درهم تنیدگی تصویر تشریح می شوند:

۱- فرم بسته برای درهم تنیدگی تصویر [۷].

۲- چندین نزدیکترین همسایه برای درهم تنیدگی تصویر [۸].

۴-۱ فرم بسته برای درهم تنیدگی تصویر

این بخش روش رسیدن به یک جواب فرم بسته را برای به دست آوردن ضریب آلفا از رابطه ۱-۱ ارائه می دهد و روش فرم بسته برای درهم تنیدگی تصویر به طور کامل تشریح نمی شود و فقط به مواردی که به هنگام استفاده از فرم جریان اطلاعات محلی که در بخش ۳-۵ برای بیان مسأله لازم بوده است اکتفا شده است. برای روشن شدن مطلب، ابتدا یافتن یک جواب فرم بسته مسأله تخمین آلفا برای تصاویر خاکستری توضیح داده شده و سپس این جواب ها برای تصاویر رنگی تعمیم داده خواهد شد [۷].

۱-۴-۱ تصاویر خاکستری

یادآوری می شود هر پیکسل ترکیب خطی از زمینه و شیء تصویر مورد نظر است، یعنی

$$I = \alpha F + (1 - \alpha)B$$

که α در بازه صفر و یک است.

همان گونه که ذکر شد برای حل مسأله درهم تنیدگی تصویر، برخی فرضیات معمول F, B, α مورد نیاز است. برای رسیدن به جواب برای ناحیه نامشخص، فرض می شود F و B هر دو روی پنجره کوچکی اطراف هر پیکسل تقریباً ثابت هستند. فرض اخیر که F و B به طور موضعی هموار است به این معنی نیست که تصویر ورودی I به طور موضعی هموار است زیرا ناپیوستگی در α می تواند ناپیوستگی در I را نتیجه دهد.

فرض این که F و B هر دو روی پنجره کوچکی اطراف هر پیکسل تقریباً ثابت هستند اجازه می دهد تا رابطه ۱-۱ را به صورت رابطه ۲-۱ بازنویسی کرده و α به عنوان یک تابع خطی از تصویر I به صورت زیر در نظر گرفته

شود:

$$\alpha_i = aI_i + b \quad \forall i \in w \quad (2-1)$$

بنا به رابطه 1-1 برای هر پیکسل i در پنجره w رابطه زیر برقرار است:

$$I_i = \alpha_i \mathcal{F}_i + (1 - \alpha_i) \mathcal{B}_i$$

$$\alpha_i (\mathcal{F}_i - \mathcal{B}_i) = I_i - \mathcal{B}_i$$

$$\alpha_i = \frac{1}{\mathcal{F}_i - \mathcal{B}_i} I_i + \frac{-\mathcal{B}_i}{\mathcal{F}_i - \mathcal{B}_i}$$

در نتیجه a و b برابر:

$$b = \frac{-\mathcal{B}_i}{\mathcal{F}_i - \mathcal{B}_i}$$

$$a = \frac{1}{\mathcal{F}_i - \mathcal{B}_i}$$

هدف پیدا کردن a, b و α و مینیمم سازی تابع هزینه زیر است:

$$J(\alpha, a, b) = \sum_{k \in I} \left(\sum_{i \in w_k} (\alpha_i - a_k I_i - b_k)^2 + \epsilon a_k^2 \right) \quad (3-1)$$

که w_k پنجره کوچکی اطراف پیکسل k می باشد. تابع هزینه بالا شامل یک عبارت منظم سازی بر حسب a میباشد. دلیل اضافه کردن این عبارت، پایداری عددی¹ است. برای مثال اگر تصویر، در k آمین پنجره ثابت باشد، a_k و b_k نمی توانند بدون داشتن یک دانش پیشین به صورت منحصر به فرد مشخص شوند. همچنین مینیمم کردن نرم a باعث پیدا شدن مقادیر آلفای هموارتر خواهد شد زیرا $a_k = 0$ به این معنا است که α در سراسر پنجره k ام ثابت است.

در عمل به طور نمونه از پنجره های 3×3 پیکسل استفاده می شود. چون اطراف هر پیکسل یک پنجره قرار داده می شود؛ پنجره های w_k در رابطه 1-3 هم پوشانی دارند. این ویژگی است که انتقال و توسعه اطلاعات بین دو پیکسل مجاور را ممکن می سازد.

تابع هزینه بالا نسبت به a, b و α درجه دوم است. برای تصویری با N پیکسل $3N$ مجهول دارد. خوشبختانه همانگونه که در قضیه بعد نشان داده می شود می توان a و b در رابطه 1-3 نادیده گرفته و تابعی با فقط N مجهول باقی بماند.

¹Numerical stability

قضیه ۱. فرض کنید $J(\alpha)$ به صورت زیر تعریف شود:

$$J(\alpha) = \min_{a,b} J(\alpha, a, b)$$

آنگاه $J(\alpha)$ را می توان به صورت $\alpha^T L \alpha$ (بدون وابستگی به a و b) نوشت که L یک ماتریس $N \times N$ است که (i, j) اُمین درایه آن به صورت زیر است:

$$\sum_{k|(i,j) \in w_k} \left(\delta_{ij} - \frac{1}{|w_k|} \left(1 + \frac{1}{\frac{\epsilon}{|w_k|} + \sigma_k^2} (I_i - \mu_k)(I_j - \mu_k) \right) \right) \quad (4-1)$$

در رابطه فوق δ_{ij} ، دلتای کرونیکر^۱ و μ_k ، میانگین و σ_k^2 ، واریانس شدت روشنایی پیکسلهای پنجره w_k حول پیکسل k و $|w_k|$ تعداد پیکسلهای این پنجره است. μ_k و σ_k^2 در روابط ۵-۱ و ۶-۱ تعریف می شوند:

$$\mu_k = \frac{\sum_{i \in w_k} I_i}{|w_k|} \quad (5-1)$$

$$\sigma_k^2 = \frac{\sum_{i \in w_k} I_i^2}{|w_k|} - \left(\frac{\sum_{i \in w_k} I_i}{|w_k|} \right)^2 \quad (6-1)$$

□

برهان. مرجع [۱۰]

۲-۴-۱ تصاویر رنگی

یک روش ساده برای به کارگیری تابع هزینه برای تصاویر رنگی، به کارگیری تابع هزینه سطح خاکستری^۲ برای هر کانال به صورت جداگانه^۳ می باشد.

مجدداً می توان مدل خطی ۲-۱ را با مدل خطی زیر جایگزین کرد:

$$\alpha_i \approx \sum_c a^c I_i^c + b \quad \forall i \in w \quad (7-1)$$

که c نمایانگر کانال های رنگی می باشد. مزیت این مدل خطی ترکیب شده این است که فرض قبلی را که \mathcal{F} و \mathcal{B} روی پنجره ثابت باشند را تخفیف می دهد. در عوض همانطور که در ذیل نشان داده می شود کافی است فرض

^۱Kronecker delta ^۲Gray level ^۳Channel separatly

شود که در هر پنجره کوچک هر کدام از \mathcal{F} و \mathcal{B} ترکیب خطی از دورنگ می باشد. به عبارت دیگر مقدار \mathcal{F}_i در هر پنجره کوچک روی یک خط واحد در فضای رنگ RGB قرار می گیرد:

$$\mathcal{F}_i = \beta_i \mathcal{F}_1 + (1 - \beta_i) \mathcal{F}_2 \quad (8-1)$$

به طور مشابه این فرض برای مقدار پس زمینه \mathcal{B}_i نیز صادق است.

قضیه ۲. اگر رنگ پیش زمینه و پس زمینه در هر پنجره در مدل خطی صدق کند می توان نشان داد:

$$\alpha_i = \sum_c a^c I_i^c + b \quad \forall i \in w$$

که c نشان دهنده کانال رنگ است.

□

برهان. مرجع [۱۰]

با استفاده از مدل خطی معادله ۷-۱ تابع هزینه ۹-۱ برای درهم تنیدگی تصاویر رنگی تعریف می شود:

$$J(\alpha, a, b) = \sum_{k \in I} \left(\sum_{i \in w_k} (\alpha_i - \sum_c a_k^c I_i^c - b_k)^2 + \epsilon \sum_c a_k^c \right) \quad (9-1)$$

مشابه نمونه خاکستری a^c و b می توانند از تابع هزینه حذف شوند. در این صورت تابع هزینه درجه دوم فقط نسبت به α مجهول بوده و به صورت زیر حاصل می شود:

$$J(\alpha) = \alpha^T L \alpha \quad (10-1)$$

که L یک ماتریس $N \times N$ است که درایه (i, j) آن به صورت زیر می باشد:

$$\sum_{k|(i,j) \in w_k} \left(\delta_{ij} - \frac{1}{|w_k|} \left(1 + I_i - \mu_k \right) \left(\Sigma_k + \frac{\epsilon}{|w_k|} I_3 \right)^{-1} (I_j - \mu_k) \right) \quad (11-1)$$

در رابطه فوق Σ ماتریس کوواریانس 3×3 و μ_k برداری 3×1 نشان دهنده میانگین رنگ پیکسل های پنجره w_k است و I_3 ماتریس همانی 3×3 است. ماتریس L در معادله ۱۰-۱ با عنوان ماتریس لاپلاسیان^۲ معرفی میشود. ماتریس لاپلاسیان درهم تنیدگی تصویر (L) را نیز می توان به صورت زیر بازنویسی کرد:

$$L = D - W \quad (12-1)$$

^۱Covariance matrix

^۲Laplacian matrix

در رابطه فوق D ماتریسی قطری به صورت $D(i, i) = \sum_j W(i, j)$ است و W ماتریسی متقارن است که درایه های غیر قطری آن با استفاده از رابطه ۱-۱۱ حاصل می شود [۷]. در واقع ماتریس (L) ، ماتریس لاپلاسیان گراف به کار رفته در روش قطعه بندی طیفی است که بر اساس تابع معرفی شده در معادله ۱-۱۱ محاسبه می شود. با ترکیب تابع هدف با محدودیتهای مشخص شده توسط کاربر در نقشه سه گانه تابع هزینه ای که باید مینیمم شود به صورت زیر حاصل می شود:

$$\operatorname{argmin} F(\alpha) = \alpha^T L \alpha + \lambda \|\alpha_{\Omega_l} - \alpha_{\Omega_l}^*\|_2^2 \quad (13-1)$$

که در آن $\alpha_{\Omega_l}^*$ برچسب از قبل مشخص شده برای داده های آموزشی است (نواحی مشخص در نقشه سه گانه) و α_{Ω_l} تخمین همین نقاط در آلفای کمینه است. D_{Ω_l} را ماتریس قطری $n \times n$ در نظر گرفته که j امین عنصر قطری مقدار ثابت ۱ را اگر $j \in \Omega_l$ باشد را اختیار و بقیه عناصر قطری صفر می باشند. بردار α^* به طول n در نظر گرفته می شود که j امین عنصر آن مقدار معلوم Ω_{lj} را اختیار می کند اگر $j \in \Omega_l$ باشد. بنابراین معادله ۱-۱۳ را می توان به صورت رابطه ۱-۱۴ فرمول بندی کرد:

$$\operatorname{argmin} F(\alpha) = \alpha^T L \alpha + \lambda (\alpha - \alpha^*)^T D_{\Omega_l} (\alpha - \alpha^*) \quad (14-1)$$

λ در رابطه فوق نشان دهنده میزان اهمیت و تأثیر اطلاعات تعیین شده توسط کاربر است. معادله ۱-۱۴ یک مسأله بهینه سازی نامقید است که به آسانی قابل حل است و می توان با انتخاب مقادیر بزرگ تر منظم ساز λ بر اطلاعات تعیین شده توسط کاربر تأکید کرد.

برای کمینه سازی تابع $F(\alpha)$ از آن نسبت به مجهول α مشتق گرفته و برابر صفر قرار داده می شود.

$$\frac{\partial F(\alpha)}{\partial \alpha} = L \alpha + L^T \alpha + \lambda D_{\Omega_l} (\alpha - \alpha^*) + D_{\Omega_l}^T \lambda (\alpha - \alpha^*) = 0 \quad (15-1)$$

با توجه به اینکه ماتریس لاپلاسیان L یک ماتریس قطری است لذا $L^T = L$ می باشد، همچنین ماتریس D_{Ω_l} نیز قطری است بنابراین $D_{\Omega_l}^T = D_{\Omega_l}$ است. حال می توان معادله ۱-۱۵ را به صورت رابطه ۱-۱۶ نوشت:

$$2L \alpha + 2\lambda D_{\Omega_l} (\alpha - \alpha^*) = 0$$

$$2(L + \lambda D_{\Omega_l}) \alpha = 2\lambda D_{\Omega_l} \alpha^* \quad (16-1)$$

در نتیجه رابطه زیر حاصل می گردد:

$$\alpha = \frac{\lambda D_{\Omega_i} \alpha^*}{L + \lambda D_{\Omega_i}} \quad (17-1)$$

معادله فوق مقدار آلفای به دست آمده بر اساس روش فرم بسته را ارایه می دهد.

۵-۱ چندین نزدیکترین همسایه برای درهم تنیدگی تصویر

همانطور که در معادله ۱-۱ گفته شد، درهم تنیدگی، مسأله ای که ترکیبی از I ، B ، F و α است را به دو لایه پیش زمینه و پس زمینه تجزیه می کند. این معادله ترکیبی وقتی که بیش از دو لایه $n \geq 2$ موجود باشد به صورت کلی به شکل زیر نوشته می شود:

$$I = \sum_{i=1}^n \alpha_i F_i, \quad \sum_{i=1}^n \alpha_i = 1$$

در واقع روش نزدیکترین همسایگی با استفاده از نشانه گذاری های کاربر (نقشه سه گانه)، مسأله استخراج چند لایه را به طور همزمان حل می کند. این روش از جستجوی K نزدیکترین همسایه در فضای ویژگی و فاصله مکانی برای تطابق نتایج خوب استفاده می کند. روش فوق الذکر عموماً در هر فضای ویژگی (مانند HSV) و در هر ابعادی موفق است [۸].

نزدیکترین همسایگی، شامل یافتن نزدیکترین داده های آزمایشی نسبت به داده های آموزشی x در قالب \hat{Y} است. به عبارت دیگر، k نزدیکترین همسایگی به صورت رابطه ۱۸-۱ تعریف می شود:

$$\hat{Y}(x) = \frac{1}{k} \sum_{x_i \in \mathcal{N}_k(x)} y_i \quad (18-1)$$

که $\mathcal{N}_k(x)$ همان همسایه های x است که با k نزدیکترین نقاط x_i در نمونه های آموزشی تعریف شده است. برای یافتن نزدیکترین همسایگی، k نزدیکترین مشاهدات x_i را نسبت به ورودی x یافته و سپس به طور متوسط پاسخ آنها سنجیده می شود.

۱-۵-۱ درهم تنیدگی مبتنی بر نزدیکترین همسایگی

در ادامه روش درهم تنیدگی نزدیکترین همسایگی تحلیل و توصیف شده است که پیاده سازی ساده ای دارد؛ از سوی دیگر تجزیه و تحلیل و نتایج تجربی نشان می دهد که این رویکرد یک راه حل ساده، سریع و بهتر را از روش

درهم تنیدگی غیر محلی [۱۲] فراهم می‌کند.

۱-۱-۵-۱ محاسبه ماتریس شباهت A با استفاده از نزدیکترین همسایگی

محاسبه ماتریس A در درهم تنیدگی نزدیکترین همسایگی شامل جمع آوری همسایگان غیر محلی z از پیکسل i است. به جای استفاده از یک هسته بزرگ در عملیات درهم تنیدگی سریع^۱ و درهم تنیدگی غیر محلی^۲ در حوزه مکانی تصویر، با توجه به پیکسل داده شده i و با استفاده از محاسبه k نزدیکترین همسایگی در فضای ویژگی اصل غیر محلی پیاده سازی می‌شود.

مقادیر معمول برای K بین ۳ تا ۱۵ است که ۳ برای تصاویر اشیاء و ۱۵ برای تصاویر طبیعت است. این پارامترها قطعی نیستند ولی در آزمایشات ثابت نگه داشته می‌شوند. شکل ۱-۴ محاسبه A را در درهم تنیدگی نزدیکترین همسایگی به تصویر در آورده است. ماتریس شباهت A در درهم تنیدگی KNN (تصویر ۱-۴(آ)) مانند روش درهم تنیدگی غیر محلی [۱۲] با شعاع ۳ (تصویر ۱-۴(ب)) قطری نیست.

۲-۱-۵-۱ بردار ویژگی X با مختصات مکانی

درهم تنیدگی نزدیکترین همسایگی می‌تواند به راحتی برای دستیابی به فضای رنگ شش بعدی (SVBRDF) یا داده‌های با ابعاد بالا با فضای رنگی غیر RGB گسترش داده شود. برای درهم تنیدگی، بردار ویژگی $X(\cdot)$ در پیکسل معین i که شامل مختصات مکانی برای تقویت انسجام مکانی است به صورت زیر می‌تواند تعریف شود:

$$X(i) = (\cos(h), \sin(h), s, v, x, y)_i \quad (19-1)$$

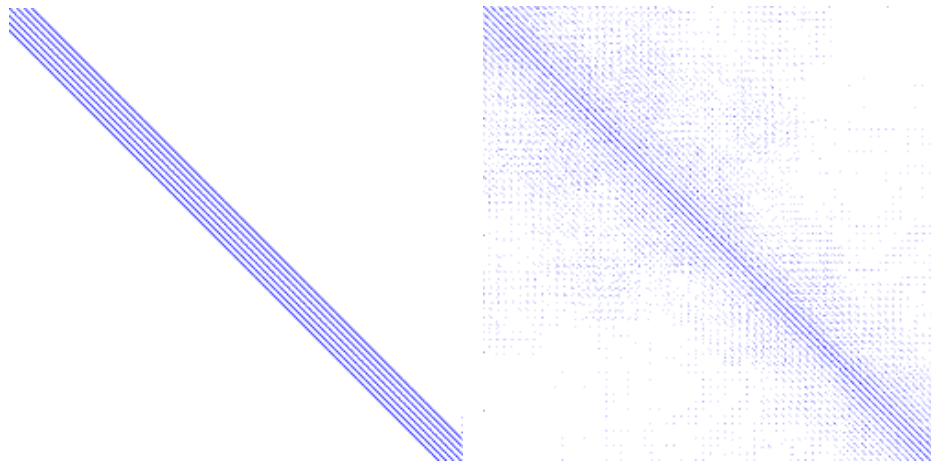
که h, s, v مختصات HSV هستند و (x, y) مختصات مکانی پیکسل i هستند. چند روش درهم تنیدگی قبلی از فضای رنگ HSV استفاده کرده‌اند. بردار ویژگی می‌تواند به طور مشابه برای درهم تنیدگی‌هایی که یک پیکسل دارای بیش از یک مشاهده است تعریف شود، مثلاً در اشیاء بدون انسجام مکانی (به عنوان مثال رنگ اسپری) میتوان مختصات مکانی را در نظر نگرفت. در درهم تنیدگی نزدیکترین همسایگی مختصات های مکانی به جای اینکه جداگانه در نظر گرفته شوند، به عنوان بخشی از بردار ویژگی مطرح شده‌اند.

۳-۱-۵-۱ حل فرم بسته با پیاده سازی سریع

از آنجاییکه خوشه بندی لاپلاسی $L = D - A$ برای خوشه های گراف مساعد است لاپلاسی $L_c = (D - A)^T(D - A)$ از این بخش

^۱Fast matting

^۲Nonlocal matting



شکل ۱-۴: مقایسه ماتریس شباهت غیر محلی A در درهم تیدگی KNN و درهم تیدگی غیر محلی [۱۲] با شعاع $K = 10$ (ب) درهم تیدگی غیر محلی [۱۲] با شعاع $K = 3$ (ا)

شکل ۱-۴: مقایسه ماتریس شباهت غیر محلی A در درهم تیدگی KNN و درهم تیدگی غیر محلی استفاده کرد (حدود ۱۰۰ بار سریعتر از L_c). هنگامی که ورودی های کاربر در قالب نقشه سه گانه هستند، می توان راه حل فرم بسته را برای استخراج بیشتر از دو لایه نشان داد که:

$$(L + \lambda D) \sum_i^n \alpha_i = \lambda m \quad (20-1)$$

در رابطه ۲۰-۱ $D = \text{diag}(m)$ ماتریسی قطری است، m یک بردار شاخص دودویی از همه پیکسل های مشخص شده است و λ یک ثابت کنترل اعتماد کاربر روی نشانه ها است. تابع بهینه $g(x)$ یک حل فرم بسته به صورت رابطه ۲۱-۱ دارد:

$$g(x) = x^T Lx + \lambda \sum_{i \in m-v} x_i^2 + \lambda \sum_{i \in v} (x_i - 1)^2 \quad (21-1)$$

در رابطه ۲۱-۱ v یک بردار دودویی از شاخص های مربوط به پیکسل نشانه گذاری شده توسط کاربر برای یک لایه معین است. در نهایت:

$$\begin{aligned} g(x) &= x^T Lx + \lambda \sum_{i \in m-v} x_i^2 + \lambda \sum_{i \in v} x_i^2 - 2\lambda v^T x + \lambda |v| \\ &= x^T Lx + \lambda \sum_{i \in m} x_i^2 - 2\lambda v^T x + \lambda |v| \\ &= \frac{1}{2} x^T 2(L + \lambda D)x - 2\lambda v^T x + \lambda |v| \\ &= \frac{1}{2} x^T Hx - c^T x + \lambda |v| \end{aligned}$$

که $|v| \lambda$ یک ثابت است. $H = \gamma(L + \lambda D)$ یک ماتریس نیمه معین مثبت است زیرا L نیمه معین مثبت است و D یک ماتریس قطری تولید شده توسط بردار دودویی m است. مشتق $g(x)$ نسبت به x گرفته و برابر صفر قرار داده می شود:

$$\frac{\partial g}{\partial x} = Hx - c = 0 \quad (22-1)$$

بنابراین حل بهینه آن برابر است با:

$$H^{-1}c = (L + \lambda D)^{-1}(\lambda v) \quad (23-1)$$

به جای استفاده از تکنیک coarse-to-fine در حل کننده $[V]$ ، چون H ماتریسی بزرگ، خلوت، متقارن و نیمه معین مثبت است از روش PCG^۱ [۱۴] که حدود ۵ ثانیه سریعتر از روش مرسوم مزدوج است برای حل تصاویر ورودی موجود در سایت ارزیابی Alpha Matting استفاده می شود.

۴-۱-۵-۱ جمع آوری ویژگی

درهم تنیدگی نزدیکترین همسایگی در شکل کلی برای استخراج بیشتر از دو لایه ($n \geq 2$) در مجموع ویژگی رضایت بخش است. به عبارتی، مجموع آلفاهای تخمین زده شده در هر پیکسل برابر ۱ می شود. با توجه به معادله ۲۲-۱:

$$(L + \lambda D)\alpha_1 = \lambda v_1$$

⋮

$$(L + \lambda D)\alpha_n = \lambda v_n$$

رابطه زیر حاصل می شود:

$$(L + \lambda D) \sum_{i=1}^n \alpha_i = \lambda \sum_{i=1}^n v_i = \lambda m \quad (24-1)$$

زیرا:

$$(L + \lambda D)\mathbf{1} = \lambda D\mathbf{1} = \lambda m \quad (25-1)$$

^۱Preconditioned Conjugate Gradients

به عنوان فضای خالی از لاپلاسین، L یک بردار ثابت تمام ۱ است زیرا $L + \lambda D$ معکوس پذیر است و $\sum_{i=1}^n \alpha_i = 1$ است.

در عمل درهم تنیدگی نزدیکترین همسایگی، آلفاهای خروجی تقریباً بین $[0, 1]$ هستند. برای بدست آوردن آلفای یک لایه، این لایه به عنوان پیش زمینه و مابقی به عنوان پس زمینه در نظر گرفته می شود. سه لایه $L_1 = (1, 0, 0)$ ، $L_2 = (0, 1, 0)$ ، $L_3 = (0, 0, 1)$ و پیکسل $I = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ را در نظر بگیرید. برای بدست آوردن آلفا از L_1 آن را به عنوان پیش زمینه \mathcal{F} و اجتماع L_2 و L_3 به عنوان پس زمینه \mathcal{B} در نظر گرفته می شود. طبق معادله ۲ در [۱۷]، $\alpha = \frac{(I-\mathcal{B})(\mathcal{F}-\mathcal{B})}{\|\mathcal{F}-\mathcal{B}\|^2}$ ، و مقدار آلفا برای L_1 برابر با $0/5$ است. به همین صورت مقدار آلفا برای L_2 یا L_3 برابر با $0/5$ است. در نتیجه مجموع آنها برابر با $1/5$ می باشد.

فصل ۲

درهم تنیدگی تصویر با استفاده از مجموعه نمونه های وسیع

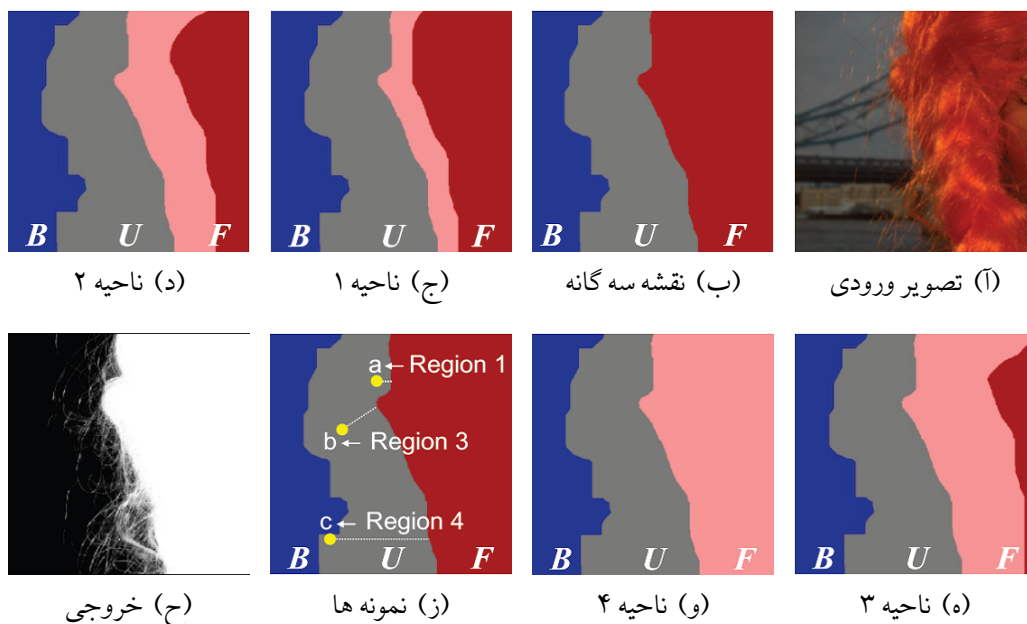
روشی که در این فصل معرفی می شود جزو روشهای بر پایه نمونه برداری است و بر اساس بدست آوردن نمونه های پیش زمینه و پس زمینه با استفاده از توزیع رنگها و گسترش مرزهای نواحی مشخص عمل می کند [۲]. روشهای بر پایه نمونه برداری مجموعه ای از نمونه های پیش زمینه و پس زمینه را برای تخمین مقادیر آلفای پیکسل های نامشخص جمع آوری می کنند. در این روشها بعد از انتخاب بهترین نمونه پیش زمینه و پس زمینه برای پیکسل z مقدار آلفای آن پیکسل به صورت زیر بدست می آید:

$$\alpha_z = \frac{(I_z - B) \cdot (F - B)}{\|F - B\|^2} \quad (1-2)$$

در رابطه فوق I_z : رنگ پیکسل z ، F و B : نمونه های پیش زمینه و پس زمینه ای هستند که در بخش ۲-۲ بدست می آیند.

روش مورد بحث هر یک از نواحی مشخص پیش زمینه و پس زمینه را که کاربر در نقشه سه گانه تعریف میکند به چندین ناحیه مجزا تقسیم کرده و با به دست آوردن نمونه های کاندید به تخمین مقدار آلفای هر پیکسل در ناحیه نامشخص می پردازد. برای هر پیکسل در ناحیه مشخص تعدادی نمونه از پیش زمینه و تعدادی نمونه از پس زمینه تعیین می گردد، با توجه به این نمونه ها می توان جفت نمونه های کاندید را که ترکیبی از پیش زمینه و پس زمینه به صورت (F, B) هستند را، مشخص کرد.

با به دست آمدن جفت نمونه های کاندید برای هر پیکسل، آلفاهایی با توجه به رابطه ۲-۱ به دست می آید که ممکن است مقادیر آلفاهای بدست آمده یا در محدوده نباشند یا نمونه ها با هم همپوشانی داشته باشند که در بخش ۲-۳ توضیح داده می شود. بنابراین نیاز به انتخاب بهترین جفت نمونه کاندید است و در بخش ۲-۴ نحوه



شکل ۱-۲: تقسیم بندی ناحیه پیش زمینه در روش درهم تیدگی مجموعه نمونه های وسیع

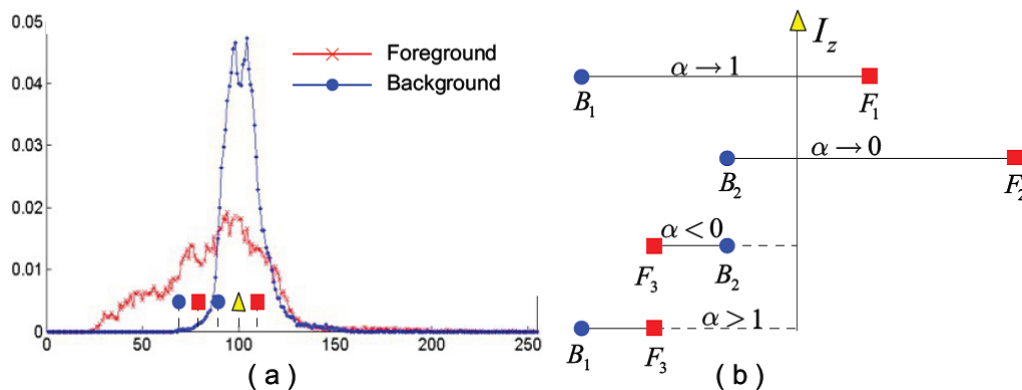
بدست آمدن آن شرح داده می شود. از این روش می توان برای بدست آوردن نقشه سه گانه ای دقیق تر با اعمال پیش پردازشی در روش های مختلف درهم تیدگی استفاده کرد (بخش ۲-۵).

۱-۲ بدست آوردن مجموعه نمونه های وسیع

هدف از این بخش پیدا کردن نمونه های بیشتری از تصویر ورودی متناظر با نقشه سه گانه است که بر اساس آن تخمینی از پیکسل های پیش زمینه و پس زمینه وجود داشته باشد. به این مجموعه جدید، مجموعه نمونه های وسیع گفته می شود. برای بدست آوردن این مجموعه نمونه ها، ابتدا نمونه های پیش زمینه بدست آورده می شوند، سپس اعمالی که برای بدست آوردن نمونه های پیش زمینه انجام شده دقیقاً برای پس زمینه تکرار و نمونه های پس زمینه نیز به دست می آیند. در الگوریتم به دست آوردن نمونه های پیش زمینه کل این ناحیه که در تصویر ۱-۲ (ب) با رنگ قرمز نشان داده شده است، به چندین قسمت تقسیم می شود. سوالی که پیش می آید این است که این نواحی چه تعداد باید باشند و میزان پهنای هر یک از این نواحی چه اندازه باید باشد؟

اندازه پهنای هر ناحیه برابر با نازکترین قسمت مرزی است که از نزدیکترین منطقه به مرز شروع شده و هر بار ناحیه جدید از اضافه شدن نازکترین قسمت مرزی به ناحیه قبلی ساخته می شود تا کل پیش زمینه پوشش داده شود و به این ترتیب تعداد نواحی نیز مشخص می شود.

در تصویر ۱-۲ (ج) قسمت صورتی رنگ بخشی از پیش زمینه است که به عنوان ناحیه ۱ مشخص گردیده است. در تصویر ۱-۲ (د) ناحیه صورتی رنگ که باید نمونه ها از آن بدست آیند نسبت به تصویر قبلی بیشتر شده



(1)

شکل ۲-۲: مشکل انتخاب نمونه وقتی که پیش زمینه و پس زمینه رنگ یکسان دارند. تصویر (a) توزیع رنگ نواحی پیش زمینه و پس زمینه و تصویر (b) تأثیر توزیع هم پوشانی بر روی آلفا را نشان می دهد.

است. در حقیقت به اندازه پهنای مشخص شده هر دفعه ناحیه صورتی رنگ بزرگتر می شود. در نتیجه نمونه های بیشتری بدست می آیند. در تصویر ۲-۱ (ه) نیز این ناحیه گسترش یافته و در نهایت در تصویر ۲-۱ (و) کل پیش زمینه به عنوان ناحیه صورتی رنگ تعیین شده است. تصویر ۲-۱ (ز) تعدادی پیکسل در ناحیه نامشخص را نشان می دهد که برای آنها باید نمونه هایی از پیش زمینه و پس زمینه مشخص شوند. تصویر ۲-۱ (ح) نتیجه عمل درهم تنیدگی تصویر را توسط روش مورد بحث نشان می دهد.

برای هر ناحیه یک خوشه بندی دو سطحی انجام می شود که در سطح اول پیکسل ها بر اساس رنگ به وسیله مدل ترکیب گوسی GMM^۱ خوشه بندی می شوند، تعداد اجزاء GMM برابر با ماکزیمم هیستوگرام رنگ پیکسل ها در آن ناحیه می باشد. در سطح دوم همان فرآیند خوشه بندی بر روی پیکسل های خوشه های بدست آمده در سطح اول با توجه به فاصله اندیسهای پیکسل ها انجام می شود. در نهایت میانگین مقدار رنگ هر خوشه که در سطح دوم به دست آمده، مجموعه نمونه های کاندید در آن ناحیه را تشکیل می دهند.

نمونه های کاندید هر پیکسل در ناحیه نامشخص \mathcal{U} در نقشه سه گانه از بین نزدیکترین ناحیه از بین نواحی تقسیم شده پیش زمینه به دست می آید. هر چه فاصله پیکسل نامشخص بیشتر باشد به ناحیه بزرگتر و نمونه های بیشتر نیاز است (برای به دست آوردن نمونه های کاندید از پس زمینه نیز دقیقاً اعمال فوق تکرار می گردد).

۲-۲ انتخاب نمونه های کاندید

با انجام خوشه بندی برای هر پیکسل نامشخص چند نمونه از پیش زمینه و چند نمونه از پس زمینه به دست می آید که تعداد کل جفت نمونه ها از ضرب آنها در یکدیگر به دست آمده و به صورت $(\mathcal{F}, \mathcal{B})$ نشان داده می شود. برای

^۱Gaussian Mixture Model

مثال اگر برای یک پیکسل نامشخص ۴ نمونه از پیش زمینه و ۳ نمونه از پس زمینه به دست آمده باشد، تعداد نمونه های کاندید ۱۲ جفت می باشد که به صورت زیر مشخص می شود:

$$\begin{aligned} & (\mathcal{F}_1, B_1), (\mathcal{F}_1, B_2), (\mathcal{F}_1, B_3) \\ & (\mathcal{F}_2, B_1), (\mathcal{F}_2, B_2), (\mathcal{F}_2, B_3) \\ & (\mathcal{F}_3, B_1), (\mathcal{F}_3, B_2), (\mathcal{F}_3, B_3) \\ & (\mathcal{F}_4, B_1), (\mathcal{F}_4, B_2), (\mathcal{F}_4, B_3) \end{aligned}$$

۳-۲ مشکل هم پوشانی توزیع های رنگ

بعد از انتخاب نمونه های کاندید، با تخمین آلفا در معادله ۲-۱ برای هر پیکسل چندین مقدار آلفا با توجه به تعداد نمونه های کاندید آن بدست می آید. برای پیکسلی در ناحیه نامشخص، ممکن است یک جفت نمونه، پیکسل را جزو پس زمینه و جفتی دیگر آنرا جزو پیش زمینه تشخیص دهند. (شکل ۲-۲(b)) این مسأله به عنوان مشکل هم پوشانی شناخته می شود. برای حل این مشکل نمونه ای که کمترین هم پوشانی را داشته باشد به عنوان بهترین نمونه برای تخمین بهترین آلفا تعیین می شود.

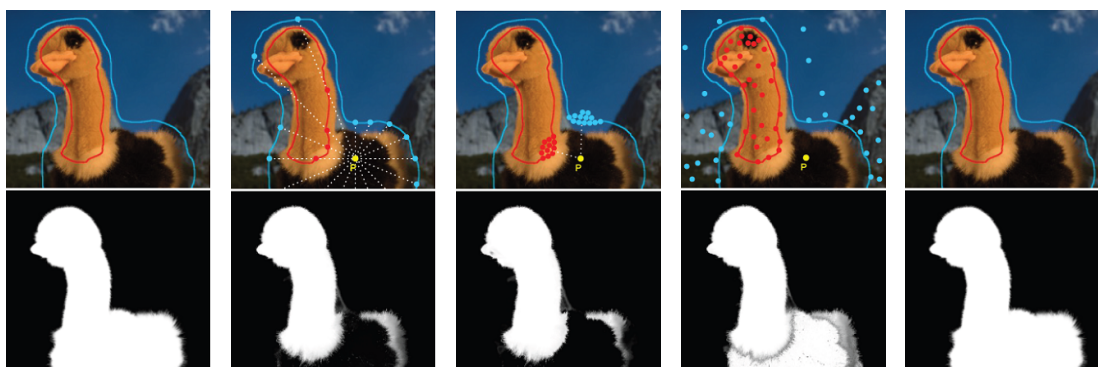
در شکل ۲-۲(b) مقدار آلفای تخمین زده شده نشان می دهد که پیکسل z بوسیله زوج (\mathcal{F}_1, B_1) پیش زمینه $(\alpha = 1)$ و بوسیله زوج (\mathcal{F}_2, B_2) پس زمینه $(\alpha = 0)$ در نظر گرفته می شوند. به علت اینکه پیکسل z در فضای رنگ نزدیک به \mathcal{F}_1 و B_2 است و دیگر مقادیر آلفای تخمین زده شده نیز خارج از محدوده $[0, 1]$ می باشند پیدا کردن بهترین جفت خیلی سخت می باشد. انتخاب بهترین زوج (\mathcal{F}, B) تولید شده ای که کمترین هم پوشانی را در توزیع رنگ نمونه های پیش زمینه و پس زمینه دارند در بخش بعد توضیح داده می شود.

۴-۲ انتخاب بهترین جفت نمونه کاندید (\mathcal{F}, B)

تابع هدفی که در زیر تشریح می شود هم پوشانی توزیع های رنگ را مشخص می کند. بهترین جفت نمونه کاندید بر اساس معادله زیر انتخاب می شود:

$$O_z(\mathcal{F}_i, B_i) = K_z(\mathcal{F}_i, B_i) \times S_z(\mathcal{F}_i, B_i) \times C_z(\mathcal{F}_i, B_i) \quad (2-2)$$

در رابطه بالا K تحریف رنگ را نشان داده و S و C آمارهایی از فاصله ها و رنگهای تصویر می باشند که مقادیر



شکل ۲-۳: تصویر اصلی (آ) تصویر اصلی (ب) روش مجموعه (ج) روش قوی [۲۷] (د) روش اشتراکی (ه) روش سراسری نمونه های وسیع [۱۷] [۲۴]

شکل ۲-۳: تصویر اصلی و استراتژی نمونه برداری روش مورد بحث در این فصل و روش های سراسری [۱۷]، اشتراکی [۲۴] و قوی [۲۷] (رنگ زرد پیکسل، رنگ قرمز نمونه های پیش زمینه و رنگ آبی نمونه های پس زمینه می باشند)

آنها به صورت زیر محاسبه می گردند:

$$K_z(\mathcal{F}_i, \mathcal{B}_i) = \exp\left(-\|I_z - (\alpha\mathcal{F}_i + (1-\alpha)\mathcal{B}_i)\|\right) \quad (3-2)$$

I_z مقدار مشاهده شده رنگ پیکسل z است.

$$S_z(\mathcal{F}_i, \mathcal{B}_i) = \exp\left(\frac{-\|z - \mathcal{F}_i^s\|}{\frac{1}{|S_z^{\mathcal{F}}|} \sum_{\mathcal{F}_k \in S_z^{\mathcal{F}}} \|z - \mathcal{F}_k^s\|}\right) \times \exp\left(\frac{-\|z - \mathcal{B}_i^s\|}{\frac{1}{|S_z^{\mathcal{B}}|} \sum_{\mathcal{B}_k \in S_z^{\mathcal{B}}} \|z - \mathcal{B}_k^s\|}\right) \quad (4-2)$$

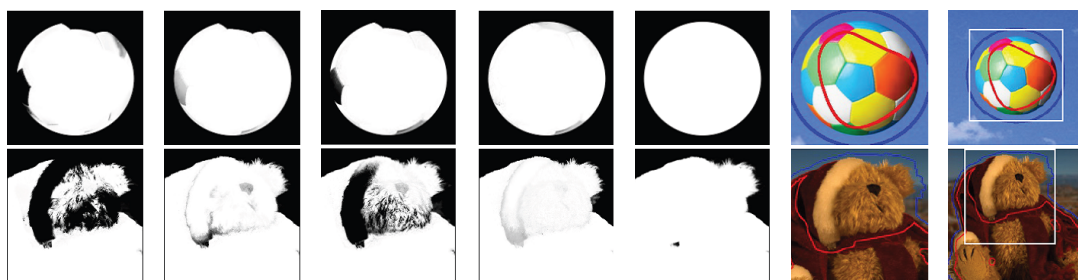
$S_z^{\mathcal{F}}$ مجموعه نمونه های پیش زمینه مشخص برای پیکسل z است. مختصات مکانی نمونه پیش زمینه \mathcal{F}_i به صورت \mathcal{F}_i^s نشان داده می شود. فاصله اقلیدوسی بین پیکسل z و نمونه پیش زمینه \mathcal{F}_i توسط $\|z - \mathcal{F}_i^s\|$ نشان داده شده است.

$$C_z(\mathcal{F}_i, \mathcal{B}_i) = \frac{d(\mathcal{F}_i, \mathcal{B}_i)}{M_z} \quad (5-2)$$

$d(\mathcal{F}_i, \mathcal{B}_i)$ ، مقدار کوهن^۱ توزیع رنگ است و M_z یک فاکتور مقیاس گذاری که بیشترین مقدار کوهن d برای مجموعه جفت نمونه های کاندید پیکسل z می باشد. مقدار کوهن d متناسب با معکوس هم پوشانی توزیعها است که مقداری بزرگ برای توزیع های به خوبی جدا شده است و توسط رابطه زیر محاسبه می شود:

$$d(\mathcal{F}_i, \mathcal{B}_i) = \frac{\mu_{\mathcal{F}_i} - \mu_{\mathcal{B}_i}}{\sqrt{\frac{(N_{\mathcal{B}_i}-1)\sigma_{\mathcal{B}_i}^2 + (N_{\mathcal{F}_i}-1)\sigma_{\mathcal{F}_i}^2}{N_{\mathcal{B}_i} + N_{\mathcal{F}_i} - 2}}} \quad (6-2)$$

^۱Cohen's d



(آ) تصویر اصلی (ب) تصویر بزرگ شده (ج) ناحیه واقعی (د) زمینه مجموعه های وسیع (ه) نمونه سراسری [۱۷] (و) روش اشتراکی [۲۴] (ز) روش قوی [۲۷]

شکل ۲-۴: مقایسه چند روش نمونه برداری و نمایش نمونه های از دست رفته

μ_{F_i} و $\sigma_{F_i}^2$ میانگین و واریانس و N_{F_i} اندازه جمعیت توزیع نمونه F_i تولید شده هستند. شکل ۲-۴ دو تصویر با جواب های مرجعشان را نشان می دهد. در تصویر بالایی در هر ۴ روش نشان داده شده نمونه هایی از دست رفته اند اما کیفیت درهم تنیدگی روش مجموعه نمونه های وسیع نسبت به روش های [۱۷، ۲۴، ۲۷] کاملاً مشهود است. در تصویر پایینی نیز برتری عملکرد روش مذکور دیده می شود. در شکل ۲-۵ نیز روش درهم تنیدگی مجموعه نمونه های وسیع با سه روش مشهور مقایسه گردیده است. در تصویر ۲-۵(د) دیده می شود که فاصله بین دو عروسک به اشتباه جزو پیش زمینه تشخیص داده شده است این مشکل در تصویر ۲-۵(ه) دیده نمی شود اما این تصویر نیز در نواحی دور با مشکل مواجه شده است. به نظر میرسد روش مورد بحث در این فصل از پس مشکلات فوق بر آمده است اما با مقایسه نمودن ناحیه مشخص شده آن با تصویر اصلی دیده میشود که برخی نواحی را جزو پس زمینه تشخیص داده است. تصویر آخر نتیجه روش درهم تنیدگی مبتنی بر جریان اطلاعات بین پیکسلی می باشد که نتیجه خوب و بهتری دارد.

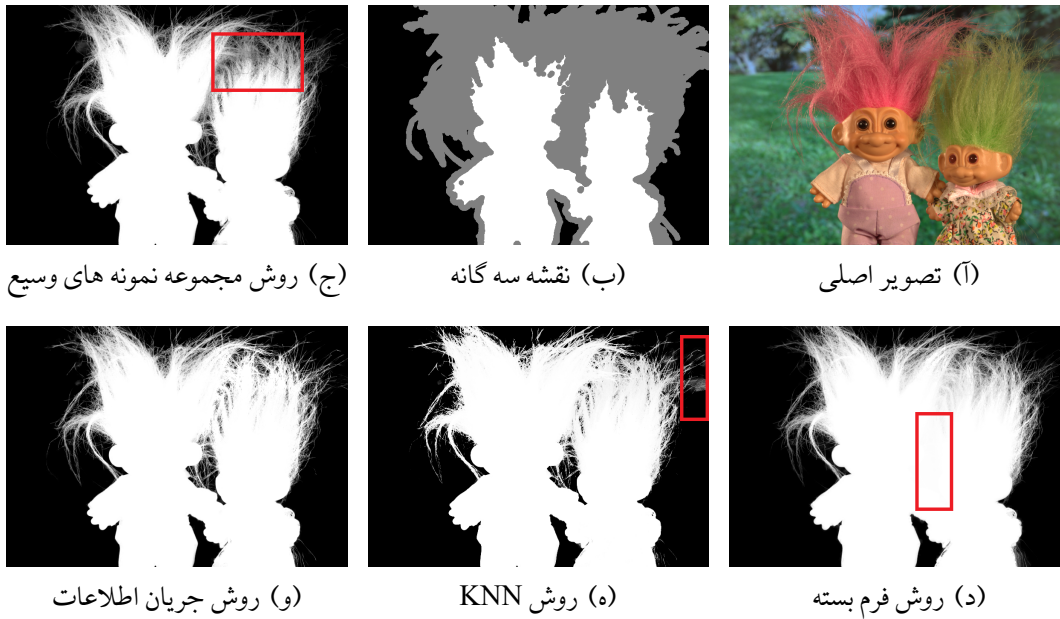
۵-۲ پیش پردازش و پس پردازش

روش بحث شده در این فصل می تواند به عنوان یک گام پیش پردازش برای گسترش نواحی مشخص به نواحی نامشخص مطابق با شرط زیر نیز استفاده می شود:

برای این منظور پیکسل نامشخص z جزو پیش زمینه علامت گذاری می شود اگر برای پیکسل $q \in F$ رابطه زیر برقرار باشد:

$$(D(z, q) < E_{thr}) \wedge (||I_z - I_q|| \leq (C_{thr} - D(z, q))) \quad (7-2)$$

$D(z, q)$ فاصله اقلیدسی بین پیکسل z و q است و E_{thr} و C_{thr} آستانه فضاها رنگ و فاصله می باشند که



شکل ۲-۵: مقایسه روش درهم تنیدگی مجموعه نمونه های وسیع با چند روش مشهور

به طور تجربی ۹ قرار داده شده است. یک معادله مشابه نیز می تواند برای مقایسه یک پیکسل با پس زمینه اعمال شود.

در نهایت مقدار آلفای بدست آمده برای هر پیکسل با استفاده از جفت $(\mathcal{F}, \mathcal{B})$ در معادله ۲-۱ برای رسیدن به مقدار آلفای هموار، با توجه به لحاظ کردن همبستگی بین پیکسلهای همسایه بیشتر پیرایش می شود. همچنین روش پس پردازش [۲۴] که تابع هزینه آن شامل یک اصطلاح $\hat{\alpha}$ و مقدار ضریب اعتماد f است به همراه یک اصطلاح همواری شامل ماتریس لاپلاسیان در [۷] که با توجه به α ، کمینه شده اتخاذ می شود.

$$\alpha = \operatorname{argmin} \quad \alpha^T L \alpha + \lambda (\alpha - \hat{\alpha})^T \Sigma (\alpha - \hat{\alpha}) + \gamma (\alpha - \hat{\alpha})^T \hat{\Gamma} (\alpha - \hat{\alpha}) \quad (۸-۲)$$

ضریب اعتماد، مقداری از تابع هدف معادله ۲-۲ برای هر جفت کاندید $(\mathcal{F}, \mathcal{B})$ انتخاب شده، می باشد. این تابع هزینه در [۲۴] داده شده است. λ پارامتر وزنی بزرگی در مقایسه با آلفای تخمین زده شده $\hat{\alpha}$ است و با ضریب اعتماد f وقتی که γ مقدار ثابت (10^{-1}) دارد مرتبط و اهمیت ارتباط با داده و اصطلاح همواری را نشان میدهد. Σ ماتریسی قطری با مقادیر ۱ برای پیکسلهای ناحیه مشخص و صفر برای ناحیه نامشخص است و ماتریس قطری $\hat{\Gamma}$ مقادیر صفر برای پیکسلهای نواحی مشخص و مقادیر f برای ناحیه نامشخص دارد.

فصل ۳

درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی

در این فصل روش اصلی مورد بحث در این پایان نامه تشریح می گردد [۱].

نقشه های سه گانه عموماً توسط کاربر تهیه می شوند و آنها شامل سه ناحیه کاملاً کدر (پیش زمینه)، کاملاً شفاف (پس زمینه) و نامشخص هستند. این نواحی به ترتیب F ، B و U نامیده شده و K اجتماع F و B را نشان می دهد. روشهای بر پایه شباهت به وسیله گسترش جریان مبهم از K به U با استفاده از انواع شباهتهایی که این روشها تعریف می کنند عمل می نمایند. این عمل **جریان اطلاعات** تعریف می شود و در چندین شکل انجام می گردد تا همه پیکسلهای داخل U اطلاعات مؤثری از نواحی مختلف در تصویر دریافت کنند:

۱- تغییرات کدری (مقادیر آلفا) در درهم تنیدگی به عنوان نتیجه ای از ترکیب رنگهای اصلی در تصویر داده شده با یکدیگر به واسطه شفافیت پیکسلها یا بخشهای پیچیده آنها رخ می دهند. از این حقیقت برای پیکسلهای رنگی مشابه و تعریف یک فرم از جریان اطلاعات که **جریان ترکیب رنگ** نامیده می شود، استفاده می گردد (بخش ۳-۱).

۲- همچنین ارتباطاتی از هر پیکسل در U به F و B برای حفظ جریان اطلاعات مستقیم از نواحی مشخص به حتی دورترین نواحی ای که مقدار کدری آنها تغییر می کند، در تصویر اضافه می شود. **جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص** (بخش ۳-۲).

۳- به منظور توزیع اطلاعات از ترکیب رنگ و جریانهای K -to- U ، یک جریان از **اطلاعات داخل U تعریف** شده که پیکسلهایی با رنگهای مشابه داخل U اطلاعاتشان را با یکدیگر به اشتراک می گذارند (بخش ۳-۴).

۴- سرانجام یک **جریان اطلاعات محلی** اضافه می شود، یک پیکسل بر کدری همسایه های مکانی پهلویی که نتایج پایانی را تضمین می کند مؤثر است (بخش ۳-۵).

۵- در انتها فرمهای منحصر به فرد جریان اطلاعات به عنوان توابع هدف فرموله شده و در معادله بهینه سازی سراسری جمع می شوند (فصل ۳-۶).

در این روش ابتدا وزنها مطابق با فرم جریان اطلاعات ترکیب رنگ به دست آورده می شود. در ادامه مجدداً مقادیری جدید از وزنها این بار توسط تعریف جریان اطلاعات محلی بدست آمده و به مقادیر قبلی اضافه می گردد. در مرحله سوم تخمین مقادیر وزنها با استفاده از فرم جریان اطلاعات داخل ناحیه نامشخص صورت گرفته و دوباره به مقادیر قبلی اضافه می شود. سپس پردازشی بر اساس هیستوگرام بر روی تصویر انجام می شود تا مشخص شود آیا تصویر شفاف است یا خیر (مطابق با توضیحات آخر فصل ۳-۶). اگر تصویر شفاف نباشد پیش پردازشی بر روی نقشه سه گانه برای پیرایش بهتر نقشه مطابق با توضیحات فصل ۳-۳ انجام می شود و مقادیر آلفا و میزان اعتماد به آنها از روی تصویر اصلی و نقشه سه گانه پیرایش شده توسط فرم جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص تخمین زده می شود. سپس پیش پردازشی بر روی نقشه سه گانه توسط روش تشریح شده در فصل ۲ انجام می گردد و تخمین مقادیر آلفا و میزان اعتماد به آنها توسط این نقشه سه گانه بهبود می یابند. در نهایت اگر تصویر شفاف باشد، با استفاده از مقادیر تخمین زده شده آلفا و میزان اعتماد به آنها و وزنه‌های بدست آمده از سه فرم دیگر جریان اطلاعات، مقادیر نهایی آلفا تخمین زده می شوند و اگر تصویر شفاف باشد تخمین مقادیر نهایی آلفا توسط وزنه‌های بدست آمده توسط فرم های عنوان شده جریان اطلاعات انجام می گردد.

۱-۳ جریان اطلاعات ترکیب رنگ

به علت شفافیت اشیاء و داشتن ساختارهای خوب و لبه های تیز یک شی که نمی تواند کاملاً در تصویر برداری ثبت شود و محدودیت رزولوشن سنسورهای تصویر برداری، بعضی پیکسلهای تصویر به ناچار شامل ترکیبی از رنگهای پیش زمینه و پس زمینه به هم پیچیده می شوند. با بررسی این ترکیبهای رنگ می توان نشان داد که چگونه مقادیر آلفا بین پیکسلها پخش شده اند. میزان رنگ اصلی پیش زمینه در یک ترکیب خاص تعیین کننده کدورت یک پیکسل است. مطابق با این حقیقت اگر رنگ یک پیکسل با ترکیب وزن دار رنگهای دیگر نشان داده شود، این وزنها باید ارتباط بین پیکسلها را نشان دهند.

به منظور استفاده از این رابطه برای هر پیکسل در ناحیه نامشخص $K_{CM} = ۲۰$ پیکسل شبیه در فضای ویژگی را بوسیله تقریب K نزدیکترین همسایه در کل تصویر پیدا می شود. یک بردار ویژگی برای این جستجو به صورت $[r, g, b, \tilde{x}, \tilde{y}]^T$ تعریف می شود که \hat{x} و \hat{y} مختصات نرمال شده به وسیله ارتفاع و پهنای تصویر است و بقیه، مقادیر RGB پیکسل می باشند. این مجموعه همسایه ها پیکسلهای رنگی مشابه به هم دارند که با \mathcal{N}_p^{CM} نشان داده میشود. سپس وزنه‌های ترکیبی $w_{p,q}^{CM}$ را پیدا کرده که مقدار جریان اطلاعات را بین $p, q \in \mathcal{N}_p^{CM}$ تعیین خواهد کرد. وزنها طوری تعیین شده اند که رنگهای پیکسلهای همسایه یک پیکسل، رنگ اصلی آن پیکسل را وقتی ترکیبی

از رابطه زیر باشند را می دهند:

$$\operatorname{argmin}_{w_{p,q}^{CM}} \|c_p - \sum_{q \in \mathcal{N}_p^{CM}} w_{p,q}^{CM} c_q\|^2 \quad (1-3)$$

در رابطه فوق برداری 1×3 از مقادیر RGB یا همان رنگ پیکسل p است و به همین ترتیب c_q نیز برداری 1×3 از مقادیر RGB و رنگ پیکسل q است که همسایه پیکسل p می باشد. معادله بالا با استفاده از روش های مختلف می تواند مینیمم شود که در [۱] از روش [۶] برای کمینه کردن آن استفاده می گردد.

برای بدست آوردن وزنها ، بعد از پیدا کردن نزدیکترین همسایه ها ، برای هر پیکسل یک ماتریس همسایه بدست می آید ، توسط ضرب ترانهاده هر ماتریس همسایه در خودش ماتریس همبستگی ساخته می شود. اگر این ماتریس A و بردار بازسازی را که حاصل ضرب بردار همسایه در پیکسل نامشخص مورد نظر است b نامیده شوند، در این صورت می توان وزنها را برای هر پیکسل به راحتی با حل دستگاه $Ax = b$ به دست آورد.

بدلیل اینکه فقط از مقادیر RGB استفاده می شود احتمال بالایی وجود دارد که ماتریس همبستگی همسایه طی کمینه کردن توسط روش [۶] محاسبه شده ، دو همسایه با رنگهای یکسان داشته باشد. بنابراین ماتریس همبستگی همسایه با اضافه کردن $10^{-3} \times I_{(K_{CM} \times K_{CM})}$ به آن قبل از معکوس کردن مقید می شود که I ماتریس همانی است. هنگامی که از روش [۶] برای کمینه کردن تابع هدف استفاده می شود، به طور کامل روش جاسازی محلی خطی (LLE) اتخاذ نمی گردد. روش LLE مجموعه ای از همسایه ها در یک فضای ویژگی را پیدا کرده و از همه متغیرها در فضای ویژگی برای محاسبه وزنها به ترتیب کاهش ابعاد داده ورودی استفاده می کند.

در مرجع [۱] فقط از ترکیب وزنها و نه مختصات مکانی آنها در گام کمینه سازی انرژی استفاده شده و از مختصات مکانی در مرحله کمینه سازی صرف نظر شده است. این ترکیب وزن های تخمین زده شده جریان های مختلف حتی در ساده ترین حالتها مانند شکل ۱-۲ می تواند مشاهده شود که نتایج روش [۹] تصویر ۱-۲(ز) و فقط جریان ترکیب رنگ (only-CMF) تصویر ۱-۲(ح) ، در گام ترکیب وزن تفاوت دارد.

جریان ترکیب رنگ توسط معادله زیر محاسبه می شود:

$$E_{CM} = \sum_{p \in \mathcal{U}} (\alpha_p - \sum_{q \in \mathcal{N}_p^{CM}} w_{p,q}^{CM} \alpha_q)^2 \quad (2-3)$$

در رابطه فوق مقدار آلفای پیکسل p و α_q مقدار آلفای پیکسل q (که همسایه پیکسل p است)، است.



شکل ۳-۱: جریان ترکیب رنگ به تنهایی

۲-۳ جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص

جریان ترکیب رنگ، اطلاعات مفیدی شامل اینکه چگونه ترکیب رنگ پیکسلها شکل گرفته اند را می دهد. با این حال تعدادی از پیکسل های ناحیه نامشخص به طور غیر مستقیم اطلاعات موجود در نقشه سه گانه را از میان همسایه هایشان دریافت می کنند، همه اینها احتمالاً می توانند در ناحیه نامشخص باشند. این جریان اطلاعات غیر مستقیم برای ناحیه هایی که از ناحیه مشخص دورتر هستند کافی نیست. به منظور حفظ جریان اطلاعات از نواحی پیش زمینه و پس زمینه به هر ناحیه ای در ناحیه نامشخص، ارتباط های هر پیکسل در ناحیه نامشخص با چندین پیکسل در ناحیه مشخص اضافه می شود.

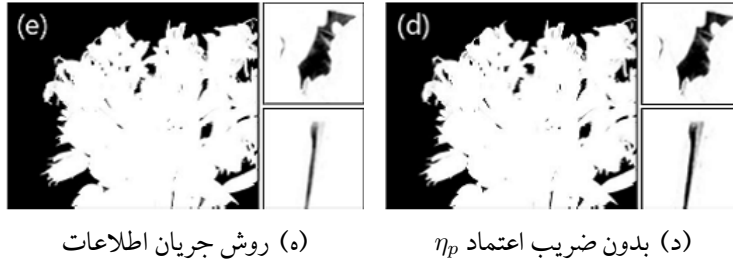
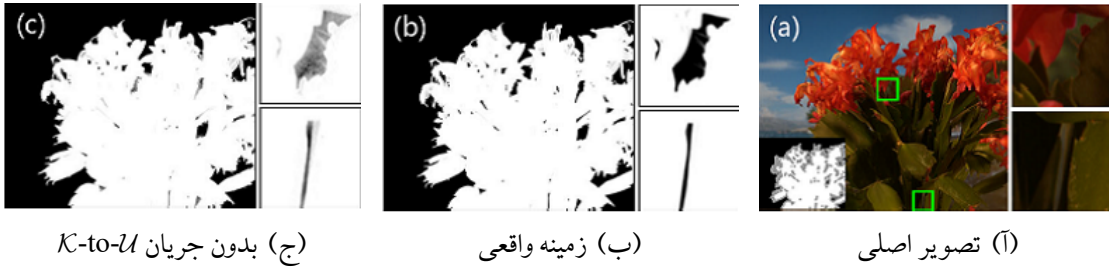
برای هر پیکسل در ناحیه نامشخص $K\mathcal{K}\mathcal{U} = \mathcal{V}$ پیکسل شبیه در نواحی پیش زمینه و پس زمینه جستجو میشود که به صورت جداگانه به فرم مجموعه ای از پیکسلهای \mathcal{N}_p^B و \mathcal{N}_p^F است و K نزدیکترین همسایه با استفاده از بردار ویژگی $[r, g, b, 10 \times \tilde{x}, 10 \times \tilde{y}]^T$ با توجه به پیکسل های نزدیک جستجو می گردد. از پیکسلهای \mathcal{N}_p^F و \mathcal{N}_p^B با هم برای نمایش رنگ c_p در معادله مینیمم شده ۳-۱ استفاده می شود. با استفاده از نتیجه وزنه های $w_{p,q}^B$ و $w_{p,q}^F$ تابع انرژی جریان \mathcal{K} -to- \mathcal{U} به صورت زیر تعریف می شود:

$$E_{\mathcal{K}\mathcal{U}} = \sum_{p \in \mathcal{U}} (\alpha_p - \sum_{q \in \mathcal{N}_p^F} w_{p,q}^F \alpha_q - \sum_{q \in \mathcal{N}_p^B} w_{p,q}^B \alpha_q)^2 \quad (3-3)$$

برای $q \in \mathcal{F}$ ، $\alpha_q = 1$ و برای $q \in \mathcal{B}$ ، $\alpha_q = 0$ است. این حقیقت اجازه می دهد که دو وزن ترکیبی یعنی اتصال یک پیکسل با پیش زمینه و دیگری اتصال آن پیکسل با پس زمینه به صورت زیر تعریف گردد:

$$w_p^F = \sum_{q \in \mathcal{N}_p^F} w_{p,q}^F \quad w_p^B = \sum_{q \in \mathcal{N}_p^B} w_{p,q}^B \quad (4-3)$$

است. $w_p^F + w_p^B = 1$



شکل ۲-۳: جریان اطلاعات \mathcal{K} -to- \mathcal{U} و ضریب اعتماد



شکل ۳-۳: جریان \mathcal{K} -to- \mathcal{U} به تنهایی

معادله ۳-۳ به صورت زیر بازنویسی می شود:

$$E_{\mathcal{K}\mathcal{U}} = \sum_{p \in \mathcal{U}} (\alpha_p - w_p^{\mathcal{F}})^2 \quad (5-3)$$

انرژی مینیمم شده در معادله ۱-۳ برای همه q ها وقتی که c_p ها شبیه به همدیگر هستند وزنه‌های مشابه می‌دهد. در نتیجه اگر $\mathcal{N}_p^{\mathcal{B}}$ و $\mathcal{N}_p^{\mathcal{F}}$ پیکسل‌های رنگی مشابه داشته باشند، وزنه‌های تخمین زده شده $w_p^{\mathcal{B}}$ و $w_p^{\mathcal{F}}$ نامعتبر و غیر قابل اعتماد می‌شوند. برای این منظور تابع انرژی ۳-۵ با مقادیر مورد اطمینان به صورت زیر تقویت و محاسبه می‌گردد:

میزان مشارکت رنگها در ترکیب تخمین زده شده معادله ۱-۳ را می‌توان با استفاده از وزنه‌های $w_{p,q}^{\mathcal{B}}$ و $w_{p,q}^{\mathcal{F}}$ تعیین

کرد:

$$c_p^{\mathcal{F}} = \frac{\sum_{q \in \mathcal{N}_p^{\mathcal{F}}} w_{p,q}^{\mathcal{F}} c_q}{w_p^{\mathcal{F}}} \quad c_p^{\mathcal{B}} = \frac{\sum_{q \in \mathcal{N}_p^{\mathcal{B}}} w_{p,q}^{\mathcal{B}} c_q}{w_p^{\mathcal{B}}} \quad (6-3)$$

همینطور می توان یک اندازه قابل اطمینان مطابق با اینکه چقدر رنگ پیش زمینه $c_p^{\mathcal{F}}$ و رنگ پس زمینه $c_p^{\mathcal{B}}$ تخمین زده شده شبیه هستند تعریف نمود:

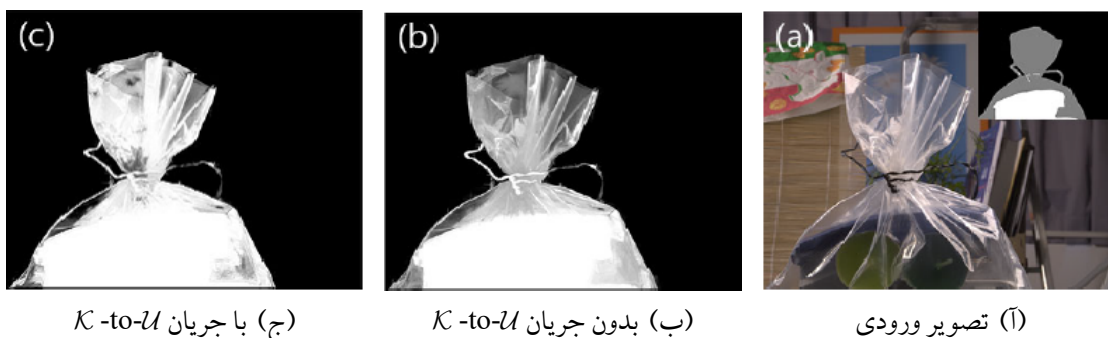
$$\eta_p = \|c_p^{\mathcal{F}} - c_p^{\mathcal{B}}\|^2 / 3 \quad (7-3)$$

در رابطه فوق تقسیم بر ۳ برای این است که مقادیر قابل اطمینان بین $[0, 1]$ بدست بیاید. انرژی جدید برای منعکس کردن مقادیر مورد اعتماد، معادله ۳-۵ را به صورت زیر بروز می کند:

$$\hat{E}_{\mathcal{K}\mathcal{U}} = \sum_{p \in \mathcal{U}} \eta_p (\alpha_p - w_p^{\mathcal{F}})^2 \quad (8-3)$$

این بروز کردن انرژی کیفیت درهم تیدگی را در ناحیه هایی با رنگهای پیش زمینه و پس زمینه مشابه، افزایش می دهد که در شکل ۳-۲ دیده می شود. باید توجه داشت که جریان اطلاعات از \mathcal{K} -to- \mathcal{U} وقتی شیء پیش زمینه شفافیت بالایی دارد (همانطور که در شکل ۳-۴ نشان شده است) قابل اطمینان نیست، که این اساساً به واسطه قدرت ضعیف نمایش $\mathcal{N}_p^{\mathcal{F}}$ و $\mathcal{N}_p^{\mathcal{B}}$ برای c_p های اطراف ناحیه های با شفافیت بالا است که به عنوان نزدیکترین همسایه های جستجو، پیکسلهایی با اندازه مناسب برای تخمین $w_{p,q}^{\mathcal{F}}$ نمی دهند. سیستم خطی نهایی مطابق با بحث فصل ۳-۶ بنا می گردد.

در شکل ۳-۲ (آ) یک تصویر و نقشه سه گانه آن و دو ناحیه بزرگ شده نشان داده شده است. تصویر ۳-۲ (ب) زمینه مرجع تصویر اصلی و قسمت های بزرگ شده را نشان می دهد. تصویر ۳-۲ (ه) روش جریان اطلاعات را هنگامی که از جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص و ضریب اعتماد مطرح شده در معادله ۳-۷ استفاده شده است را نشان می دهد. جریان اطلاعات از هم پیش زمینه و هم پس زمینه حتی در دورترین نواحی نامشخص، کارایی را در اطراف سوراخها افزایش می دهد (قسمت انتخاب شده بالایی). استفاده از ضریب اعتماد باعث افزایش کارایی بخصوص در اطراف ناحیه هایی از پیش زمینه و پس زمینه که رنگهای مشابه دارند می شود (قسمت انتخاب شده پایینی).



شکل ۳-۴: جریان اطلاعات K -to- U هنگامی که شیء پیش زمینه شفافیت بالایی دارد عملکرد خوبی ندارد.

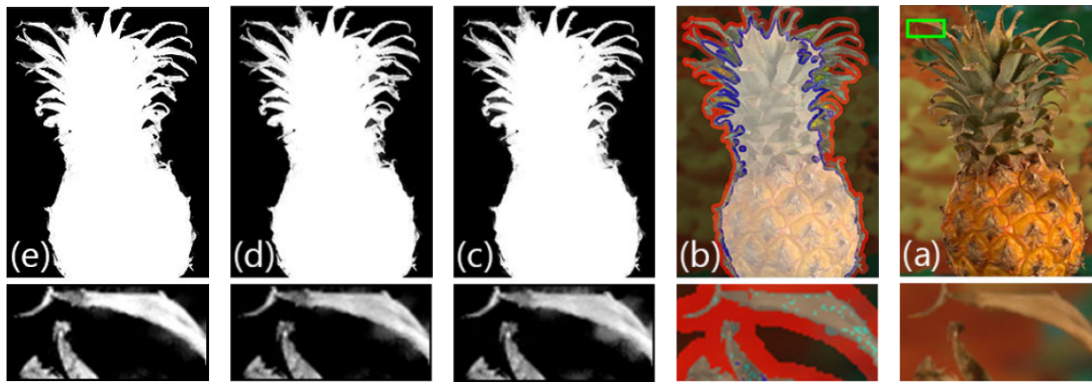
۳-۳ پیش پردازش نقشه سه گانه

در این بخش بر اساس توضیحات داده شده در فصل ۲، پیش پردازشی در نقشه سه گانه به منظور تسهیل پیدا کردن همسایه های قابل اطمینان بیشتر که به نوبه خود تاثیرات جریان K -to- U را افزایش می دهد انجام داده می شود. ترسیم یک نقشه سه گانه مفصل کاری پر زحمت و مستعد خطا است. دوروش معرفی شده در [۲۳، ۲۵] نقشه سه گانه را به عنوان یک گام پیش پردازش به وسیله گسترش \mathcal{F} و \mathcal{B} با شروع از مرزهای آنها با U پیرایش میکنند که در [۲] پیشنهاد شده است. ترکیب این تکنیک نتایج روش جریان اطلاعات را بهبود می بخشد که در تصویر ۳-۵ (د) نشان داده شده است. همچنین این ناحیه های گسترش شده \mathcal{F} و \mathcal{B} بعد از تخمین درهم تنیدگی نیز در پس پردازش اعمال می شوند.

برای این منظور یک توزیع نرمال از تصویر به پنجره 3×3 اطراف هر پیکسل p (توسط میانگین گرفتن مقادیر پیکسل های اطراف هر پیکسل و قرار دادن آن به عنوان مقدار جدید) متناسب می شود. سپس با پیدا کردن ۲۰ نزدیکترین همسایه در ناحیه پیش زمینه و به همین ترتیب پس زمینه برای هر پیکسل در ناحیه نامشخص تصویر نرمال شده و با استفاده از ماتریس کواریانس، فاصله باتاچاریای آن پیکسل تا نزدیکترین همسایه ها تعیین میشود. با داشتن فاصله باتاچاریا، امتیاز تطابق پیش زمینه به صورت $b_p^{\mathcal{F}} = \min_{q \in \mathcal{F}} B(\mathcal{N}_p, \mathcal{N}_q)$ و امتیاز تطابق پس زمینه به صورت $b_p^{\mathcal{B}} = \min_{q \in \mathcal{B}} B(\mathcal{N}_p, \mathcal{N}_q)$ تعریف می شود که $B(.,.)$ فاصله باتاچاریا است. سپس یک ناحیه برای پیکسل p مطابق با قانون زیر انتخاب می شود:

$$p \in \begin{cases} \hat{\mathcal{F}} & \text{if } b_p^{\mathcal{F}} < \tau_c \quad \text{and} \quad b_p^{\mathcal{B}} > \tau_f \\ \hat{\mathcal{B}} & \text{if } b_p^{\mathcal{B}} < \tau_c \quad \text{and} \quad b_p^{\mathcal{F}} > \tau_f \\ \hat{U} & \text{otherwise} \end{cases} \quad (9-3)$$

یک پیکسل نامشخص در ناحیه علامت گذاری شده $\hat{\mathcal{F}}$ پس از پیرایش متعلق پیش زمینه قرار می گیرد اگر آن



(آ) تصویر ورودی (ب) نقشه سه گانه (ج) بدون پیش (د) *CSTrimm* (ه) هر دو پیرایش پردازش

شکل ۳-۵: نقشه سه گانه به صورت چند لایه که در تصویر (ب) که نواحی پیش زمینه با رنگ آبی (CS Trimming) [۲] و فیروزه ای (جستجوی تکه ای) و نواحی پس زمینه با رنگ قرمز (CS Trimming) و زرد (جستجوی تکه ای) نشان داده شده اند. (CS Trimming) نواحی کاملاً کدر یا شفاف را مشخص می کند در حالیکه پیرایش روش جریان اطلاعات آورده شده در ۳-۳، نتایج اطراف ساختارهای دور را بهبود می بخشد.

پیکسل یک همخوانی بالا در \mathcal{F} داشته و در \mathcal{B} نداشته باشد که توسط ثابتهای $\tau_c = 0.25$ و $\tau_f = 0.9$ تعیین شده اند. به این ترتیب نقشه سه گانه پیرایش و کارایی درهم تنیدگی در نواحی چالش بر انگیز افزایش داده می شود (شکل ۳-۶).

۴-۳ جریان اطلاعات درون ناحیه نامشخص

هر پیکسل منحصر به فرد در \mathcal{U} اطلاعاتی از جریان رنگ و جریان \mathcal{K} -to- \mathcal{U} دریافت می کند. علاوه بر این اطلاعات داخل \mathcal{U} را نیز می توان به طور مؤثر توزیع کرد تا بوسیله تقویت پیکسل هایی با رنگهای مشابه داخل \mathcal{U} به داشتن مقدار آلفای مشابه رسید. برای هر پیکسل در \mathcal{U} ، $K\mathcal{U} = 5$ نزدیکترین همسایه فقط در داخل \mathcal{U} با تعیین $\mathcal{N}_p^{\mathcal{U}}$ با استفاده از بردار ویژگی تعریف شده $v = [r, g, b, \hat{x}/20, \hat{y}/20]^T$ پیدا می شود.

مقدار جریان اطلاعات درون ناحیه نامشخص با استفاده از فاصله بین بردارهای ویژگی تعیین می گردد:

$$w_{p,q}^{\mathcal{U}} = \max(1 - \|v_p - v_q\|_1, 0) \quad \forall q \in \mathcal{N}_p^{\mathcal{U}} \quad (10-3)$$



شکل ۳-۶: پیش پردازش نقشه سه گانه

حال می توان اصطلاح انرژی برای جریان اطلاعات داخل \mathcal{U} را تعریف کرد:

$$E_{\mathcal{U}} = \sum_{p \in \mathcal{U}} \sum_{q \in \mathcal{N}_p^{\mathcal{U}}} w_{p,q}^{\mathcal{U}} (\alpha_p - \alpha_q)^2 \quad (11-3)$$

اطلاعات به اشتراک گذاشته بین پیکسلهای نامشخص کیفیت درهم تنیدگی را در اطراف ساختارهای پیچیده (تارهای مو) که در شکل ۳-۱(ه) نشان داده شده است افزایش می دهد.

روش KNN Matting [۸] از یک تعریف شباهت مشابه برای ساختن پیکسل های با رنگ مشابه که کدورت شبیه دارند استفاده می کند. با این حال فقط تکیه بر این فرم جریان اطلاعات، برای کل تصویر، برخی تصنیعات معمولی در نتیجه تخمین آلفا ایجاد می کند. اعتماد و وابستگی به تعریف بردار ویژگی و رنگهای تصویر نتیجه مقادیر آلفا ممکن است اشتباهاً زیر مجموعه ای هموار شود. (حالت شکل ۳-۱(ه)). وقتی همسایه های پیکسلهای داخل \mathcal{U} اغلب در \mathcal{F} یا \mathcal{B} باشند یا نواحی آلفای صاف و ثابت به جای شبیهی نامحسوس و دقیق و تیز ایجاد کنند (حالتی از شکل ۳-۱(و))، جریان اطلاعات محدود که صرفاً بر پایه شباهت رنگ اعمال شده باشد مجموعه تغییرات آلفا یا نواحی پهن با شیب آلفا را نمی تواند نشان دهد.

۳-۵ جریان اطلاعات محلی

اتصال مکانی یکی از نشانه های اصلی برای جریان اطلاعات است. هر پیکسل در \mathcal{U} به ۸ همسایه مجاورش وصل می شود و با \mathcal{N}_p^L برای تضمین درهم تنیدگی هموار مکانی نشان داده می شود. مقدار جریان اطلاعات محلی همچنین باید منطبق با لبه های قوی و فراوان در تصویر باشد. برای تعیین مقدار جریان محلی خوب، به تعریف درهم تنیدگی بر پایه شباهت پیشنهاد شده در [۷] اعتماد و از آن استفاده می شود.

درهم تنیدگی بر پایه شباهت از آمارهای تکه ای برای تعیین وزنه های $w_{p,q}^L$ ، $q \in \mathcal{N}_p^L$ استفاده می کند. اصطلاح



شکل ۳-۷: جریان IntraU به تنهایی

انرژی جریان اطلاعات محلی مرتبط به صورت زیر تعریف می گردد:

$$E_L = \sum_{p \in \mathcal{U}} \sum_{q \in \mathcal{N}_p^{\mathcal{U}}} w_{p,q}^L (\alpha_p - \alpha_q)^2 \quad (12-3)$$

علی رقم نمایش جریان اطلاعات محلی خوب، تنها درهم تنیدگی بر پایه شباهت، نواحی بزرگ تغییرات (شکل ۲-۱ بالا)، یا نواحی جدا شده که اتصال مکانی ندارند یا اتصال مکانی ضعیفی به \mathcal{F} یا \mathcal{B} دارند را نمی تواند نشان دهد (شکل ۲-۱ پایین).

۳-۶ سیستم خطی و مینیمم سازی انرژی

تابع هدف نهایی ترکیبی از ۴ انرژی تعریف شده است که هر فرم از جریان اطلاعات را نشان می دهد:

$$E_{\lambda} = E_{CM} + \sigma_{\mathcal{K}\mathcal{U}} E_{\mathcal{K}\mathcal{U}} + \sigma_{\mathcal{U}\mathcal{U}} E_{\mathcal{U}\mathcal{U}} + \sigma_L E_L + \lambda E_{\tau} \quad (13-3)$$

در رابطه بالا $\sigma_{\mathcal{K}\mathcal{U}} = 0.05$ و $\sigma_{\mathcal{U}\mathcal{U}} = 0.01$ و $\sigma_L = 1$ و $\lambda = 100$ ثابتهای الگوریتمی و تعیین کننده قوت همبستگی جریان های اطلاعات هستند و

$$E_{\tau} = \sum_{p \in \mathcal{F}} (\alpha_p - 1)^2 + \sum_{p \in \mathcal{B}} (\alpha_p - 0)^2$$

انرژی ای برای ثابت نگهداشتن مقادیر آلفاهای مشخص است. برای یک تصویر با N پیکسل، بوسیله تعریف ماتریسهای اسپارس $N \times N$ ، W_{CM} و $W_{\mathcal{U}\mathcal{U}}$ و W_L که برای جفت پیکسلهای با همبستگی جریان اطلاعات مقادیر غیر صفر دارند و بردار $w^{\mathcal{F}}$ برای $p \in \mathcal{U}$ مقدار $w_p^{\mathcal{F}}$ ، برای $p \in \mathcal{F}$ مقدار ۱ و برای $p \in \mathcal{B}$ مقدار ۰ دارد،



شکل ۳-۸: جریان اطلاعات محلی به تنهایی

میتوان معادله ۳-۱۳ را به فرم ماتریسی نوشت:

$$E_{\gamma} = \alpha^T \mathcal{L}_{IFM} \alpha + (\alpha - w^{\mathcal{F}})^T \sigma_{\mathcal{K}\mathcal{U}} \mathcal{H} (\alpha - w^{\mathcal{F}}) + (\alpha - \alpha_{\mathcal{K}})^T \lambda \tau (\alpha - \alpha_{\mathcal{K}}) \quad (14-3)$$

τ یک ماتریس قطری $N \times N$ با مقدار ورودی (p, p) برای $p \in \mathcal{K}$ برابر با ۱ و در غیر اینصورت ۰ است. \mathcal{H} یک ماتریس اسپارس با مقادیر قطری η_p که در معادله ۳-۷ تعریف شده، می باشد. $\alpha_{\mathcal{K}}$ یک بردار سطری با ورودی p^{th} که $p \in \mathcal{F}$ مقدارش ۱ و در غیر اینصورت ۰ است. α یک بردار سطری از مقادیر آلفای تخمین زده شده است و \mathcal{L}_{IFM} به صورت زیر تعریف می شود:

$$\mathcal{L}_{IFM} = (D_{CM} - W_{CM})^T (D_{CM} - W_{CM}) + \sigma_{\mathcal{U}\mathcal{U}} (D_{\mathcal{U}\mathcal{U}} - W_{\mathcal{U}\mathcal{U}}) + \sigma_L (D_L - W_L) \quad (15-3)$$

$D_{(\cdot)}(i, i)$ ماتریسی قطری برابر $\sum_j W_{(\cdot)}(i, j)$ است. با حل رابطه زیر انرژی معادله ۳-۱۴ می تواند کمینه شود:

$$(\mathcal{L}_{IFM} + \lambda \tau + \sigma_{\mathcal{K}\mathcal{U}} \mathcal{H}) \alpha = (\lambda \tau + \sigma_{\mathcal{K}\mathcal{U}} \mathcal{H}) w^{\mathcal{F}} \quad (16-3)$$

همچنین تابع هدف دومی نیز برای زمانی که از انرژی جریان اطلاعات \mathcal{K} -to- \mathcal{U} استفاده نمی شود تعریف میگردد:

$$E_{\gamma} = E_{CM} + \sigma_{\mathcal{U}\mathcal{U}} E_{\mathcal{U}\mathcal{U}} + \sigma_L E_L + \lambda E_{\tau} \quad (17-3)$$

رابطه فوق را می توان در یک فرم ماتریسی نوشت:

$$E_{\gamma} = \alpha^T \mathcal{L}_{IFM} \alpha + (\alpha - \alpha_{\mathcal{K}})^T \lambda \tau (\alpha - \alpha_{\mathcal{K}}) \quad (18-3)$$

رابطه بالا همان معادله ۱۴-۱ مطرح شده در فصل ۱-۴ روش [۷] است و با حل رابطه زیر می توان آنرا کمینه کرد:

$$(\mathcal{L}_{IFM} + \lambda\tau)\alpha = \lambda\tau\alpha_K \quad (19-3)$$

سیستم های معادلات خطی ۱۶-۳ و ۱۹-۳ را می توان با استفاده از روش های مختلف حل کرد. قبلاً گفته شد که جریان اطلاعات \mathcal{K} -to- \mathcal{U} بر روی اشیاء با شفافیت بالا تاثیر گذار نیست. برای تعیین اینکه آیا یک تصویر باید شامل جریان اطلاعات \mathcal{K} -to- \mathcal{U} باشد و E_1 را حل کند یا مانع جریان اطلاعات \mathcal{K} -to- \mathcal{U} شود و E_2 را حل کند از یک طبقه بندی ساده بر اساس هیستوگرام که تعیین کننده این است که آیا نتیجه شفافیت بالا است یا خیر، استفاده می شود. اگر شفافیت بالا باشد، انتظار می رود پیکسل های درون \mathcal{U} بیشترشان رنگهایی که ترکیبی از رنگهای \mathcal{F} و \mathcal{B} هستند را داشته باشند. از سوی دیگر اگر مقادیر واقعی آلفا برای تغییرات نرم (تصویر هموار) بیشتر ۰ یا ۱ باشند، هیستوگرام \mathcal{U} شبیه به ترکیب خطی هیستوگرام های \mathcal{F} و \mathcal{B} خواهد شد که \mathcal{U} بیشتر شامل رنگهای خیلی شبیه به \mathcal{K} خواهد بود.

بر طبق این مشاهدات هیستوگرام پیکسلهای درون \mathcal{U} ، که ترکیبی خطی از: $D_{\mathcal{F}}$ ، هیستوگرام پیکسلهای درون \mathcal{F} و $D_{\mathcal{B}}$ ، هیستوگرام پیکسلهای درون \mathcal{B} است، $D_{\mathcal{U}}$ نامیده می شود. خطا e تعریف شده که اندازه ای از چگونگی خوبی ترکیب خطی هیستوگرام واقعی را نشان می دهد:

$$e = \min_{a,b} \|aD_{\mathcal{F}} + bD_{\mathcal{B}} - D_{\mathcal{U}}\|^2 \quad (20-3)$$

مقادیر بالای e یک شفافیت بالا را نشان می دهند که در این حالت E_2 بر E_1 ترجیح داده می شود.

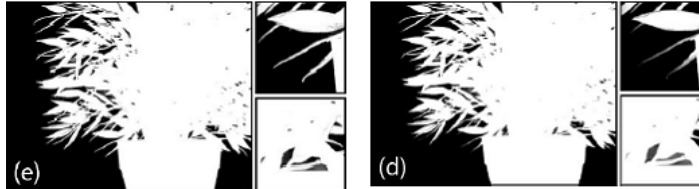
۷-۳ تنظیم برای روشهای برش بر پایه نمونه برداری

روشهای درهم تنیدگی بر پایه نمونه برداری معمولاً نمونه هایی برای هر پیکسل در \mathcal{U} که یا مستقل و یا وابستگی مکانی کمی دارند را انتخاب می کنند. به منظور بدست آوردن یک فضای درهم تنیدگی منسجم، عمل معمول، ترکیب مقادیر آلفای حدس زده اولیه با اندازه همواری است.

روشهای مختلف [۲۳، ۲۴، ۲۵، ۲] روش پس پردازش پیشنهاد شده [۲۴] را که شباهت درهم تنیدگی [۷] را با مقادیر آلفای بر پایه نمونه برداری و اطلاعات مورد اطمینان ترکیب می کند را اتخاذ کرده اند. این تکنیک پس پردازش منجر به بهبود درهم تنیدگی می شود، اما چون فقط همواری محلی را شامل می شود نتایج آن می تواند هنوز زیر مجموعه ای از شکل ۳-۹(د) باشد.



(آ) تصویر ورودی (ب) زمینه واقعی (ج) بر پایه نمونه برداری $\hat{\alpha}$ [۲] [۲۴]



(د) تنظیم شده بوسیله [۲۴] (ه) تنظیم سازی با روش جریان اطلاعات

شکل ۳-۹: روش تنظیم سازی گاستال و آلیویرا [۲۴] جزئیات دور را گم می کند (ناحیه انتخاب شده بالایی) و یا سوراخها را پر می نماید (ناحیه انتخاب شده پایینی)، در حالیکه روش تنظیم سازی جریان اطلاعات قادر است که این جزئیات بریده شده در روش بر پایه نمونه برداری را نشان دهد.

روش بحث شده در این فصل با انواع مختلف جریان اطلاعات می تواند از پس پردازش در یک روش مشابه در [۲۴] استفاده کند. مقادیر اولیه آلفای داده شده $\hat{\alpha}_p$ و $\hat{\eta}_p$ مورد اطمینان پیدا شده بوسیله یک روش بر پایه نمونه برداری، انرژی تنظیم درهم تنیدگی تعریف می گردد:

$$E_R = E_\gamma + \sigma_R \sum_{p \in \mathcal{U}} \hat{\eta}_p (\alpha_p - \hat{\alpha}_p)^2 \quad (21-3)$$

در رابطه فوق $\sigma_R = 0.05$ تعیین می کند که ثبات قدم (میزان تأثیر گذاری) برای مقادیر اولیه داده شده چقدر باید باشد. این انرژی می تواند در فرم ماتریسی نوشته شود:

$$E_R = \alpha^T \mathcal{L}_{IFM} \alpha + (\alpha - \hat{\alpha})^T \sigma_R \hat{\mathcal{H}} (\alpha - \hat{\alpha}) + (\alpha - \hat{\alpha})^T \lambda \tau (\alpha - \alpha_K) \quad (22-3)$$

و با حل رابطه زیر مینیمم می شود:

$$(\mathcal{L}_{IFM} + \lambda \tau + \sigma_R \hat{\mathcal{H}}) \alpha = (\lambda \tau + \sigma_R \hat{\mathcal{H}}) \hat{\alpha} \quad (23-3)$$

شکل ۳-۹ نشان می دهد که تنظیم درهم تنیدگی، به خصوص در اطراف ساختارهای پیش زمینه چالش بر انگیز مانند برگهای دراز یا سوراخهای دیده شده تأثیر گذار است. در فصل بعد به صورت عددی بهبود این گام تنظیم سازی بر روی چندین روش بر پایه نمونه برداری مورد کاوش قرار گرفته می شود و نتایج آن نشان داده میشود.

۸-۳ نکات پیاده سازی

برنامه های متلب روش درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی شامل توابع زیر است:

- تابع اصلی **informationFlowMatting** است که شامل پارامترهای ورودی **image** تصویر ورودی، **trimap** نقشه سه گانه، **params** ثابت هایی مانند λ و τ و K و ... و **suppressMessages** برای تعیین نمایش پیامها می باشد و پارامتر خروجی آن **alpha** است که مقدار تخمین زده شده آلفا می باشد. در ابتدای کد تابع تعیین میزان شفافیت تصویر ورودی و تابع پیش پردازش نقشه سه گانه فراخوانی میشود. سپس تابع جریان ترکیب اطلاعات فراخوانی شده و وزنه های به دست آمده در یک ماتریس خلوت از خروجی این تابع با فراخوانی تابع تبدیل ماتریس شباهت به لاپلاسیان در ماتریس لاپلاسیان **Lap** ذخیره میشود. سپس تابع جریان اطلاعات محلی فراخوانی و دوباره با همان ترتیب فوق ماتریس لاپلاسیان جدید شکل گرفته و بعد از اعمال ضرب ثابت قدم با ماتریس لاپلاسیان قبلی **Lap** جمع می شود. بعد از آن تابع جریان اطلاعات درون ناحیه نامشخص فراخوانی و نتیجه به دست آمده بعد از اعمال ضرب ثابت قدم با مقدار **Lap** قبلی جمع می شود. اگر تصویر شفافیت بالایی نداشته باشد ابتدا پیش پردازشی بر روی نقشه سه گانه توسط فراخوانی تابع پیش پردازش اعمال و سپس تابع جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص فراخوانی می شود. خروجی تابع جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص توسط اولین پیش پردازش انجام شده پیرایش می شود و در نهایت اگر شفافیت تصویر بالا نباشد تابع حل آلفا با مقادیر به دست آمده از تابع جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص و **Lap** فراخوانی میشود و در غیر اینصورت فقط تابع به دست آوردن مقادیر آلفا با مقادیر **Lap** فراخوانی و مقادیر آلفا تخمین زده میشوند.

- تابع جریان ترکیب رنگ (**colorMixtureAffinities**):

این تابع پیاده سازی روش توضیح داده شده در بخش ۳-۱ می باشد. پارامترهای ورودی **image** تصویر ورودی، **K** تعداد همسایه ها، **inMap** و **outMap** نقشه ورودی و خروجی، **xyWeight** اهمیت مختصات مکانی و **useXYinLLEcomp** استفاده از وزنه های **LLE** است و پارامتر خروجی آن **Wcm** وزنه های بدست آمده از این روش می باشند.

ابتدا با فراخوانی تابع پیدا کردن نزدیکترین همسایه، **K** نزدیکترین همسایه برای پیکسل های ناحیه نامشخص پیدا می شوند. سپس با فراخوانی تابع جاسازی خطی محلی وزنه ها به دست می آیند. در نهایت پارامتر خروجی یعنی **Wcm** ماتریسی خلوت است که در سطر اندیس پیکسل های ناحیه نامشخص و ستون اندیس پیکسل های همسایه وزنه های به دست آمده قرار می گیرد.

- تابع جریان اطلاعات محلی (**mattingAffinity**):

این تابع پیاده سازی روش توضیح داده شده در بخش ۳-۵ می باشد. پارامترهای ورودی image تصویر ورودی، inMap نقشه ورودی، windowRadius شعاع پنجره و epsilon ضریب تنظیم سازی قبل از معکوس کردن ماتریس کوواریانس می باشند و پارامتر خروجی W وزنه‌های حاصل می باشد.

ابتدا تابع توزیع نرمال رنگ فراخوانی شده و ماتریس کوواریانس و تصویر نرمال شده به دست می آید. سپس برای هر پیکسل در ناحیه نامشخص با استفاده از روش فرم بسته وزنه‌ها محاسبه می شوند. در نهایت ماتریس خلوت W ساخته و با ترانهاده خودش برای اطمینان از قطری بودن جمع می شود و سپس نرمالیزه می شود.

- تابع جریان اطلاعات داخل ناحیه نامشخص (colorSimilarityAffinities):

این تابع پیاده سازی روش توضیح داده شده در فصل ۳-۴ می باشد. پارامترهای ورودی image تصویر ورودی، K تعداد همسایه ها، inMap نقشه ورودی، outMap نقشه خروجی، xyWeight اهمیت مختصات مکانی و useHSV استفاده از فضای رنگ HSV می باشند و پارامتر خروجی Wcs وزنه‌های حاصل می باشد. ابتدا با فراخوانی تابع پیدا کردن نزدیکترین همسایه، K نزدیکترین همسایه برای پیکسل های ناحیه نامشخص پیدا می شوند. سپس توسط معادله ۳-۱۰ وزنه‌ها تعیین می گردند و در نهایت خروجی Wcs که ماتریسی خلوت است شکل می گیرد.

- تابع جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص (knownToUnknownColorMixture):

این تابع پیاده سازی روش توضیح داده شده در بخش ۳-۲ می باشد. پارامترهای ورودی image تصویر ورودی، trimap نقشه سه گانه پیرایش شده، K تعداد همسایه ها و xyWeight اهمیت مختصات مکانی می باشند و پارامترهای خروجی alphaEst آلفاهای تخمین زده شده و conf میزان اعتماد به آنها می باشند. ابتدا تابع پیدا کردن نزدیکترین همسایه ها فراخوانی و K نزدیکترین همسایه در پیش زمینه پیدا می شود و به همین ترتیب K نزدیکترین همسایه در پس زمینه نیز پیدا می شود. سپس با فراخوانی تابع جاسازی خطی محلی وزنه‌ها تعیین شده مقادیر تخمین آلفاها به دست می آیند. توسط معادله ۳-۶ میزان مشارکت رنگها در پیش زمینه و پس زمینه بدست آمده و توسط معادله ۳-۷ میزان اعتماد مشخص می گردد.

- تابع بدست آوردن مقادیر آلفا (solveForAlphas):

این تابع پیاده سازی روش توضیح داده شده در انتهای بخش ۳-۶ می باشد. پارامترهای ورودی Lap ماتریس لاپلاسیان به دست آمده از فرم های جریان اطلاعات، trimap نقشه سه گانه، lambda ثابت الگوریتمی، usePCG تعیین حل با روش گرادینان مزدوج، alphaHat آلفاهای تخمین زده شده در روش جریان اطلاعات ناحیه مشخص به نامشخص، conf میزان اعتماد به آنها و aHatMult ثابت الگوریتمی و خروجی تابع alphas مقادیر تخمین زده شده نهایی آلفا می باشد.

- تابع تعیین میزان شفافیت تصویر (**detectHighlyTransparent**):

این تابع پیاده سازی روش توضیح داده شده در انتهای بخش ۳-۶ می باشد. پارامترهای ورودی `image` تصویر اصلی و `trimap` نقشه سه گانه می باشند و پارامتر خروجی آن `ht` شفاف بودن یا نبودن تصویر است. با تعیین هیستوگرام تصویر در پیش زمینه و پس زمینه میزان شفافیت تعیین و با یک مقدار آستانه برای شفاف بودن یا نبودن تصویر مقایسه می شود.

- تابع پیرایش نقشه سه گانه (**trimmingFromKnownUnknownEdges**):

این تابع پیاده سازی روش توضیح داده شده در فصل ۲ می باشد. پارامترهای ورودی `image` تصویر اصلی، `trimap` نقشه سه گانه، `paramU` ماکزیمم آستانه رنگ، `paramD` تعیین میزان کم شدن آستانه در هر تکرار و `iterCnt` تعداد تکرار می باشند و پارامتر خروجی آن `trimap` نقشه سه گانه پیرایش شده است. ابتدا لبه های نواحی مشخص تعیین گردیده و فاصله اقلیدوسی هر پیکسل جزو لبه از پنجره اطرافش به دست می آید. اگر فاصله به دست آمده از حد آستانه کمتر باشد به عنوان نمونه کاندید در نظر گرفته می شود. آلفای کمترین مقدار در بین نمونه های کاندید به جای مقدار آلفای جدید پیکسل روی لبه قرار میگیرد.

- تابع پیدا کردن همسایه ها (**findNonlocalNeighbors**):

پارامترهای ورودی این تابع `image` تصویر ورودی، `K` تعداد همسایه ها، `xyWeight` اهمیت مختصات مکانی، `inMap` نقشه ورودی، `outMap` نقشه خروجی، `eraseSelfMatches` حذف خود تطابقی میباشد و پارامترهای خروجی آن `inInd` اندیس پیکسل های ناحیه نامشخص، `neighInd` اندیس پیکسل های همسایه و `features` ویژگی ها می باشند.

این تابع توسط دستور `knnsearch` متلب `K` همسایه `inMap` داخل `outMap` پیدا می کند.

- تابع جاسازی خطی محلی (**localLinearEmbedding**):

این تابع پیاده سازی یادداشت شماره ۷ مرجع [۶] می باشد. پارامترهای ورودی `pt` پیکسل، `neighbors` همسایه های آن و `conditionerMult` ضریب مقید سازی می باشند و پارامتر خروجی آن `w` وزنه های محاسبه شده است.

ابتدا ماتریس همبستگی از ضرب بردار همسایه از ترانهاده خودش محاسبه و ضریب مقید سازی به تمام عناصر اضافه می شود. سپس بازسازی پیکسل از طریق ضرب پیکسل داده شده در بردار همسایه صورت می گیرد. بعد از آن ضریب لاگرانژ λ محاسبه شده و به پیکسل بازسازی شده اضافه می شود. از تقسیم این حاصل بر ماتریس همبستگی وزنها به دست می آیند.

- تابع توزیع نرمال محلی رنگ (**localRGBnormalDistributions**):

پارامترهای ورودی `image` تصویر ورودی، `inMap` نقشه ورودی، `windowRadius` شعاع پنجره و `epsilon` ضریب تنظیم سازی قبل از معکوس کردن ماتریس کوواریانس می باشند و پارامترهای خروجی `meanImage` تصویر نرمال شده و `covarMat` ماتریس کوواریانس می باشند.

ابتدا توسط دستور `imboxfilt` متلب تصویر ورودی نرمال می شود و با تصویر ورودی و تصویر نرمال شده ماتریس کوواریانس محاسبه می شود.

تابع پیرایش نقشه سه گانه (`patchBasedTrimming`):

این تابع پیاده سازی روش توضیح داده شده در بخش ۳-۳ می باشد. پارامترهای ورودی `image` تصویر اصلی، `trimap` نقشه سه گانه، `minDist` و `maxDist` ثابت های T_c و T_f در معادله ۳-۹ هستند و `windowsRadius` شعاع پنجره و `K` تعداد همسایه ها می باشند. پارامتر خروجی `trimap` نقشه سه گانه پیرایش شده است.

ابتدا تابع توزیع نرمال رنگ فراخوانی شده و ماتریس کوواریانس و تصویر نرمال شده به دست می آید. سپس با فراخوانی تابع پیدا کردن نزدیکترین همسایه `K` نزدیکترین همسایه به پیکسل های نامشخص در ناحیه پیش زمینه از تصویر نرمال شده پیدا شده و به همین ترتیب در ناحیه پس زمینه نیز همین تعداد همسایه پیدا می شود. با محاسبه فاصله باتاچاریا و امتیاز تطابق پیش زمینه و پس زمینه نقشه سه گانه پیرایش می شود.

• تابع تبدیل ماتریس شباهت به لاپلاسیان (`affinityMatrixToLaplacian`):

پارامتر ورودی این تابع `aff` ماتریسی خلوت شامل وزنها و پارامتر خروجی `Lap` ماتریس لاپلاسیان است.

فصل ۴

نتایج پیاده سازی

۱-۴ بحث و مقایسه

به طور کمی الگوریتم توضیح داده شده در فصل ۳ با استفاده از معیار درهم تیدگی آلفای عمومی^۲ [۲۹] ارزیابی شده است. در زمان ارسال مرجع [۱]، رتبه های این روش برحسب تفاوت های مجموع قدر مطلق ها (SAD)^۳ و ماتریسهای میانگین های خطا (MSE)^۴ در جایگاه اول قرار دارد. نتایج در جدول ۴-۱ دیده می شوند. کد این روش که در متلب نوشته شده با میانگین ۵۰ ثانیه تصویر را پردازش می کند.

مقایسه بصری چند روش برای تصویر عروسک در تصویر ۴-۱ آمده است. برای همه روشها بجز [۹] که نگارنده مرجع [۱] پیاده سازی کرده است از نتایجی که برای معیار درهم تیدگی موجود هستند استفاده شده است. تصویر ۴-۱ استفاده از فقط یکی از فرمهای جریان اطلاعات که در تعداد سناریوها همانند نواحی نامشخص وسیع یا سوراخها در اشیاء پیش زمینه موثر نیستند را نشان می دهد.

استراتژی درهم تیدگی DCNN [۲۸] مستقیماً از نتایج فرم بسته و درهم تیدگی چند همسایگی و نه از فرمول ترکیب انرژی استفاده شده در تعریف شباهت پیروی می کند. هر دو روش با شکست روبرو می شوند، نتیجه ترکیب نیز خطاهایی آشکار در نمونه های Pineapple و Troll دارد. شبکه عصبی ای که مرجع [۲۸] مطرح کرده است درهم تیدگی هایی مات تهیه می کند. از سوی دیگر درهم تیدگی LNSP [۲۶] اطراف نواحی سوراخدار (مانند Pineapple) یا وقتی رنگهای پیش زمینه و پس زمینه شبیه هستند (مانند Troll و Donkey) مشکلاتی دارد (تصویر ۴-۳).

جدول ۴-۱ امتیازات روش جریان اطلاعات در معیار درهم تیدگی آلفا [۲۹] به همراه بهترین روشهای منتشر شده از لحاظ کارایی در زمان ارسال مرجع [۱]. S، L و U نقشه سه گانه در سه معیار به ترتیب کوچک، بزرگ و

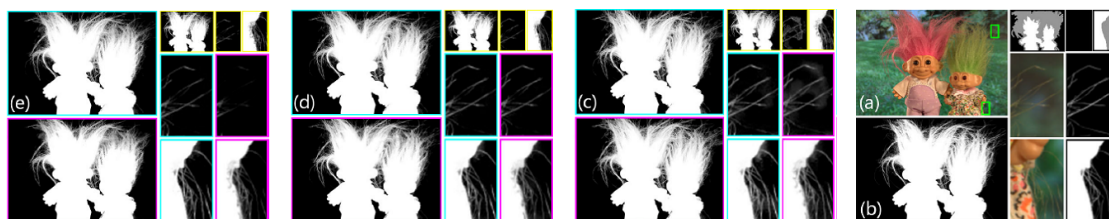
^۲public alpha matting benchmark

^۳Sum Absolute Difference

^۴Mean Squared Error

جدول ۴-۱: مقایسه روش جریان اطلاعات با چند روش قوی درهم تنیدگی توسط ۸ نمونه

روش	میانگین رتبه																											
	Net			bag Plastic			Pineapple			Plant			Elephant			Donkey			Doll			Troll			overall			
	U	L	S	U	L	S	U	L	S	U	L	S	U	L	S	U	L	S	U	L	S	U	L	S	U	L	S	
مجموع قدر مطلق تفاوت ها																												
جریان اطلاعات	۲۲.۳	۲۲.۲	۲۰.۲	۱۵/۸	۱۹.۳	۱۹.۳	۴/۶	۵/۷	۳/۶	۸/۶	۷/۱	۵/۹	۲/۰	۲/۳	۱/۲	۳/۰	۴/۱	۳/۸	۷/۳	۷/۳	۵/۶	۱۲/۵	۱۱/۲	۱۰/۳	۲/۰	۱/۶	۲/۸	۲/۱
DCNN [۲۸]	۲۱.۲	۲۰.۲	۱۹.۴	۱۹.۱	۱۹.۲	۱۹.۹	۵/۳	۶/۰	۴/۰	۹/۱	۶/۹	۶/۰	۲/۲	۲/۵	۱/۶	۳/۴	۴/۵	۳/۹	۶/۸	۶/۴	۵/۳	۱۴/۵	۱۴/۱	۱۲/۰	۳/۴	۱/۶	۴/۶	۳/۰
CSC [۲۳]	۲۵.۲	۲۷.۲	۲۶.۳	۲۱.۰	۲۳.۰	۲۳.۷	۶/۹	۷/۶	۴/۶	۹/۷	۷/۳	۵/۵	۲/۵	۲/۷	۱/۸	۴/۲	۴/۸	۴/۶	۸/۱	۷/۵	۶/۲	۱۴/۵	۱۵/۶	۱۳/۶	۱۰/۳	۶/۴	۱۳/۵	۱/۰
LNSP [۲۶]	۲۷.۸	۲۴.۴	۲۲.۵	۱۶.۳	۲۰.۸	۲۱.۵	۶/۴	۷/۱	۴/۰	۱۰/۷	۸/۱	۶/۲	۳/۱	۳/۵	۱/۵	۳/۶	۵/۹	۴/۶	۸/۸	۸/۱	۵/۶	۱۹/۵	۲۲/۵	۱۲/۲	۱۴/۶	۱۰/۳	۱۰/۷	۱۰/۷
مجموع مربعات خطا																												
جریان اطلاعات	۰.۹	۰.۸	۰.۸	۰.۸	۱.۲	۱.۳	۰.۳	۰.۲	۰.۲	۰.۶	۰.۴	۰.۴	۰.۱	۰.۱	۰.۱	۰.۲	۰.۳	۰.۳	۰.۵	۰.۴	۰.۳	۰.۵	۰.۴	۰.۲	۲.۴	۲.۱	۵.۰	۳.۵
DCNN [۲۸]	۰.۹	۰.۷	۰.۷	۱.۰	۱.۲	۱.۳	۰.۳	۰.۴	۰.۲	۰.۸	۰.۴	۰.۴	۰.۱	۰.۱	۰.۱	۰.۲	۰.۳	۰.۳	۰.۴	۰.۳	۰.۲	۰.۷	۰.۵	۰.۴	۴.۶	۱.۹	۴.۵	۳.۷
LNSP [۲۶]	۱.۵	۱.۱	۱.۰	۰.۸	۱.۲	۱.۴	۰.۴	۰.۳	۰.۲	۰.۸	۰.۵	۰.۴	۰.۲	۰.۱	۰.۱	۰.۴	۰.۳	۰.۳	۰.۵	۰.۴	۰.۲	۱.۲	۱.۹	۰.۵	۱۲.۴	۸.۶	۶.۶	۹.۲
KL-D [۲۵]	۰.۹	۰.۸	۰.۸	۲.۱	۲.۰	۱.۷	۰.۶	۰.۶	۰.۴	۱.۲	۰.۴	۰.۴	۰.۱	۰.۲	۰.۱	۰.۳	۰.۳	۰.۳	۰.۵	۰.۵	۰.۳	۰.۷	۰.۹	۰.۴	۱۳.۳	۱۰.۵	۱۱.۰	۱۱.۶



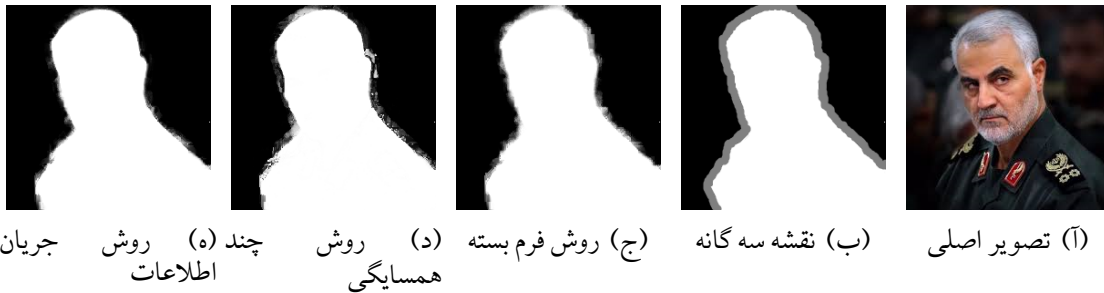
شکل ۴-۱: مقایسه روش جریان اطلاعات با چند روش قوی درهم تنیدگی توسط ۸ نمونه

شکل ۴-۱: استفاده از گام تنظیم سازی برای روش ۳ بر پایه نمونه برداری

کاربر است. اعداد با زمینه خاکستری، بهترین امتیازات معیارها را نشان می دهند. (امتیازات نشان داده شده در اینجا فقط منحصر با زمان ارسال می باشد. می توانید رتبه ها را به صورت بروز از سایت Alphamatting.com تهیه کرده و مشاهده کنید).

تصویر ۴-۱ یک مثال برای منظم سازی توضیح داده شده در ۳-۷ استفاده شده در سه روش بر پایه نمونه برداری را نشان می دهد. تنظیم سازی با استفاده از روش جریان اطلاعات با رنگ (فیروزه ای)، روش [۲۴] (ارغوانی) و روش بر پایه نمونه برداری (زرد) نشان داده شده است. روش جریان اطلاعات قادر است اجزاء دور از دسترس را هنگامی که درهم تنیدگی تولید شده (قسمت انتخاب شده بالایی) و میزان وضوح تصاویر را حتی در اطراف مناطق بافت حفظ کند (قسمت انتخاب شده پایینی). اطلاعاتی که از مقادیر آلفا و میزان اعتماد به آنها بوسیله روش بر پایه نمونه برداری بدست آمده بوسیله روش جریان اطلاعات موثرتر توزیع شده است. نواحی چالش بر انگیز مانند سوراخها بوسیله روش های بر پایه نمونه برداری وقتی روش جریان اطلاعات برای پس پردازش استفاده شده است شناسایی شده و حفظ می گردند.

در تصویر ۴-۲ (آ) رنگ پیش زمینه در بعضی نقاط شبیه به رنگ پس زمینه می باشد که عمل درهم تنیدگی ممکن است نتایج خوبی تولید نکند. تصویر ۴-۲ (ب) نقشه سه گانه به صورت دستی تهیه شده توسط گردآورنده می باشد. تصاویر ۴-۲ (ج) و ۴-۲ (د) نتایج دو روش مشهور مطرح شده در فصل ۱ را نشان می دهند. برتری کیفیت روش درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی در تصویر ۴-۲ (ه) نسبت به دو تصویر قبلی مشهود است.



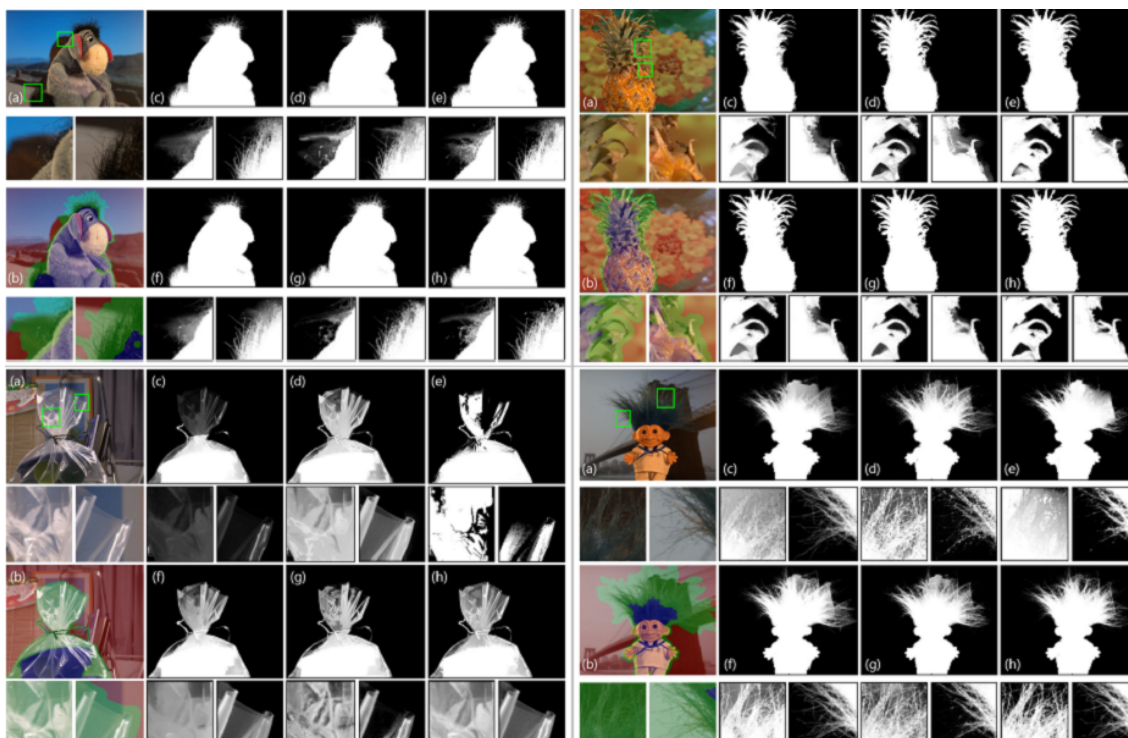
(آ) تصویر اصلی (ب) نقشه سه گانه (ج) روش فرم بسته (د) روش همسایگی (ه) چند (ه) روش جریان اطلاعات

شکل ۴-۲: مقایسه روش جریان اطلاعات با روش های فرم بسته و نزدیکترین همسایگی

۲-۴ طبقه بندی مبتنی بر درهم تنیدگی

در این بخش چندین مجموعه داده غیر تصویری مطابق با جدول؟؟ شامل ۲ کلاس ۰ و ۱ و تعداد ویژگی های متفاوت با روش های KNN، SVM و جریان اطلاعات طبقه بندی و نتایج آنها با هم از لحاظ دقت و زمان اجرا مقایسه شده و در جداول؟؟ تا ۴-۵ نشان داده شده است. برای انجام طبقه بندی از روش k-fold cross Validation استفاده می شود. در این روش داده ها بر اساس نمونه های آموزشی تعیین شده به چندین قسمت تقسیم شده و هر بار یک قسمت از داده ها به عنوان داده های آموزشی و بقیه به عنوان داده های تست در نظر گرفته می شود. برای پیاده سازی طبقه بندی با استفاده از روش جریان اطلاعات، داده های آموزشی در نقشه سه گانه به عنوان ناحیه مشخص (کلاس ۱ پیش زمینه و کلاس ۰ پس زمینه) و داده های تست به عنوان ناحیه نامشخص (مقدار ۰.۵) تعریف می شوند. همچنین از ویژگی مکان نیز صرف نظر می شود. با انجام طبقه بندی کلاس داده های تست مشخص می گردد.

در جداول؟؟ تا ۴-۵ برای بررسی طبقه بندی داده ها با روش جریان اطلاعات هر بار یکی از جریانهای توضیح داده شده در فصل ۳ حذف و روش بدون آن اجرا شده است. ستون های جداول به ترتیب SVM روش طبقه بندی بردار پشتیبان، KNN روش طبقه بندی چندین همسایگی، InfFlow-all استفاده از روش جریان اطلاعات برای طبقه بندی، InfFlow-noCM روش جریان اطلاعات بدون ترکیب رنگ، InfFlow-noLoc روش جریان اطلاعات بدون جریان محلی، InfFlow-noIU روش جریان اطلاعات بدون جریان اطلاعات داخل ناحیه نامشخص، InfFlow-noKtoU روش جریان اطلاعات بدون جریان اطلاعات ناحیه مشخص به نامشخص و InfFlow-onlyIU روش جریان اطلاعات فقط با استفاده از جریان اطلاعات داخل ناحیه نامشخص هستند. نتایج نشان می دهد که در بعضی از داده های آموزشی طبقه بندی روش جریان اطلاعات فقط با استفاده از جریان داخل ناحیه نامشخص می تواند مؤثر باشد.



شکل ۴-۳: چندین مثال برای معیار درهم تنیدگی آلفا [۲۹]. a تصویر اصلی، b نقشه سه گانه، c روش فرم بسته، d روش چند همسایگی، e روش حفظ خمینه، f روش LNSP، g روش DCNN، h روش جریان اطلاعات.

شماره	مجموعه داده	تعداد ویژگی ها	تعداد نمونه ها	تعداد نمونه های کلاس صفر	تعداد نمونه های کلاس یک
۱	Fourcluster	۲	۱۴۴	۸۰	۶۴
۲	Twomoon	۲	۲۲۵	۱۰۰	۱۲۵
۳	Tworing	۲	۲۲۵	۱۳۷	۸۸
۴	Fisheriris	۴	۱۰۰	۵۰	۵۰
۵	Ionospher	۳۴	۳۲۴	۱۲۶	۱۹۸
۶	Sonar	۸۰	۱۹۶	۹۷	۹۹

۳-۴ نتیجه گیری

این پایان نامه روش درهم تنیدگی بر پایه شباهت با جریان ترکیب رنگ طراحی شده از وزنهای LLE برای درهم تنیدگی تصویر را معرفی کرد و با طراحی دقیق جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص به خوبی توزیع اطلاعات داخل ناحیه نامشخص، چندین چالش معمول در درهم تنیدگی بهبود بخشید. همچنین نشان داده شد سیستم خطی ای که فرموله شد از بروزترین کارها در معیار درهم تنیدگی آلفا بهتر عمل می کند. نشان داده شد که فرمول یاد شده می تواند عموماً جایگزین گام اصلاح درهم تنیدگی در روشهای درهم تنیدگی بر پایه نمونه برداری برای رسیدن و افزایش کیفیت نهایی درهم تنیدگی شود [۱].

MethodName								No.
InfFlow-onlyIU	InfFlow-noKtoU	InfFlow-noIU	InfFlow-noLoc	InfFlow-noCM	InfFlow-all	KNN	SVM	No.
○/۹۳۱	○/۴۶۶	○/۴۲۱	○/۴۰۲	○/۴۴۰	○/۴۲۲	۱/۰۰۰	○/۹۰۲	۱
○/۹۵۱	○/۳۵۱	○/۳۶۶	○/۴۰۵	○/۴۷۹	○/۳۶۵	۱/۰۰۰	○/۶۰۵	۲
○/۹۱۱	○/۴۶۳	○/۳۸۵	○/۳۴۴	○/۳۹۲	○/۳۸۳	۱/۰۰۰	○/۷۰۹	۳
○/۸۳۳	○/۴۱۳	○/۴۷۰	○/۷۰۰	○/۴۶۰	○/۴۷۰	۱/۰۰۰	۱/۰۰۰	۴
○/۹۴۴	○/۴۹۶	○/۵۰۲	○/۵۹۲	○/۴۶۹	○/۴۸۸	○/۸۴۶	○/۸۶۱	۵
○/۹۲۸	○/۴۰۷	○/۴۵۰	○/۴۱۶	○/۳۹۲	○/۴۱۲	○/۷۸۱	○/۷۹۶	۶
○/۹۱۶	○/۴۳۳	○/۴۳۲	○/۴۷۷	○/۴۳۹	○/۴۲۳	○/۹۳۸	○/۸۱۲	Avg

جدول ۴-۲: دقت برای ۳۰ نمونه تست

MethodName								No.
InfFlow-onlyIU	InfFlow-noKtoU	InfFlow-noIU	InfFlow-noLoc	InfFlow-noCM	InfFlow-all	KNN	SVM	No.
○/۲۵۲	○/۲۷۳	○/۳۳۴	○/۳۲۹	○/۳۲۸	○/۳۴۰	○/۰۰۰۸	○/۰۲۳	۱
○/۲۶۱	○/۲۷۰	○/۳۵۶	○/۳۳۵	○/۳۳۸	○/۳۴۲	○/۰۰۰۹	○/۰۲۷	۲
○/۲۴۸	○/۲۶۳	○/۳۳۱	○/۳۱۷	○/۳۲۲	○/۳۱۸	○/۰۰۰۸	○/۰۲۳	۳
○/۲۵۱	○/۲۶۵	○/۳۲۳	○/۳۱۴	○/۳۲۱	○/۳۲۸	○/۰۰۰۸	○/۰۲۰	۴
○/۲۶۲	○/۵۵۹	○/۹۴۹	○/۶۵۹	○/۹۳۴	○/۹۵۵	○/۰۰۲۲	○/۰۳۲	۵
○/۲۵۷	۱/۱۱۸	۲/۰۸۰	۱/۱۹۷	۲/۰۱۶	۲/۰۳۴	○/۰۰۰۹	○/۰۲۳	۶
○/۲۵۵	○/۴۵۸	○/۷۲۹	○/۵۲۵	○/۷۱۰	○/۷۱۹	○/۰۰۱۱	○/۰۲۵	Avg

جدول ۴-۳: زمان اجرا برای ۳۰ نمونه تست

MethodName								No.
InfFlow-onlyIU	InfFlow-noKtoU	InfFlow-noIU	InfFlow-noLoc	InfFlow-noCM	InfFlow-all	KNN	SVM	No.
○/۸۶۱	○/۴۴۰	○/۳۸۷	○/۳۶۱	○/۳۵۸	○/۳۸۴	۱/۰۰۰	○/۹۰۲	۱
○/۹۰۱	○/۲۱۴	○/۳۰۴	○/۴۵۶	○/۳۹۰	○/۳۰۳	۱/۰۰۰	○/۶۰۴	۲
○/۷۷۸	○/۳۹۲	○/۳۶۸	○/۳۶۱	○/۳۶۵	○/۳۶۵	۱/۰۰۰	○/۶۶۷	۳
○/۷۵۰	○/۴۲۰	○/۴۴۵	○/۵۲۵	○/۳۷۵	○/۴۴۰	۱/۰۰۰	۱/۰۰۰	۴
○/۸۷۸	○/۴۶۶	○/۴۵۴	○/۵۵۷	○/۳۶۱	○/۴۴۳	○/۸۳۷	○/۸۴۹	۵
○/۸۳۲	○/۳۱۸	○/۳۹۱	○/۳۴۷	○/۲۹۹	○/۳۲۸	○/۷۶۵	○/۷۵۵	۶
○/۸۳۳	○/۳۷۳	○/۳۹۲	○/۴۳۵	○/۳۵۸	○/۳۷۷	○/۹۳۴	○/۷۹۶	Avg

جدول ۴-۴: دقت برای ۶۰ نمونه تست

MethodName								No.
InfFlow-onlyIU	InfFlow-noKtoU	InfFlow-noIU	InfFlow-noLoc	InfFlow-noCM	InfFlow-all	KNN	SVM	No.
○/۲۵۰	○/۲۷۱	○/۳۴۳	○/۳۴۶	○/۳۳۸	○/۳۴۴	○/۰۰۰۹	○/۰۲۱	۱
○/۲۵۸	○/۲۷۸	○/۳۶۲	○/۳۴۱	○/۳۳۷	○/۳۵۱	○/۰۰۰۸	○/۰۲۳	۲
○/۲۴۲	○/۲۶۸	○/۳۳۶	○/۳۳۰	○/۳۲۴	○/۳۳۴	○/۰۰۰۷	○/۰۲۰	۳
○/۲۴۵	○/۲۶۳	○/۳۲۲	○/۳۲۲	○/۳۳۱	○/۳۲۷	○/۰۰۰۹	○/۰۲۰	۴
○/۲۴۷	○/۵۶۳	○/۹۷۱	○/۶۷۱	○/۹۴۵	○/۹۹۷	○/۰۱۰	○/۰۲۹	۵
○/۲۴۹	۱/۰۹۹	۲/۰۸۶	۱/۳۶۱	۲/۰۶۱	۲/۰۸۵	○/۰۱۰	○/۰۲۲	۶
○/۲۴۹	○/۴۵۷	○/۷۳۷	○/۵۶۲	○/۷۲۲	○/۷۴۰	○/۰۰۰۹	○/۰۲۲	Avg

جدول ۴-۵: زمان اجرا برای ۶۰ نمونه تست

فهرست منابع

- [1] Aksoy, Yagız, Aydın, Tunç Ozan, Pollefeys, Marc, and Zürich, ETH. Designing effective inter-pixel information flow for natural image matting. in *Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [2] Shahrian, Ehsan, Rajan, Deepu, Price, Brian, and Cohen, Scott. Improving image matting using comprehensive sampling sets. in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 636–643. IEEE, 2013.
- [۳] محمد، صنیعی آبا،. داده کاوی کاربردی. انتشارات نیاز دانش، ۱۳۹۳.
- [۴] گلزاریان محمودرضا، کاظمی فاطمه، حاجی ابوالحسنی زینب. پردازش تصویر از اصول تا اجرا با کمک نرم افزار متلب. انتشارات دانشگاه فردوسی مشهد، ۱۳۹۳.
- [5] Wang, Jue and Cohen, Michael F. Image and video matting: A survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(2):97–175, 2008.
- [6] Roweis, Sam T and Saul, Lawrence K. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.
- [7] Levin, Anat, Lischinski, Dani, and Weiss, Yair. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):228–242, 2008.
- [8] Chen, Qifeng, Li, Dingzeyu, and Tang, Chi-Keung. Knn matting. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2175–2188, 2013.
- [9] Chen, Xiaowu, Zou, Dongqing, Zhao, Qinqing, and Tan, Ping. Manifold preserving edit propagation. *ACM Transactions on Graphics (TOG)*, 31(6):132, 2012.
- [۱۰] رشید آبادی فرزانه، امین طوسی محمود. برش هوشمند تصویر. پایان نامه کارشناسی ارشد، دانشکده ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری، ۱۳۹۴.
- [11] Buades, Antoni, Coll, Bartomeu, and Morel, Jean-Michel. Nonlocal image and movie denoising. *International journal of computer vision*, 76(2):123–139, 2008.
- [12] Lee, Philip and Wu, Ying. Nonlocal matting. in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 2193–2200. IEEE, 2011.

- [13] He, Kaiming, Sun, Jian, and Tang, Xiaoou. Fast matting using large kernel matting laplacian matrices. in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 2165–2172. IEEE, 2010.
- [14] Barrett, Richard, Berry, Michael W, Chan, Tony F, Demmel, James, Donato, June, Dongarra, Jack, Eijkhout, Victor, Pozo, Roldan, Romine, Charles, and Van der Vorst, Henk. *Templates for the solution of linear systems: building blocks for iterative methods*, vol. 43. Siam, 1994.
- [15] Singaraju, Dheeraj and Vidal, Rene. Estimation of alpha mattes for multiple image layers. *IEEE transactions on pattern analysis and machine intelligence*, 33(7):1295–1309, 2011.
- [16] Levin, Anat, Rav-Acha, Alex, and Lischinski, Dani. Spectral matting. *IEEE transactions on pattern analysis and machine intelligence*, 30(10):1699–1712, 2008.
- [17] He, Kaiming, Rhemann, Christoph, Rother, Carsten, Tang, Xiaoou, and Sun, Jian. A global sampling method for alpha matting. in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 2049–2056. IEEE, 2011.
- [18] Aksoy, Yağiz, Aydin, Tunç Ozan, Pollefeys, Marc, and Smolić, Aljoša. Interactive high-quality green-screen keying via color unmixing. *ACM Transactions on Graphics (TOG)*, 35(5):152, 2016.
- [19] Aksoy, Yağiz, Aydin, Tunç Ozan, Smolić, Aljoša, and Pollefeys, Marc. Unmixing-based soft color segmentation for image manipulation. *ACM Transactions on Graphics (TOG)*, 36(2):19, 2017.
- [20] Tan, Jianchao, Lien, Jyh-Ming, and Gingold, Yotam. Decomposing images into layers via rgb-space geometry. *ACM Transactions on Graphics (TOG)*, 36(1):7, 2017.
- [21] Shih, YiChang, Krishnan, Dilip, Durand, Fredo, and Freeman, William T. Reflection removal using ghosting cues. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3193–3201, 2015.
- [22] Pan, Jinshan, Hu, Zhe, Su, Zhixun, Lee, Hsin-Ying, and Yang, Ming-Hsuan. Soft-segmentation guided object motion deblurring. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 459–468, 2016.
- [23] Feng, Xiaoxue, Liang, Xiaohui, and Zhang, Zili. A cluster sampling method for image matting via sparse coding. in *European Conference on Computer Vision*, pp. 204–219. Springer, 2016.
- [24] Gastal, Eduardo SL and Oliveira, Manuel M. Shared sampling for real-time alpha matting. in *Computer Graphics Forum*, vol. 29, pp. 575–584. Wiley Online Library, 2010.

- [25] Karacan, Levent, Erdem, Aykut, and Erdem, Erkut. Image matting with kl-divergence based sparse sampling. in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 424–432, 2015.
- [26] Chen, Xiaowu, Zou, Dongqing, Zhou, Steven Zhiying, Zhao, Qinpeng, and Tan, Ping. Image matting with local and nonlocal smooth priors. in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pp. 1902–1907. IEEE, 2013.
- [27] Wang, Jue and Cohen, Michael F. Optimized color sampling for robust matting. in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–8. IEEE, 2007.
- [28] Cho, Donghyeon, Tai, Yu-Wing, and Kweon, Inso. Natural image matting using deep convolutional neural networks. in *European Conference on Computer Vision*, pp. 626–643. Springer, 2016.
- [29] Rhemann, Christoph, Rother, Carsten, Wang, Jue, Gelautz, Margrit, Kohli, Pushmeet, and Rott, Pamela. A perceptually motivated online benchmark for image matting. in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 1826–1833. IEEE, 2009.

پیوست آ

کدهای مربوط به پیاده سازی روش درهم تنیدگی تصویر مبتنی بر جریان اطلاعات بین پیکسلی در متلب برگرفته از سایت

www.alphamattng.com

در بین هر یک از توابع کدهایی که اهمیت بیشتری دارند توضیح داده شده است.

آ-۱ تابع اصلی

برنامه آ-۱: کد MATLAB تابع اصلی

```
% تابع زیر روش درهم تنیدگی تصویر توضیح داده شده در این پایان نامه را پیاده سازی می کند.
% پارامتر ورودی 'params' می تواند ویرایش شود. مقدار پیش فرض آن 'IFM' است.
% پارامتر 'useKnownToUnknown' تصمیم می گیرد که تابع هدف E1 یا E2 توضیح داده شده در پایان نامه اجرا شود.
% یک مقدار منفی نشانه تصمیم گیری اتوماتیک می باشد.
% پارامترهای **K تعداد همسایه های غیر محلی جستجو هستند.
% هنگامیکه پارامترهای **xyw تعریف شوند مختصات فضایی برای جریانهای ترکیب رنگ و K-to-U و داخل U، تاثیر
% گذار خواهد بود.
% پارامترهای **mult وزن هر جریان اطلاعات را نشان می دهند.
% ***loc پارامترهای درهم تنیدگی لاپلاسی هستند.
% -mattePostTrim تعیین می کند که پیرایش بر اساس لبه به عنوان پس پردازش در تخمین آلفا اعمال شود یا خیر.
% مقدار پیش فرض در اینجا false است که در مقاله مرجع اصلی true گزارش شده است.
```

```

function alpha = informationFlowMatting(image, trimap, params,
    suppressMessages)
    abmtSetup
    tic;
    if ~exist('image','var')
        image = imread('training4.png');
    end
    if ~exist('trimap','var')
        trimap = imread('training4trimap.png');
    end
    if ~exist('params', 'var') || isempty(params)
        params = getMattingParams('IFM');
    end
    if ~exist('suppressMessages', 'var') || isempty(suppressMessages)
        suppressMessages = false;
    end
    if(~suppressMessages) display('Information-Flow Matting started...')
    end

    image = im2double(image);
    trimap = im2double(trimap(:,:,1));

    %%%%% تعیین می کند که از جریان K-to-U استفاده شود یا خیر
    if params.useKnownToUnknown < 0
        useKU = ~detectHighlyTransparent(image, trimap);
    else
        useKU = params.useKnownToUnknown > 0;
    end
    if(~suppressMessages)
        if useKU
            display('      Known-to-unknown information flow will be
                used.');
```

```

if(~suppressMessages) display('      Computing matting Laplacian...')
); end
%%%%% جریان اطلاعات محلی
Lap = Lap + params.loc_mult * affinityMatrixToLaplacian(
    mattingAffinity(image, dilUnk, params.loc_win, params.loc_eps))
;
if(~suppressMessages) display('      Computing intra-U flow...');
end
%%%%% U جریان اطلاعات داخل U
Lap = Lap + params.iu_mult * affinityMatrixToLaplacian(
    colorSimilarityAffinities(image, params.iu_K, unk, unk, params.
    iu_xyw));

if useKU
    %%%%% K-to-U محاسبه جریان
    if(~suppressMessages) display('      Trimming trimap using patch
        similarity...'); end
    %%%%% پیش پردازش
    patchTrimmed = patchBasedTrimming(image, trimap, [], [], [], 5)
    ;
    %%%%% در اینجا k=5 برای داشتن زمان محاسبه کمتر گرفته می شود.
    if(~suppressMessages) display('      Computing K-to-U flow...');
    end
    %%%%% جریان اطلاعات K-to-U
    [kToU, kToUconf] = knownToUnknownColorMixture(image,
        patchTrimmed, params.ku_K, params.ku_xyw);
    kToU(edgeTrimmed < 0.2) = 0;
    kToU(edgeTrimmed > 0.8) = 1;
    kToUconf(edgeTrimmed < 0.2) = 1;
    kToUconf(edgeTrimmed > 0.8) = 1;
    if(~suppressMessages) display('      Solving for alphas...');
    end
    alpha = solveForAlphas(Lap, trimap, params.lambda, params.
        usePCGtoSolve, kToU, kToUconf, params.ku_mult);
else
    if(~suppressMessages) display('      Solving for alphas...');
    end
    alpha = solveForAlphas(Lap, trimap, params.lambda, params.
        usePCGtoSolve);
end

alpha = reshape(alpha, [size(image, 1), size(image, 2)]);

if params.mattePostTrim %%%%% پس پردازش
    alpha(edgeTrimmed < 0.2) = 0;
    alpha(edgeTrimmed > 0.8) = 1;
end

dur = toc;
if(~suppressMessages) display(['Done. It took ' num2str(dur) '
    seconds.']); end
end

```

۲-آ تابع جریان اطلاعات ترکیب رنگ

برنامه آ-۲: کد MATLAB تابع جریان اطلاعات ترکیب رنگ

```
% این تابع جریان اطلاعات ترکیب رنگ در : Pollefeys Marc Aydin, Ozan Tunc Aksoy, Yagiz
% "Designing "Image Natural for Flow Information Inter-Pixel Effective" CVPR, Matting", ۲۰۱۷
% وقتی که پارامتر ورودی 'useXYinLLEcomp' باشد false (default)
% و تعریف شباهت استفاده شده در : "Manifold Tan, Ping Zhao, Qinqing Zou, Dongqing Chen, Xiaowu
% TOG, ACM propagation", edit preserving ۲۰۱۲
% وقتی که پارامتر ورودی 'useXYinLLEcomp' true باشد را پیاده سازی می کند.
% همه پارامترها غیر از Image اضافی هستند.
% خروجی ماتریس خلوتی است که برای همسایه های غیر محلی نقشه داده شده عناصر غیر صفر دارد.
% - K تعداد همسایه ها است
% - outMap نقشه سه گانه ای است که نزدیکترین همسایه ها تعریف شده اند.
% - xyWeight میزان اهمیت فضایی را در مختصات نزدیکترین همسایه های انتخابی تعیین می کند
function Wcm = colorMixtureAffinities(image, K, inMap, outMap, xyWeight
, useXYinLLEcomp)

[h, w, ~] = size(image);
N = h * w;

if ~exist('K', 'var') || isempty(K)
    K = 20;
end
if ~exist('inMap', 'var') || isempty(inMap)
    inMap = true(h, w);
end
if ~exist('outMap', 'var') || isempty(outMap)
    outMap = true(h, w);
end
if ~exist('xyWeight', 'var') || isempty(xyWeight)
    xyWeight = 1;
end
if ~exist('useXYinLLEcomp', 'var') || isempty(useXYinLLEcomp)
    useXYinLLEcomp = false;
end

[inInd, neighInd, features] = findNonlocalNeighbors(image, K,
xyWeight, inMap, outMap);
% inInd --> اندیس پیکسلهای ناحیه نامشخص
% neighInd --> اندیسهای ۲۰ همسایه در کل تصویر برای هر پیکسل در ناحیه نامشخص
% features --> ویژگی های [r,g,b,x,y] کل تصویر
if ~useXYinLLEcomp
    features = features(:, 1 : end - 2);
end
flows = zeros(size(inInd, 1), size(neighInd, 2));

for i = 1 : size(inInd, 1)
    flows(i, :) = localLinearEmbedding(features(inInd(i), :)',
features(neighInd(i, :), :)', 1e-10);
end
% sum(flows, 2) --> flows مجموع سطرهای ماتریس
flows = flows ./ repmat(sum(flows, 2), [1, K]);
```

```

inInd = repmat(inInd, [1, K]);
Wcm = sparse(inInd(:), neighInd(:), flows, N, N);
end

```

آ-۳ تابع جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص

برنامه آ-۳: کد MATLAB جریان اطلاعات از ناحیه مشخص به ناحیه نامشخص

```

% Effective "Designing Pollefeys, Marc Aydin, Ozan Tunc Aksoy, Yagiz
% Image Natural for Flow Information Inter-Pixel" CVPR, 2017. را پیاده سازی می کند.
% همه پارامترها غیر از image و trimap اضافی هستند.
% خروجی ها وزنهایی از مقادیر تخمین آلفا و مقادیر ضریب اعتماد آنها هستند.
% K- تعداد همسایه های مورد جستجو در نواحی پیش زمینه و پس زمینه که توسط وزنهایی LLE محاسبه شده اند است.
% xyWeight- میزان اهمیت مختصات فضایی را در همسایه های مورد جستجو را تعیین می کند.

function [alphaEst, conf] = knownToUnknownColorMixture(image, trimap, K, xyWeight)

    if ~exist('K', 'var') || isempty(K)
        K = 7;
    end
    if ~exist('xyWeight', 'var') || isempty(xyWeight)
        xyWeight = 10;
    end

    image = im2double(image);
    trimap = im2double(trimap(:,:,1));
    bg = trimap < 0.2;
    fg = trimap > 0.8;
    unk = ~(bg | fg);

    %%%%% همسایه ها را در پیش زمینه و پس زمینه پیدا می کند
    [inInd, bgInd, features] = findNonlocalNeighbors(image, K, xyWeight, unk, bg);
    [~, fgInd] = findNonlocalNeighbors(image, K, xyWeight, unk, fg);
    neighInd = [fgInd, bgInd];

    %%%%% وزنهایی LLE را محاسبه و رنگهای پیش زمینه و پس زمینه بدست آمده از ترکیب را تخمین می زند
    features = features(:, 1 : end - 2);
    flows = zeros(size(inInd, 1), size(neighInd, 2));
    fgCols = zeros(size(inInd, 1), 3);
    bgCols = zeros(size(inInd, 1), 3);
    for i = 1 : size(inInd, 1)
        flows(i, :) = localLinearEmbedding(features(inInd(i), :)', features(neighInd(i, :), :)', 1e-10);
        fgCols(i, :) = sum(features(neighInd(i, 1 : K), :) .* repmat(flows(i, 1 : K)', [1 3]), 1);
        %%%%% معادله ۳-۶
        bgCols(i, :) = sum(features(neighInd(i, K + 1 : end), :) .* repmat(flows(i, K + 1 : end)', [1 3]), 1);
        %%%%% معادله ۳-۶
    end
end

```

```

%%%%% آلفای تخمین زده شده از مجموع وزنه‌های همسایه های پیش زمینه
alphaEst = trimap;
alphaEst(unk) = sum(flows(:, 1 : K), 2); %%%%% معادله ۵-۳

%%%%% محاسبه ضریب اعتماد بر اساس تفاوت رنگ پیش زمینه و پس زمینه
unConf = fgCols - bgCols; %%%%% معادله ۷-۳
unConf = sum(unConf .* unConf, 2) / 3; %%%%% معادله ۷-۳
conf = double(fg | bg);
conf(unk) = unConf;
end

```

۴-آ تابع جریان اطلاعات درون ناحیه نامشخص

برنامه ۴-آ: کد MATLAB جریان اطلاعات درون ناحیه نامشخص

```

%Chi-Keung Li, Dingzeyu Chen, Qifeng Tang, "KNN Tang", TPAMI, IEEE Matting, ۲۰۱۳. پیاده سازی می کند.
% این تابع تفاوت‌های رنگ را بر اساس شباهت مقاله درهم تنیدگی تصویر
% همه پارامترها غیر از image اضافی هستند.
% خروجی ماتریسی خلوت است که عناصر غیر صفر برای همسایه های غیر محلی پیکسل‌های داده شده در نقشه دودویی
% inMap دارد.
% K-تعداد همسایه ها که از طریق LLE محاسبه می شوند، است.
% -outMap نقشه سه گانه ای است که نزدیکترین همسایه ها تعریف شده اند.
% -xyWeight میزان اهمیت مختصات فضایی را در انتخاب نزدیکترین همسایه انتخابی تعیین می کند.
% وقتی useHSV مقدار false (default) دارد جستجو در فضای [r,g,b,x,y] انجام می شود، در غیر اینصورت فضای
% ویژگی [cos(h),sin(h),s,v,x,y] است.

function Wcs = colorSimilarityAffinities(image, K, inMap, outMap,
xyWeight, useHSV)

[h, w, ~] = size(image);
N = h * w;

if ~exist('K', 'var') || isempty(K)
    K = 5;
end
if ~exist('inMap', 'var') || isempty(inMap)
    inMap = true(h, w);
end
if ~exist('outMap', 'var') || isempty(outMap)
    outMap = true(h, w);
end
if ~exist('xyWeight', 'var') || isempty(xyWeight)
    xyWeight = 0.05;
end
if ~exist('useHSV', 'var') || isempty(useHSV)
    useHSV = false;
end

if useHSV
    image = rgb2hsv(image);

```

```

        image = cat(3, cos(image(:, :, 1)) * 2 * pi, sin(image(:, :, 1))
            ) * 2 * pi, image(:, :, 2:3));
    end

    [~, neighInd, ~] = findNonlocalNeighbors(image, K, xyWeight, inMap,
        outMap);
    % خطوط زیر xy-weight را کاهش داده و مجموعه ای جدید از همسایه ها پیدا می کند که از پیاده سازی روش
    % KNN by Chen et al. برگرفته شده است.
    [inInd, neighInd2, features] = findNonlocalNeighbors(image, ceil(K
        / 5), xyWeight / 100, inMap, outMap);
    neighInd = [neighInd, neighInd2];
    features(:, end-1 : end) = features(:, end-1 : end) / 100;

    inInd = repmat(inInd, [1, size(neighInd, 2)]);
    flows = max(1 - sum(abs(features(inInd(:), :) - features(neighInd
        (:), :)), 2) / size(features, 2), 0);
    %%% معادله ۱۰-۲
    Wcs = sparse(inInd(:), neighInd(:), flows, N, N);
    Wcs = (Wcs + Wcs') / 2; %%% اگر p همسایه q باشد، q را همسایه p قرار می دهد.
end

```

۵-آ تابع جریان اطلاعات محلی

برنامه ۵-آ: کد MATLAB جریان اطلاعات محلی

```

% تابع زیر روش درهم تیدگی تصویر تشریح شده در "A Weiss, Yair Lischinski, Dani Levin, Anat TPAMI, IEEE Matting", Image Natural to Solution Form
% همه پارامترها غیر از image اضافی هستند.
% خروجی ماتریسی خلوت است که عناصر غیر صفر برای همسایه های غیر محلی پیکسلهای داده شده در نقشه دودویی
% inMap دارد.
% windowRadius - اندازه پنجره را برای توزیع های نرمال محلی تخمین زده شده را تعیین می کند.
% epsilon - ضریب منظم سازی را قبل از معکوس کردن ماتریس کوواریانس تعیین می کند (برای تصاویر نویزی باید
% بزرگتر باشد).
function W = mattingAffinity(image, inMap, windowRadius, epsilon)

    if ~exist('windowRadius', 'var') || isempty(windowRadius)
        windowRadius = 1;
    end
    if ~exist('epsilon', 'var') || isempty(epsilon)
        epsilon = 1e-7;
    end

    windowSize = 2 * windowRadius + 1;
    neighSize = windowSize^2;
    [h, w, c] = size(image);
    N = h * w;
    epsilon = epsilon / neighSize;

    %%% اگر نقشه سه گانه تعریف شده باشد نیازی به محاسبه شباهت در نواحی مشخص نیست.
    if nargin < 2 || isempty(inMap)

```

```

inMap = true(size(image, 1), size(image, 2));
end

[meanImage, covarMat] = localRGBnormalDistributions(image,
windowRadius, epsilon);

%%%%%% پیکسلها و همسایه های محلی آنها را تعیین می کند.
indices = reshape((1 : h * w), [h w]);
neighInd = im2col(indices, [windowSize windowSize], 'sliding');
inMap = inMap(windowRadius + 1 : end - windowRadius, windowRadius +
1 : end - windowRadius);
neighInd = neighInd(inMap, :); % همه ۹ همسایه محلی برای هر پیکسل در ناحیه نامشخص
inInd = neighInd(:, (neighSize + 1) / 2); % اندیس پیکسلها در ناحیه نامشخص
pixCnt = size(inInd, 1);

%%%%%% آماده سازی داده ورودی و خروجی
image = reshape(image, [N, c]); %%% ویژگی های تصویر
meanImage = reshape(meanImage, [N, c]); %%% ویژگی های تصویر میانگین گرفته
شده
flowRows = zeros(neighSize, neighSize, pixCnt);
flowCols = zeros(neighSize, neighSize, pixCnt);
flows = zeros(neighSize, neighSize, pixCnt);

%%%%%% محاسبه شباهت کدری
for i = 1 : size(inInd, 1)
neighs = neighInd(i, :);
shiftedWinColors = image(neighs, :) - repmat(meanImage(inInd(i)
, :), [size(neighs, 2), 1]);
%%%%%% معادله ۴-۱
flows(:, :, i) = shiftedWinColors * (covarMat(:, :, inInd(i)) \
shiftedWinColors');
%%%%%% معادله ۴-۱
neighs = repmat(neighs, [size(neighs, 2), 1]);
flowRows(:, :, i) = neighs;
flowCols(:, :, i) = neighs';
end
flows = (flows + 1) / neighSize; %%% معادله ۴-۱
W = sparse(flowRows(:), flowCols(:), flows(:), N, N);

%%%%%% اطمینان از اینکه W متقارن است.
W = W + W';

%%%%%% نرمال سازی
sumW = full(sum(W, 2));
sumW(sumW < 0.05) = 1;
W = spdiags(1 ./ sumW(:), 0, N, N) * W;
end

```

آ-۶ تابع پیش پرازش نقشه سه گانه ۱

برنامه آ-۶: کد MATLAB پیش پرازش نقشه سه گانه ۱


```

% Scott Price, Brian Rajan, Deepu Shahrian, Ehsan مقاله در تشریح سه گانه نقشه پیراش روش پیراش نقشه سه گانه تشریح شده در مقاله "Improving Cohen, Sampling Comprehensive using Matting Image" CVPR Sets ۲۰۱۳ را پیاده سازی می کند.
% این پیاده سازی از کد منبع عمومی که توسط نویسنده تهیه شده بعنوان یک راهنما استفاده می کند و ساختاری تکراری دارد که در مقاله توضیح داده نشده است.
% پارامترهای ورودی غیر از image و trimap اضافی هستند.
% maxManhaDist=sum(1:iterCnt) مشخص می شود که با IterCnt منتهن پیرایش با IterCnt مشخص می شود که چقدر آستانه رنگ کم شده است.
% -paramU ماکزیمم آستانه رنگ در تکرارها است.
% -paramD مشخص می کند که چقدر آستانه در هر تکرار کاهش یابد.

function trimap = trimmingFromKnownUnknownEdges(image, trimap, paramU,
paramD, iterCnt)

    if ~exist('iterCnt', 'var') || isempty(iterCnt)
        iterCnt = 9;
    end
    if ~exist('paramD', 'var') || isempty(paramD)
        paramD = 1 / 256; %%% چقدر آستانه رنگ شده است
    end
    if ~exist('paramU', 'var') || isempty(paramU)
        paramU = 9 / 256;
    end

    image = im2double(image);
    trimap = im2double(trimap);
    bg = (trimap < 0.2);
    fg = (trimap > 0.8);
    paramD = paramU - paramD;

    for i = 1 : iterCnt
        iterColorThresh = paramU - i * paramD / iterCnt; % color
        threshold = paramU - iterNo * (paramU - paramD) / maxIter
        trimap = LabelExpansion(image, trimap, i, iterColorThresh);
        %%% فاصله آستانه ۱ تا iterCnt
    end
end

function [extendedTrimap] = LabelExpansion(image, trimap, maxDist,
colorThresh)
    [h, w, ~] = size(image);

    fg = trimap > 0.8;
    bg = trimap < 0.2;
    knownReg = (bg | fg);
    extendedTrimap = trimap;

    searchReg = ((imdilate(fg, ones(2 * maxDist + 1)) & ~fg) | (imdilate
        (bg, ones(2 * maxDist + 1)) & ~bg));
    %%% لبه های نواحی مشخص
    [cols, rows] = meshgrid(1 : w, 1 : h);
    % cols --> 563 Rows Repetitious Contaning Numbers 1 : 800
    % rows --> 800 Columns Repetitious Contaning Numbers 1 : 563
    cols = cols(searchReg(:)); %%% ستونهایی که لبه ها در آنها وجود دارند
    rows = rows(searchReg(:)); %%% سطرهایی که لبه ها در آنها وجود دارند

```

```

winCenter = (2 * maxDist) / 2 + 1;
distPlane = repmat((1 : 2 * maxDist + 1)', [1, 2 * maxDist + 1]);
distPlane = sqrt((distPlane - winCenter) .^ 2 + (distPlane' -
winCenter) .^ 2);

for pixNo = 1 : size(cols, 1)
r = rows(pixNo);
c = cols(pixNo);
minR = max(r - maxDist, 1); %%%% محدود پیکسل
minC = max(c - maxDist, 1);
maxR = min(r + maxDist, h);
maxC = min(c + maxDist, w);
winMinR = winCenter - (r - minR); %%%% محدود پیکسل در پنجره
winMinC = winCenter - (c - minC);
winMaxR = winCenter + (maxR - r);
winMaxC = winCenter + (maxC - c);

pixColor = image(r, c, :);
imgWin = image(minR : maxR, minC : maxC, :); % رنگهای پیکسلها در پنجره
trimapWin = trimap(minR : maxR, minC : maxC);
%%%%% رنگهای پیکسلهای نقشه سه گانه در پنجره
%%%%% تفاوت رنگ بین پیکسل
winColorDiff = imgWin(:, :, 1) - pixColor(1);
winColorDiff(:, :, 2) = imgWin(:, :, 2) - pixColor(2);
winColorDiff(:, :, 3) = imgWin(:, :, 3) - pixColor(3);
winColorDiff = sqrt(sum(winColorDiff .* winColorDiff, 3));
%%%%% فاصله اقلیدوسی بین پیکسل و پنجره

candidates = (winColorDiff < colorThresh) & knownReg(minR : maxR
, minC : maxC);
%%%%% پیکسلهای شناخته شده زیر حد آستانه
if sum(candidates(:)) > 0
distWin = distPlane(winMinR : winMaxR, winMinC : winMaxC);
%%%%% فاصله سطح
distWin = distWin(candidates); %%%% فاصله های شناخته شده
[~, minDistInd] = min(distWin); %%%% اندیس کمترین فاصله
trimapWin = trimapWin(candidates);
extendedTrimap(r, c) = trimapWin(minDistInd);
end
end
end
end

```

۷-آ تابع پیش پرازش نقشه سه گانه ۲

برنامه ۷-آ: کد MATLAB پیش پرازش نقشه سه گانه ۲

% تابع زیر پیرایش نقشه سه گانه توضیح داده شده در Pollefeys, Marc Aydin, Ozan Tunc Aksoy, Yagiz
CVPR, Matting, Image Natural for Flow Information Inter-Pixel Effective Designing
۲۰۱۷. را پیاده سازی می کند.
% پارامترهای ورودی غیر از image و trimap اضافی هستند.

```

% minDist و maxDist میزان خوبی تطابق با پیش زمینه و پس زمینه را مشخص می کنند.
% windowRadius - اندازه پنجره را برای توزیع های نرمال محلی تخمین زده شده را تعیین می کند.
% K - تعداد نزدیکترین همسایه های پیدا شده برای استفاده بردارها قبل از مقایسه فاصله باتاچاریا است.
function trimap = patchBasedTrimming(image, trimap, minDist, maxDist,
    windowRadius, K)

    if ~exist('minDist', 'var') || isempty(minDist)
        minDist = 0.25;
    end
    if ~exist('maxDist', 'var') || isempty(maxDist)
        maxDist = 0.90;
    end
    if ~exist('windowRadius', 'var') || isempty(windowRadius)
        windowRadius = 1;
    end
    if ~exist('K', 'var') || isempty(K)
        K = 10;
    end

    image = im2double(image);
    trimap = im2double(trimap(:,:,1));
    [h, w, ~] = size(image);

    epsilon = 1e-8;

    fg = trimap > 0.8;
    bg = trimap < 0.2;
    unk = ~(fg | bg);

    [meanImage, covarMat] = localRGBnormalDistributions(image,
        windowRadius, epsilon);

    [unkInd, fgNeigh] = findNonlocalNeighbors(meanImage, K, -1, unk, fg
    );
    %%% K تا نزدیکترین همسایه از پیش زمینه برای هر پیکسل در ناحیه نامشخص پیدا می کند.
    [~, bgNeigh] = findNonlocalNeighbors(meanImage, K, -1, unk, bg);
    %%% K تا نزدیکترین همسایه از پس زمینه برای هر پیکسل در ناحیه نامشخص پیدا می کند.

    meanImage = reshape(meanImage, [h * w, size(meanImage, 3)]);
    %%% ویژگی های تصویر میانگین شده

    fgBhatt = zeros(K, 1);
    bgBhatt = zeros(K, 1);
    for i = 1 : size(unkInd, 1)
        pixMean = meanImage(unkInd(i), :)' ;
        pixCovar = covarMat(:, :, unkInd(i));
        pixDet = det(pixCovar);
        for n = 1 : K
            nMean = meanImage(fgNeigh(i, n), :)' - pixMean;
            nCovar = covarMat(:, :, fgNeigh(i, n));
            nDet = det(nCovar);
            nCovar = (pixCovar + nCovar) / 2;
            fgBhatt(n) = 0.125 * nMean' * (nCovar \ nMean) + 0.5 * log(
                det(nCovar) / sqrt(pixDet * nDet));
            %%% فاصله باتاچاریا
        end
    end
end

```

```

end
for n = 1 : K
    nMean = meanImage(bgNeigh(i, n), :) - pixMean;
    nCovar = covarMat(:, :, bgNeigh(i, n));
    nDet = det(nCovar);
    nCovar = (pixCovar + nCovar) / 2;
    bgBhatt(n) = 0.125 * nMean' * (nCovar \ nMean) + 0.5 * log(
        det(nCovar) / sqrt(pixDet * nDet));
    %%% فاصله باتاچاریا
end
minFGdist = min(fgBhatt);
minBGdist = min(bgBhatt);
%%%% معادله ۹-۳
if minFGdist < minDist
    if minBGdist > maxDist
        trimap(unkInd(i)) = 1;
    end
elseif minBGdist < minDist
    if minFGdist > maxDist
        trimap(unkInd(i)) = 0;
    end
end
end
end
end

```

آ-۸ تابع تعیین میزان شفافیت تصویر

برنامه آ-۸: کد MATLAB تعیین میزان شفافیت تصویر

% Pollefeys, Marc Aydin, Ozan Tunc Aksoy, Yagiz در "Designing "CVPR, Matting", Image Natural for Flow Information Inter-Pixel Effective" ۲۰۱۷. را پیاده سازی می کند.

% عملکرد آن بر اساس یک طبقه بندی ساده هیستوگرام است.

```

function ht = detectHighlyTransparent(image, trimap)

    image = reshape(im2double(image), [size(image, 1) * size(image, 2),
        3]);
    %%% ویژگی های تصویر
    trimap = im2double(trimap(:, :, 1));

    fg = trimap > 0.8; %%% ناحیه پیش زمینه
    bg = trimap < 0.2; %%% ناحیه پس زمینه
    unk = ~(fg | bg); %%% ناحیه نامشخص
    fg = fg & imdilate(unk, ones(20)); %%% لبه های ناحیه پیش زمینه
    bg = bg & imdilate(unk, ones(20)); %%% لبه های ناحیه پس زمینه

    fgi = image(fg, :); % مقادیر لبه های تصویر در پیش زمینه
    bgi = image(bg, :); % مقادیر لبه های تصویر در پس زمینه
    uni = image(unk, :); % مقادیر پیکسل های ناحیه نامشخص

```

```

fgh = [imhist(fgi(:, 1), 10); imhist(fgi(:, 2), 10); imhist(fgi(:, 3), 10);] / sum(fg(:));
%%%%% احتمال ۱۰ سطح در هر رنگ RGB از پیش زمینه
bgh = [imhist(bgi(:, 1), 10); imhist(bgi(:, 2), 10); imhist(bgi(:, 3), 10);] / sum(bg(:));
%%%%% احتمال ۱۰ سطح در هر رنگ RGB از پس زمینه
unh = [imhist(uni(:, 1), 10); imhist(uni(:, 2), 10); imhist(uni(:, 3), 10);] / sum(unk(:));
%%%%% احتمال ۱۰ سطح در هر رنگ RGB در ناحیه نامشخص
%%%%%  $Ax = b \rightarrow x = A \setminus b$ 
weights = ([fgh bgh]' * [fgh bgh]) \ ([fgh bgh]' * unh);
%%%%% بدست آوردن a و b در معادله ۲۰-۳
recError = [fgh bgh] * weights - unh; %%%% معادله ۲۰-۳
recError = sqrt(sum(recError(:) .* recError(:))) / size(recError(:), 1);
%%%%% معادله ۲۰-۳

ht = recError > 0.0099;

end

```

۹-آ تابع پیدا کردن همسایه ها

برنامه ۹-آ: کد MATLAB پیدا کردن همسایه ها

```

% تابع زیر همسایه ها را با استفاده از رنگ و مختصات فضایی پیکسلها پیدا می کند.
% -K تعداد همسایه های مورد جستجو است.
% پارامترهای غیر از K و image اضافی هستند.
% -xyWeight میزان اهمیت مختصات فضایی است.
% inMap و outMap نقشه هایی دودویی که ناحیه جستجو را تعیین می کنند، هستند.
% اگر eraseSelfMatches مقدار true داشته باشد، مطابقت با خود شناسایی و حذف می شوند.
% -inInd اندیس پیکسلهای مورد جستجو و neighInd اندیس همسایه های پیدا شده را بر می گردانند.
% -features ویژگی های تصویر است.
function [inInd, neighInd, features] = findNonlocalNeighbors(image, K, xyWeight, inMap, outMap, eraseSelfMatches)

[h, w, c] = size(image);

if ~exist('xyWeight', 'var') || isempty(xyWeight)
    xyWeight = 1;
end
if ~exist('inMap', 'var') || isempty(inMap)
    inMap = true(h, w);
end
if ~exist('outMap', 'var') || isempty(outMap)
    outMap = true(h, w);
end
if ~exist('eraseSelfMatches', 'var') || isempty(eraseSelfMatches)
    eraseSelfMatches = true;
end
end

```

```

features = reshape(image, [h*w, c]);
if xyWeight > 0
    [x, y] = meshgrid(1 : w, 1 : h);
    % x --> 563 Rows Repetitious Of Values = 1 : 800
    % y --> 800 columns Repetitious Of Vlaues = 1 : 563
    x = xyWeight * double(x) / w;
    y = xyWeight * double(y) / h;
    features = [features x(:) y(:)]; % [r,g,b,x,y] For All Pixels
end

inMap = inMap(:);
outMap = outMap(:);
indices = (1 : h * w)';
inInd = indices(inMap);
%%%%% اندیس پیکسلهای ناحیه نامشخص در نقشه سه گانه
outInd = indices(outMap);

if eraseSelfMatches
    %%%%% همسایه برای خود تطبیقی پیدا می کند.
    neighbors = knnsearch(features(outMap, :), features(inMap, :),
        'K', K + 1);
    %%%%% همسایه برای هر پیکسل ناحیه نامشخص پیدا می کند
    %%%%% خود تطبیقی را حذف می کند
    validNeighMap = true(size(neighbors));
    validNeighMap(inMap(inInd) & outMap(inInd), 1) = 0;
    validNeighMap(:, end) = ~validNeighMap(:, 1);
    validNeighbors = zeros(size(neighbors, 1), size(neighbors, 2) -
        1);
    for i = 1 : size(validNeighbors, 1)
        validNeighbors(i, :) = neighbors(i, validNeighMap(i, :));
    end
    neighInd = outInd(validNeighbors);
else
    neighbors = knnsearch(features(outMap, :), features(inMap, :),
        'K', K);
    neighInd = outInd(neighbors);
end
end
end

```

آ-۱۰ تابع پیدا کردن وزن ها

برنامه آ-۱۰: کد MATLAB پیدا کردن وزن ها

```
% "Nonlinear Saul. K. Lawrence Roweis, T. Sam
%   Science. Embedding", Linear Local by Reduction Dimensionality
%   را پیاده سازی می کند.
% -w وزنهایی برای نمایش بردار سطری است.
% -pt یک پیکسل داده شده است.
% -neighbors همسایه های پیکسل pt است.
% -conditionerMult ضریب مقید سازی برای اضافه کردن به ماتریس همبستگی همسایه قبل از معکوس کردن است.

function w = localLinearEmbedding(pt, neighbors, conditionerMult)
    %%% pt یک بردار سطری است
    % کد زیر مطابق با مقاله Reduction Dimensionality Nonlinear قسمت یادداشت و مراجع (۷) نوشته شده
    % است.
    corr = neighbors' * neighbors + conditionerMult * eye(size(
        neighbors, 2));
    ptDotN = neighbors' * pt; %%% پیکسل با استفاده از همسایه هایش بازسازی می شود
    %%% Ax=b --> x=A\b
    %%% A Inverse * b <---> A\b
    alpha = 1 - sum(corr \ ptDotN);
    beta = sum(corr \ ones(size(corr, 1), 1));
    %%% مجموع عناصر معکوس ماتریس همبستگی
    lagrangeMult = alpha / beta;
    w = corr \ (ptDotN + lagrangeMult);
end
```

آ-۱۱ تابع حل سیستم خطی برای پیدا کردن آلفا

برنامه آ-۱۱: کد MATLAB حل سیستم خطی برای پیدا کردن آلفا ها

```
% تابع زیر سیستم خطی را حل می کند.

function alphas = solveForAlphas(Lap, trimap, lambda, usePCG, alphaHat,
    conf, aHatMult)
    if ~exist('usePCG', 'var') || isempty(usePCG)
        usePCG = true;
    end
    [h, w, ~] = size(trimap);
    N = h * w;
    known = trimap > 0.8 | trimap < 0.2;
    A = lambda * spdiags(double(known(:)), 0, N, N);
    %%% A = lambda * E_tau
    if exist('alphaHat', 'var')
        if ~exist('conf', 'var') || isempty(conf)
            conf = ones(size(alphaHat));
        end
    end
    if ~exist('aHatMult', 'var') || isempty(aHatMult)
        aHatMult = 0.1;
    end
end
```

```

    conf(known(:)) = 0;
    A = A + aHatMult * spdiags(conf(:), 0, N, N);
    %%%%% همان ماتریس H در معادله ۱۶-۳ است
    b = A * alphaHat(:); %%%%% سمت راست معادله ۱۶-۲
else
    b = A * double(trimap(:) > 0.8);
end
A = A + Lap; %%%%% سمت چپ معادله ۱۶-۳
if usePCG %%%%% استفاده از روش گرادیان مزدوج با استفاده از تابع pcg متلب
    [alphas, ~] = pcg(A, b, [], 2000); %%%%% حل معادله ۱۶-۳
else
    alphas = A \ b;
end
alphas(alphas < 0) = 0;
alphas(alphas > 1) = 1;
end

```


پیوست ب

کدهای مربوط به پیاده سازی طبقه بندی با استفاده از روش درهم تنیدگی جریان اطلاعات بین پیکسلی در متلب

ب-۱ تابع اصلی ، فراخوانی و اجرای روش ها برای طبقه بندی

برنامه ب-۱: کد MATLAB تابع اصلی فراخوانی و اجرای روش ها برای طبقه بندی

```
کد زیر روش های طبقه بندی و مجموعه داده ها را برای طبقه بندی مقدار دهی می کند.  
همچنین جریانهای مختلف اطلاعات را برای طبقه بندی مشخص می سازد  
% Mahmood Amintoosi, HSU 2018  
% Modified by: Farzad Zandi  
close all  
clear  
clc  
  
masir = 'doc/';  
datasets = {'twomoon','tworing','fourcluster','fisheriris','ionosphere'  
, 'sonar'};  
nDataset = numel(datasets);  
  
flows = {'all','noCM','noLoc','noIU','noKtoU','onlyIU'};  
methods = {'SVM','KNN'}; % Methods Numbers  
nMethods = numel(methods);  
for i=1:numel(flows)  
    methods{nMethods+1} = ['InfFlow-',flows{i}];  
    nMethods = nMethods+1;  
end  
% nMethods = numel(methods);  
sampleNumbers = [20 40 60]; %%% تعداد نمونه ها را تعیین می کند  
nSN = numel(sampleNumbers); % Number of sample numbers' list  
  
Results = [];  
for i = 1 : nDataset  
    datasetFileName = datasets{i};  
    [X,y] = loadDataset(datasetFileName); %%% مجموعه داده را فراخوانی می کند
```

```

problem.X = X;
problem.y = y;
numC = numel(unique(y)); % Number of classes
N = numel(y);
for snIdx=1 : nSN
    k = round(N/sampleNumbers(snIdx));
    CVO = cvpartition(y,'k',k);
    Errors = zeros(CVO.NumTestSets,1);
    RunTimes = zeros(CVO.NumTestSets,1);
    NSamples = zeros(CVO.NumTestSets,1);
    iTr = 0;
    for iCV = 1:CVO.NumTestSets
        % MAT % CV training set is also our training set
        trainIdx = CVO.training(iCV);
        testIdx= CVO.test(iCV);
        numCTr = numel(unique(y(trainIdx))); % Number of classes in
            training set
        if numC ~= numCTr
            disp('Partitions are discarded...')
            pause(0.1)
            continue
        end
        iTr = iTr+1;
        problem.testIdx = testIdx;
        problem.trainIdx = trainIdx;
        for methodNo = 1 : nMethods
            ravesh = methods{methodNo};
            fprintf(' Dataset : %s, Number of train samples : %d,
                Method : %s\n',datasetFileName,...
                    sampleNumbers(snIdx),ravesh)
            [err, runTime] = dasteband(problem,ravesh) ;
            %%%% روش مورد نظر را بر روی داده فراخوانی شده اجرا می کند %%%%
            Errors(iTr,methodNo) = err;
            RunTimes(iTr,methodNo) = runTime;
        end
        NSamples(iTr) = sum(trainIdx);
    end
    if iTr == 0
        error('There is not any training set containing all
            categories')
    end
    %%%% نتایج را در متغیر Result ذخیره می کند %%%%
    for methodNo = 1 : nMethods
        Results(i,snIdx).kCV = k;
        Results(i,snIdx).NRuns = iTr;
        Results(i,snIdx).NSamples = round(mean(NSamples));
        Results(i,snIdx).method(methodNo).rt = mean(RunTimes(:,
            methodNo));
        Results(i,snIdx).method(methodNo).er = mean(Errors(:,
            methodNo));
        Results(i,snIdx).method(methodNo).acc = 1-Results(i,snIdx).
            method(methodNo).er;
        Results(i,snIdx).method(methodNo).methodName = methods{
            methodNo};
    end
end
end

```

```
end
```

```
resFileName = sprintf('Results');  
save(resFileName, 'Results', 'methods', 'sampleNumbers');  
writeLatexResults(Results, methods); %%%% کند می تولید می کند
```

ب-۲ تابع پیاده سازی جداول نتایج طبقه بندی در لاتک

برنامه ب-۲: کد MATLAB تابع پیاده سازی جداول نتایج طبقه بندی در لاتک

تابع زیر نتایجی که از طبقه بندی بدست آمده را به صورت فایل‌هایی متنی برای فراخوانی در لاتک ایجاد می‌کند.

% Mahmood Amintoosi, HSU 2018

% Modified by: Farzad Zandi

```
function writeLatexResults(Results, methods)  
mahaks = {'ACC', 'ERR', 'RT'};  
mahakNames = {'Accuracy', 'Error', 'RunTime'};  
mahakOptFunc = {'max', 'min', 'min'};  
mahakNonOptFunc = {'min', 'max', 'max'};  
  
N = size(Results, 1);  
nSN = size(Results, 2);  
nMethods = numel(methods);  
  
for snIdx=1:nSN  
    NSamples = Results(1, snIdx).NSamples;  
    kCV = Results(1, snIdx).kCV;  
  
    fVal = zeros(N, nMethods);  
    ERR = fVal;  
    RT = fVal; % RunTime  
    ACC = fVal;  
    for ii=1:N  
        for j=1:nMethods, eval(['RT(ii, j)=Results(ii, snIdx).method(j).  
            rt;']); end  
        for j=1:nMethods, eval(['ACC(ii, j)=Results(ii, snIdx).method(j).  
            acc;']); end  
        for j=1:nMethods, eval(['ERR(ii, j)= Results(ii, snIdx).method(j)  
            .er;']); end  
    end  
  
    % ravesh = 'KNN';  
    for mm = 1:numel(mahaks)  
        mahak = mahaks{mm};  
        mahakName = mahakNames{mm};  
  
        masir = 'doc\';  
  
        data = eval(mahak);  
  
        meanData = mean(data);  
        optData = eval(['mahakOptFunc{mm} '(meanData)']);
```

```

nonOptData = eval([mahakNonOptFunc{mm} '(meanData)']);

tableFileName = sprintf('%s_k%02d_Table.txt',mahak,kCV);
fprintf('\n\\input{%s}\n',tableFileName)
tableFileName = sprintf('%s%s',masir,tableFileName);

fid = fopen(tableFileName,'wt');
fprintf(fid,'\\begin{table}[t]\n');
fprintf(fid,'\\centering\n');
fprintf(fid,'\\scalebox{0.72}{\n');
fprintf(fid,'\\begin{tabular}{|c|}');
for m=1:nMethods
    fprintf(fid,'c|');
end
fprintf(fid,')\n\\hline\n');
fprintf(fid,' & \\multicolumn{%d}',nMethods);
fprintf(fid,' >{\\columncolor{gray}c|}{\\color{white}\\textbf
{MethodName}}\\\\');

fprintf(fid,'No. ');
for m=1:nMethods
    fprintf(fid,'%s ',methods{m});
end
fprintf(fid,'\\\\ \n');
fprintf(fid,'\\hhline{|=}');
for m=1:nMethods
    fprintf(fid,'=');
end
fprintf(fid,'|}\n');

% [~,bestIdxInRow] = max(data ');
[bestVal,bestIdxInRow] = eval([mahakOptFunc{mm} '(data')']);

for dbNo = 1:N
    fprintf(fid,'%d ',dbNo);
    for m=1:nMethods
        if data(dbNo,m)==bestVal(dbNo)
            fprintf(fid,'%& \\multicolumn{1}{>{\\cellcolor[gray
] {.8}c|}{\\mathbf{%.3f}}$ ',data(dbNo,m));
        else
            fprintf(fid,'%& \\t %.3f$ ',data(dbNo,m));
        end
    end
    fprintf(fid,'\\\\ \n');
end
fprintf(fid,'\\hhline{|=}');
for m=1:nMethods
    fprintf(fid,'=');
end
fprintf(fid,'|}\n');

fprintf(fid,'Avg');
for m=1:nMethods
    if meanData(m)==optData
        fprintf(fid,'%& \\multicolumn{1}{>{\\cellcolor[gray
] {.8}c|}{\\mathbf{%.3f}}$ ',meanData(m));
    end
end

```

```

                %           elseif meanData(m)==nonOptData
                %           fprintf(fid,'%s \\multicolumn{1}{>{\\
                %           cellcolor{black}}c/}&{\\color{white}$\\mathbf{\\%5.3f}
                %           $} ',meanData(m));
        else
            fprintf(fid,'%& \\t \\%5.3f$ ',meanData(m));
        end
    end
    fprintf(fid,'\\\\\\\\ \\n');

    fprintf(fid,'\\\\hline\\n');
    fprintf(fid,'\\\\end{tabular}} \\n');
    fprintf(fid,'\\\\caption{%s:k%d}\\n',mahakName,kCV);%
        native2unicode(mahakName,'UTF-8');
    fprintf(fid,'\\\\label{tab:%s:%d}\\n\\\\end{table}\\n',mahak,NSamples
    );
    fclose(fid);
end
end
end

```

واژه‌نامه فارسی به انگلیسی

Threshold	آستانه گیری
Cross validation	اعتبار سنجی متقابل
Ill-posed	بد طرح
Ill-conditioned	بد وضع
Overfitting	بیش برازش
Background	پس زمینه
Foreground	پیش زمینه
Labeled pixels	پیکسل های برچسب گذاری شده
Unlabeled pixels	پیکسل های برچسب گذاری نشده
Neighboring pixels	پیکسل های مجاور
Prediction function	تابع پیش بینی
Objective function	تابع هدف
Cost function	تابع هزینه
Convert	تبدیل کردن
Extracting	تخمین
Transpose	ترانهاد
Compositing	ترکیب
Approximation	تقریب
Propagation	تعمیم
Constant	ثابت
Substituting	جای گذاری
Inner producte	حاصل ضرب داخلی
Eliminated	حذف کردن

Scribble	نشان گذاری
Coefficients	ضرایب
Training data	داده های آموزشی
Image Matting	در هم تنیدگی تصویر
Gray level	سطح خاکستری
Pseudo-inverse	شبه معکوس
Intensities	شدت روشنایی
Assumption	فرض
Image segmentation	قطعه بندی تصویر
Least squares errors	کم ترین مربعات خطا
Covariance	کواریانس
Minimizing	کمینه سازی
Diagonal matrix	ماتریس قطری
Training set	مجموعه آموزشی
Test set	مجموعه آموزشی
Data set	مجموعه داده ها
Unknown	مجهول
Evaluation	محاسبه
Linear model	مدل خطی
Local	محلی
Compar	مقایسه
Independent	مستقل
Normal equation	معادلات نرمال
Criterion	معیار
Regularisation	منظم سازی
Invertible	معکوس پذیر
Non smooth	ناهموار
Region	ناحیه
Consequence	نتیجه
Feature mapping	نگاشت طرح

Trimap	نقشه سه گانه
Representation	نمایش
Semi-supervised	نیمه نظارتی
Goal	هدف
Overlap	هم پوشانی
Machine learning	یادگیری ماشین
Variance	واریانس

واژه‌نامه انگلیسی به فارسی

Approximation	تقریب
Assumption	فرض
Background	پس زمینه
Coefficients	ضرایب
Compar	مقایسه
Compositing	ترکیب
Constant	ثابت
Convert	تبدیل کردن
Cost function	تابع هزینه
Consequence	نتیجه
Covariance	کواریانس
Criterion	معیار
Cross validation	اعتبار سنجی متقابل
Data set	مجموعه داده ها
Diagonal matrix	ماتریس قطری
Eliminated	حذف کردن
Evaluation	محاسبه
Extracting	تخمین
Foreground	پیش زمینه
Goal	هدف
Gray level	سطح خاکستری
Ill-conditioned	بد وضع
Ill-posed	بد طرح

Image Matting	در هم تنیدگی تصویر
Image segmentation	قطعه بندی تصویر
Independent	مستقل
Inner product	حاصل ضرب داخلی
Intensities	شدت روشنایی
Interpolates	درونیابی
Invertible	معکوس پذیر
Labeled pixels	پیکسل های برچسب گذاری شده
Least squares errors	کم ترین مربعات خطا
Linear model	مدل خطی
Local	محلی
Machine learning	یادگیری ماشین
Minimizing	کمینه سازی
Neighboring pixels	پیکسل های مجاور
Non smooth	ناهموار
Normal equation	معادلات نرمال
Objective function	تابع هدف
Overfitting	بیش برآزش
Overlap	هم پوشانی
Partition	افراز
Prediction function	تابع پیش بینی
Propagation	تعمیم
Pseudo-inverse	شبه معکوس
Region	ناحیه
Regularisation	منظم سازی
Representation	نمایش
Semi-supervised	نیمه نظارتی
Scribble	نشان گذاری
Substituting	جای گذاری
Threshold	آستانه گیری

Training data	داده های آموزشی
Training set	مجموعه آموزشی
Test set	مجموعه تست
Transpose	ترانهاده
Trimap	نقشه سه گانه
Unknown	مجهول
Unlabeled pixels	پیکسل های برچسب گذاری نشده
Variance	واریانس

Hakim Sabzevari University

An Outline of MSc. Thesis



Surname : Zandi

Name : Farzad

Student No. : 9513137048

Supervisor : Dr. Mahmood Amintoosi

Advisor : Dr. Alireza Ghodsi

Faculty of Mathematics and Computer Science

Program: Computer Science Field : Decision Science and Knowledge Engineering

Title of thesis : Designing Effective Inter-Pixel Information Flow for Natural Image Matting

Keywords : Image matting , Information flow , Known and Unknown regions , K Nearest Neighborhood

Abstract : The exporting foregrounds from images have important role in images and videos. The separating foreground from background is an important semi-supervised learning classification problem and is related to image processing. semi-supervised learning means a problem be modeling with user's input finites. One of the relatively new ways in this domain , the so-called Image Matting , which has focused on several studies. In this research , introduce several image matting methods. One of the newest image matting methods is Designing Effective Inter-Pixel Information Flow for Natural Image Matting that is major purpose of this thesis. In addition to with define similarities of pixels information flows will describe and with combine them this method construct.

Many image matting methods have weaknesses , so with regularization equations parameters of above method it can be use at a post-processing step for improve the weaknesses.

At last Inter-pixel information flow method used for classification several non image datasets with different features and train samples in a new job and it's results compare with Support Vector Machine (SVM) and K-Nearest Neighborhood (KNN) methods and obtain to better results from others in some datasets .



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 Computer Science**

Designing Effective Inter-Pixel Information Flow for Natural Image Matting

Supervisor
Dr. Mahmood Amintoosi

Advisor
Dr. Alireza Ghodsi

By
Farzad Zandi

July 2018