

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ
الْحَمْدُ لِلَّهِ الَّذِي
خَلَقَ السَّمَوَاتِ وَالْأَرْضَ
وَالَّذِي يُضَوِّبُ الْمَوْتَى
إِنَّ رَبَّهُ لَسَدِيدٌ
إِلَىٰ عَرْشِهِ الرَّحِيمُ
الَّذِي يُرْسِلُ الرِّيَّاحَ
تُضَوِّبُ السَّحَابَ الْمَوْبِقَ
الَّذِي يُرْسِلُ السَّمَكَ
الْمَوْبِقَ وَهُوَ الْعَلِيمُ
الْحَكِيمُ الَّذِي يُرْسِلُ
الْمَوْتَىٰ بِرُوحِهِ
الْحَمْدُ لِلَّهِ الَّذِي
خَلَقَ السَّمَوَاتِ وَالْأَرْضَ
وَالَّذِي يُضَوِّبُ الْمَوْتَى
إِنَّ رَبَّهُ لَسَدِيدٌ
إِلَىٰ عَرْشِهِ الرَّحِيمُ
الَّذِي يُرْسِلُ الرِّيَّاحَ
تُضَوِّبُ السَّحَابَ الْمَوْبِقَ
الَّذِي يُرْسِلُ السَّمَكَ
الْمَوْبِقَ وَهُوَ الْعَلِيمُ
الْحَكِيمُ الَّذِي يُرْسِلُ
الْمَوْتَىٰ بِرُوحِهِ



دانشکده‌ی مهندسی برق و رباتیک

گروه الکترونیک

رفع اعوجاج های فتومتریک و هندسی در تصاویر به کمک اطلاعات روشنایی پس زمینه و اطلاعات هندسی خطوط متن

محمد امین طلوع بیدختی

استاد راهنما:

دکتر علیرضا احمدی فرد

استاد مشاور:

دکتر حسین خسروی

پایان نامه جهت اخذ درجه‌ی کارشناسی ارشد در رشته مهندسی برق، الکترونیک

زمستان ۱۳۹۳

تعهد نامه

اینجانب **محمدامین طلوع بیدختی** دانشجوی دوره کارشناسی ارشد رشته **مهندسی برق - الکترونیک** دانشکده **مهندسی برق و رباتیک** دانشگاه صنعتی شاهرود نویسنده پایان نامه **رفع اعوجاج های فتومتریک و هندسی در تصاویر به کمک اطلاعات روشنایی پس زمینه و اطلاعات هندسی خطوط متن** تحت راهنمایی **دکتر علیرضا احمدی** فرد متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه های رایانه ای، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

تقدیر و تشکر

پس از حمد و ثنای بی‌حد بر آستان صفات بی‌همتای احدیت که در کمال رأفت و درنهایت عطوفت رخصت اتمام این پایان‌نامه را به بنده عطا فرموده است، از اساتید گرانقدر، دکتر علی‌رضا احمدی‌فرد و دکتر حسین خسروی که بارانمایی‌ها و کمک‌های ارزشمند خود، من را در انجام این پایان‌نامه یاری رساندند؛ تشکر می‌کنم.

تقدیم به

و بعد از مدت‌ها، پس از پیمودن راه‌های فراوان با لطف و عنایت خدا همراه با حضور شیرین اساتید عزیزم در کنارم و باراهنمایی‌ها و صبوری‌های فراوانشان، نگاه‌های پدر مادرم، با چشم‌های پر از برق شوق، و زیبایی حضورشان در کنارم، که خستگی‌های این راه را به امید و روشنی راه تبدیل کرد و امیدوارم بتوانم در آینده‌ای نزدیک جوابگوی این همه محبت آن‌ها باشم، با احترام فراوان برای این همه تلاش این عزیزان برای موفقیت من، این پایان‌نامه را به پدر و مادرم و اساتید عزیزم تقدیم می‌کنم. باشد که گوشه‌ای از محبت‌شان جبران شود.

چکیده

تصاویر سند تهیه شده توسط اسکنر یا دوربین دیجیتال، همواره با دو نوع اعوجاج فتومتریک و اعوجاج هندسی همراه هستند. هر دو نوع اعوجاج، باعث کاهش دقت عملکرد نرم افزارهای OCR می شوند. در این پایان نامه سعی بر این است که با ارائه ی روشی نوین و همچنین بهبود روش های گذشته، به رفع اعوجاج اسناد به منظور بهبود عملکرد نرم افزارهای OCR دست یابیم.

در این پایان نامه به منظور تصحیح اعوجاج فتومتریک از روش درون نگاری هارمونیک استفاده شده است. به منظور افزایش دقت این الگوریتم پیشنهاد نموده ایم تا از همسایه های قطری هر پیکسل نیز استفاده شود؛ این عمل موجب بهبود تخمین روشنایی پس زمینه می گردد. همچنین برای کاهش پیچیدگی محاسباتی پیشنهاد نمودیم تا در اجرای الگوریتم درون نگاری یکبار از همسایه های چهارتایی اصلی و بار دیگر از همسایه های قطری استفاده نماییم. با این روش پیچیدگی محاسباتی روش پیشنهادی نیز افزایش نمی یابد. علاوه بر این به منظور بهبود کیفیت در تخمین پس زمینه از فیلترهای ملایم کننده ی گوسی و میانگین گیر نیز استفاده شده است.

همچنین به منظور تصحیح اعوجاج هندسی پس از استخراج خطوط از تصویر، هر خط به چندین نوار عمودی تقسیم می شود. در هر نوار مختصات نقطه ای که مقدار افکنش افقی آن بیشینه است؛ ذخیره می شود. انحنای هر خط به کمک این مجموعه نقاط توسط تابع درجه ی سوم مشخص می گردد. در نهایت با استفاده از تخمین تبدیل پرسپکتیو، اعوجاج هندسی هر خط بر طرف می گردد. در ضمن به منظور تشخیص انحنای خطوط با طول کم از انحنای خطوط با طول بزرگ تر خط قبلی یا بعدی استفاده می شود.

روش های پیشنهادی بر روی پایگاه های داده ی فارسی و انگلیسی پیاده سازی شده و با روش های دیگر مقایسه گردیده است. نتایج بیانگر قدرت و دقت روش پیشنهادی در رفع این دو نوع اعوجاج است.

کلمات کلیدی: درون نگاری، اعوجاج فتومتریک، اعوجاج هندسی، پردازش دوبعدی اسناد، نویسه خوان نوری

مقالات استخراج شده از پایان نامه

- ۱- م. طلوع بیدختی، ع. احمدی فرد، "رفع اعوجاج فتومتریک از تصویر اسناد به کمک درون نگاری بهبودیافته"، ششمین کنفرانس فناوری اطلاعات و دانش، دانشگاه صنعتی شاهرود، شاهرود، ۱۳۹۲.

فهرست

فصل ۱- مقدمه.....	۱
۱-۱- سند چیست ؟	۲
۲-۱- نویسه خوان نوری	۲
۳-۱- تعریف مسئله.....	۳
۴-۱- اهداف پایان نامه.....	۴
۵-۱- ساختار پایان نامه	۵
فصل ۲- بررسی روش های موجود.....	۷
۱-۲- مقدمه	۸
۲-۲- اعوجاج فتومتریک	۸
۱-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر شکل هیستوگرام	۹
۲-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر خوشه بندی.....	۱۰
۳-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر آنتروپی	۱۱
۴-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر تشابه نسبی	۱۲
۵-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر اطلاعات فضایی	۱۲
۶-۲-۲- روش های حذف اعوجاج فتومتریک مبتنی بر اطلاعات محلی	۱۳
۳-۲-۲- اعوجاج هندسی	۱۵
۱-۳-۲- تکنیک های مبتنی بر بازسازی سه بعدی تصاویر	۱۵
۲-۳-۲- تکنیک های مبتنی بر شناسایی خطوط متن	۱۹
فصل ۳- تئوری.....	۲۷
۱-۳- مقدمه	۲۷
۱-۲-۳- فرسایش	۲۷

۲۸ گسترش ۲-۲-۳
۲۹ بستن ۳-۲-۳
۲۹ درون نگاری ۳-۳
۳۰ فیلتر گذاری مکانی ۴-۳
۳۰ فیلترهای مکانی ملایم کننده ۱-۴-۳
۳۰ فیلتر ملایم کننده ی میانگین گیر ۱-۱-۴-۳
۳۱ فیلتر ملایم کننده ی گوسی ۲-۱-۴-۳
۳۲ فیلترهای تیز کننده ۲-۴-۳
۳۲ لاپلاسین ۱-۲-۴-۳
۳۳ گرادیان ۲-۲-۴-۳
۳۴ روش های مبتنی بر افکنش ۵-۳
۳۶ تخمین تبدیل پرسپکتیو ۶-۳
۴۱	فصل ۴-روش پیشنهادی.....
۴۱ مقدمه ۱-۴
۴۱ معرفی پایگاه های داده ی مورد استفاده ۲-۴
۴۱ پایگاه داده ی خسروی راد ۱-۲-۴
۴۱ پایگاه داده ی شامقلی ۲-۲-۴
۴۴ پایگاه داده ی DFKI ۳-۲-۴
۴۵ اعوجاج فتومتریک ۳-۴
۴۵ ساختن ماسک درون نگاری ۱-۳-۴
۴۶ حذف پیش زمینه ۲-۳-۴
۴۷ نرم کردن سایه ی استخراج شده ۳-۳-۴
۴۸ بازتاب تصویر استخراج شده ۴-۳-۴

۴-۴-۴-۴	روش پیشنهادی به منظور رفع اعوجاج فتومتریک	۴۹
۴-۴-۴-۱	ساختن ماسک درون نگاری	۴۹
۴-۴-۲	حذف پیش زمینه با استفاده از درون نگاری هارمونیک بهبود یافته	۵۰
۴-۴-۳	عملیات ملایم کردن / تغییر متناوب ماسک	۵۲
۴-۴-۴	بازتاب تصویر استخراج شده	۵۴
۴-۵-۵	اعوجاج هندسی	۵۳
۴-۵-۱	تشخیص خطوط متن	۵۴
۴-۵-۲	تشخیص نقشه‌ی خمیدگی سند	۶۱
۴-۵-۲-۱	تقسیم هر خط به نوارهای عمودی	۶۱
۴-۵-۲-۲	افکنش افقی با تابع گرادیان و یافتن مختصات بیشترین مقدار	۶۲
۴-۵-۲-۳	گذراندن منحنی درجه‌ی سه از نقاط مشخص شده	۶۸
۴-۵-۳	تعیین نقشه‌ی انتقال نقاط	۷۰
۴-۵-۳-۱	قرار دادن هر خط بین دو منحنی	۷۰
۴-۵-۳-۲	تقسیم ناحیه‌ی بین دو منحنی به تعدادی چهارضلعی	۷۱
۴-۵-۳-۳	انتقال نقاط از چهارضلعی به مستطیل	۷۱
۴-۵-۴	همپوشانی نقاط انتقال نیافته	۷۲
۷۷	فصل ۵- بررسی نتایج روش پیشنهادی	
۵-۱-۱	مقدمه	۷۵
۵-۲-۲	نتایج روش پیشنهادی به منظور رفع اعوجاج فتومتریک	۷۵
۵-۲-۱	بررسی نتایج روش‌های نیبلک و اتسو	۸۱
۵-۲-۲	بررسی نتایج روش ژانگ و روش پیشنهادی	۸۲
۵-۳-۳	نتایج روش پیشنهادی به منظور رفع اعوجاج هندسی	۸۶
۵-۳-۱	نتایج تجربی روش پیشنهادی بر روی تصاویر	۸۶

۵-۳-۲-مقایسه‌ی نتایج بر روی اسناد انگلیسی ۸۸

۵-۳-۳-مقایسه‌ی نتایج بر روی اسناد فارسی ۹۴

فصل ۶- نتیجه‌گیری و پیشنهادات ۱۰۱

۱-۶-مقدمه ۱۰۱

۲-۶-نتیجه‌گیری ۱۰۱

۳-۶-کارهای آینده ۱۰۳

مراجع ۱۰۸

فهرست شکل‌ها

- شکل ۱-۱: اسناد الکترونیکی مشکل‌دار ۳
- شکل ۱-۲: نتیجه‌ی اعمال روش [۴] الف) تصویر ورودی ب) نتیجه‌ی روش [۴] ج) هیستوگرام تصویر ورودی و مقدار آستانه ۱۰
- شکل ۲-۲: استفاده از روش درون‌نگاری به‌منظور بازسازی تصاویر آسیب‌دیده الف) تصویر ورودی ب) نتیجه‌ی روش [۲۵] ۱۴
- شکل ۳-۲: رفع اعوجاج فتومتریک به کمک روش [۲۹] الف) تصویر ورودی دارای اعوجاج فتومتریک ب) استخراج لایه پیش‌زمینه ج) تصویر خروجی بدون اعوجاج ۱۵
- شکل ۴-۲: نتیجه‌ی حاصل از روش اولگس [۳۳] الف) تصویر دارای اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج ۱۷
- شکل ۵-۲: تشکیل سطح استوانه‌ای تصویر سند [۳۴] ۱۷
- شکل ۶-۲: اسناد پیچش دار [۳۵] ۱۸
- شکل ۷-۲: مدل سازی بازتاب داخلی سطوح کتاب [۳۸] ۱۹
- شکل ۸-۲: نتایج حاصل از روش ژانگ [۲۹] ۱۹
- شکل ۹-۲: نتیجه‌ی اعمال روش لاوایلی [۳۹] الف) تصویر ورودی دارای اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی ۲۰
- شکل ۱۰-۲: نتیجه‌ی روش ماسالوویچ [۴۰] الف) تخمین بدفرمی تصویر ب) تصویر بازسازی شده ۲۱
- شکل ۱۱-۲: نتیجه‌ی روش اولگس [۴۱] الف) تصویر دارای اعوجاج هندسی ب) نتیجه‌ی روش اولگس ۲۱
- شکل ۱۲-۲: نتیجه‌ی روش کاو [۴۲] الف) تصویر ورودی با اعوجاج هندسی ب) نتیجه‌ی روش اولگس ۲۲
- شکل ۱۳-۲: نتیجه‌ی روش ده بوید [۴۴] الف) تصویر ورودی با اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی ۲۴
- شکل ۱۴-۲: نتیجه‌ی روش گل پرداز [۴۵] الف) تصویر ورودی با اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی ۲۴
- شکل ۱۵-۲: نتیجه‌ی روش خسروی راد [۴۶] الف) تصویر ورودی با اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی ۲۵

شکل ۲-۱۶: نتیجه‌ی روش شامقلی [۴۷] الف) تصویر ورودی با اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی ۲۶

شکل ۳-۱: الف) تصویر ورودی ب) نتیجه‌ی فرسایش با جزء ساختاری مربعی 3×3 ج) نتیجه‌ی گسترش با جزء ساختاری ستونی به طول ۳ پیکسل د) جزء ساختاری مربعی 3×3 ه) جزء ساختاری ستونی به طول ۳ پیکسل ۲۹

شکل ۳-۲: فیلتر ملایم‌کننده‌ی میانگین‌گیر الف) فیلتر ملایم‌کننده‌ی میانگین‌گیر 3×3 ب) فیلتر ملایم‌کننده‌ی میانگین‌گیر سطری به طول ۵ پیکسل ۳۲

شکل ۳-۳: نقاب‌های فیلتر ملایم‌کننده‌ی گوسی الف) فیلتر ملایم‌کننده‌ی گوسی 3×3 با واریانس ۰,۵ ب) فیلتر ملایم‌کننده‌ی گوسی سطری به طول ۵ پیکسل با واریانس ۰,۵ ۳۳

شکل ۳-۴: نقاب‌های فیلتر لاپلاسین الف) نقاب فیلتری رابطه‌ی $3-14$ ب) نقاب فیلتری با المان‌های قطری که نسبت به چرخش حساس نیست ۳۴

شکل ۳-۵: عملگرهای سو بل الف) عملگر سو بل در راستای X ب) عملگر سو بل در راستای Y ۳۵

شکل ۳-۶: نتیجه‌ی افکنش افقی مبتنی بر شدت روشنایی [۵۳] الف) تصویر ورودی ب) نمودار افکنش افقی ج) خطوط یافته شده بر اساس نمودار افکنش افقی (تصاویر برای فهم بهتر 90° درجه چرخانده شده‌اند) ۳۶

شکل ۳-۷: انتقال نقاط از چهارضلعی به مستطیل و بالعکس الف) چهارضلعی ب) مستطیل ۳۷

شکل ۴-۱: نمونه تصاویر مجموعه‌ی داده‌ی خسرویراد ۴۳

شکل ۴-۲: نمونه تصاویر مجموعه‌ی داده‌ی شامقلی ۴۴

شکل ۴-۳: دو نمونه از تصاویر پایگاه داده‌ی DFKI ۴۵

شکل ۴-۴: فلوجارت حذف اعوجاج فتومتریک در روش [۲۹] ۴۶

شکل ۴-۵: نتیجه‌ی اعمال روش ژانگ [۲۹] به منظور تولید ماسک الف) تصویر ورودی ب) ماسک تولیدی تصویر ۲-۴ الف ۴۷

شکل ۴-۶: نتیجه‌ی اعمال روش ژانگ [۲۹] به منظور حذف پیشزمینه و نرم کردن سایه‌ی استخراج‌شده شکل ۴-۵ الف) نتیجه‌ی حذف پیشزمینه به روش درون‌نگاری ب) نتیجه‌ی نرم کردن سایه به روش RBF ۴۸

شکل ۴-۷: نتیجه‌ی نهایی روش ژانگ [۲۹] الف) تصویر ورودی ب) تصویر نهایی بدون اعوجاج فتومتریک ۴۹

شکل ۴-۸: فلوجارت روش پیشنهادی ۵۰

شکل ۴-۹: نتیجه‌ی تولید ماسک درون‌نگاری الف) تصویر اولیه با اعوجاج فتومتریک ب) ماسک روش درون‌نگاری ۵۱

شکل ۴-۱۰: نتیجه‌ی روش درون‌نگاری هارمونیک بهبودیافته الف) بعد از اولین تکرار ب) بعد از ۳۰ تکرار ۵۲

شکل ۴-۱۱: روش پیشنهادی جهت کاهش حلقه‌های تودرتو الف) تصویر ورودی ب) پر کردن سطر و ستون اضافه‌شده با پیکسل‌های همسایه ج) شیف‌ت به بالا د) شیف‌ت به راست ۵۳

شکل ۴-۱۲: نتیجه‌ی نهایی روش پیشنهادی الف) به کمک ملایم‌کننده‌ی میانگین‌گیر 3×3 ب) به کمک ملایم‌کننده‌ی میانگین‌گیر سطری به طول ۲۰ پیکسل ۵۴

شکل ۴-۱۳: فلوجارت روش پیشنهادی رفع اعوجاج هندسی ۵۵

شکل ۴-۱۴: ماسک تولیدی توسط عملگر گسترش و مشکلات رایج آن الف) تصویر ورودی ب) ماسک با جزء ساختاری نامناسب ج) ماسک با جزء ساختاری مناسب ۵۷

شکل ۴-۱۵: حذف نقاط از تصویر الف) تصویر ورودی ب) به اشتباه برخی از نقاط یا قسمتی از حروف حذف‌شده‌اند ج) تصویر اصلاح شده با روش پیشنهادی ۵۹

شکل ۴-۱۶: ماسک تولیدی تصویر ۴-۱۵ الف) ۶۰

شکل ۴-۱۷: نمونه‌ای دیگر از نتیجه‌ی تولید ماسک الف) تصویر ورودی ب) ماسک تولیدی ۶۰

شکل ۴-۱۸: آینه کردن افقی و تولید ماسک الف) آینه‌ی افقی تصویر ۴-۱۵ الف) ب) ماسک تولیدی ۶۱

شکل ۴-۱۹: تقسیم خطوط به نوارهای افقی الف) تقسیم خط چهارم شکل ۴-۹ الف) به نوارهای افقی ب) تقسیم خط دوم شکل ۴-۱۰ الف) به نوارهای افقی ۶۲

شکل ۴-۲۰: چرخش نوار بین زاویه‌های -16 تا 16 درجه الف) چرخش با زاویه 16 - درجه ب) چرخش با زاویه 12 - درجه پ) چرخش با زاویه 8 - درجه ت) چرخش با زاویه 4 - درجه ج) چرخش با زاویه 0 درجه چ) چرخش با زاویه 4 درجه ح) چرخش با زاویه 8 درجه خ) چرخش با زاویه 16 درجه ... ۶۳

شکل ۴-۲۱: نمودار افکنش افقی الف) زاویه‌ی چرخش 16 - ب) زاویه‌ی چرخش 8 - ۶۴

شکل ۴-۲۲: نمودار افکنش افقی الف) زاویه‌ی چرخش 0 ب) زاویه‌ی چرخش 16 ۶۵

شکل ۴-۲۳: نمودار افکنش افقی Gy الف): گردایان تصویر ب) نمودار افکنش افقی ۶۶

شکل ۴-۲۴: تغییر اندازه‌ی تصویر بر اثر چرخش الف) چرخش به سمت راست ($\theta < 0$) ب) چرخش به سمت چپ ($\theta > 0$) ۶۸

شکل ۴-۲۵: پیدا کردن نقاط جهت تشخیص انحنای هر خط الف) نتیجه‌ی پیدا کردن نقاط جهت تشخیص انحنای ب) نتیجه‌ی روش پیشنهادی جهت یافتن نقطه‌ی پایانی ۶۹

شکل ۴-۲۶: چند نمونه از تخمین انحنای خطوط الف و ب) تشخیص انحنای خطوط انگلیسی ج) تشخیص انحنای خط فارسی (از آیینیهی خط فارسی استفاده شده است). ۷۰.....

شکل ۴-۲۷: نقاط رئوس چهارضلعی ها الف) متن انگلیسی ب) متن فارسی ۷۲.....

شکل ۴-۲۸: نتیجهی تصویر رفع اعوجاج هندسی شده الف) تصویر بدون اعوجاج هندسی ب) نمایش نقاط انتقال نیافته قسمتی از شکل ۴-۲۸ الف ۷۳.....

شکل ۴-۲۹: همپوشانی نقاط انتقال نیافته الف) تصویر خروجی همپوشانی شده ب) همپوشانی نقاط انتقال نیافته قسمتی از تصویر ۴-۲۹ الف ۷۴.....

شکل ۵-۱: نمونه تصاویر از پایگاههای دادهی مختلف الف و ب) پایگاه دادهی شامقلی ج و د) پایگاه دادهی خسروی راد ۷۷.....

شکل ۵-۲: نتایج روشهای مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ الف، الف) روش اتسو ب) روش نیبلک ج) روش ژانگ د) روش پیشنهادی ۷۸.....

شکل ۵-۳: نتایج روشهای مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ ب، الف) روش اتسو ب) روش نیبلک ج) روش ژانگ د) روش پیشنهادی ۷۹.....

شکل ۵-۴: نتایج روشهای مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ ج، الف) روش اتسو ب) روش نیبلک ج) روش ژانگ د) روش پیشنهادی ۸۰.....

شکل ۵-۵: نتایج روشهای مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ د، الف) روش اتسو ب) روش نیبلک ج) روش ژانگ د) روش پیشنهادی ۸۱.....

شکل ۵-۶: هیستوگرام تصویر ۵-۱ ب به همراه مقدار آستانه محاسبه شده به روش اتسو ۸۲.....

شکل ۵-۷: دو نمونه از مشکلات روش نیبلک الف) حروف شکسته قسمتی از تصویر ۵-۴ ب) نوفه فلفلی قسمتی از تصویر ۵-۲ ب ۸۳.....

شکل ۵-۸: نتایج مرحلهی درون نگاری دو تصویر ۵-۱ الف و ۵-۱ ب، الف و ج) روش پیشنهادی ب و د) روش ژانگ ۸۶.....

شکل ۵-۹: نمونه تصاویر دارای اعوجاج هندسی الف، ب، پ) پایگاه دادهی DFKI ت، ج، د) پایگاه دادهی خسرویراد ۸۸.....

شکل ۵-۱۰: نتیجهی اعمال روشهای مختلف بر روی تصویر ۵-۹ پ الف) روش SEG ب) روش SKEL ج) روش شامقلی د) روش پیشنهادی ۹۰.....

شکل ۵-۱۱: نتیجهی اعمال روشهای مختلف بر روی تصویر ۵-۹ پ الف) روش SEG ب) روش SKEL ج) روش شامقلی د) روش پیشنهادی ۹۲.....

شکل ۵-۱۲: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ پ الف) روش SEG ب) روش SKEL ج) روش شامقلی د) روش پیشنهادی ۹۳

شکل ۵-۱۳: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ ت الف) روش خسروی‌راد ب) روش شامقلی ج) روش پیشنهادی ۹۷

شکل ۵-۱۴: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ ج الف) روش خسروی‌راد ب) روش شامقلی ج) روش پیشنهادی ۹۸

شکل ۵-۱۵: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ د الف) روش خسروی‌راد ب) روش شامقلی ج) روش پیشنهادی ۹۹

فهرست جدول‌ها

- جدول ۱-۲: جدول روش‌های رایج رفع اعوجاج فتومتریک از تصویر ۸
- جدول ۱-۵: مشخصات تصاویر نمونه ۷۷
- جدول ۲-۵: مدت زمان اجرای روش پیشنهادی و روش ژانگ ۸۵
- جدول ۳-۵: مدت زمان اجرای روش پیشنهادی و روش ژانگ بدون کاهش حلقه‌های تودرتو ۸۵
- جدول ۴-۵: بررسی درصد دقت صحت کلمات اسناد انگلیسی ۹۵
- جدول ۵-۵: مدت زمان اجرای روش شامقلی و روش پیشنهادی ۹۶
- جدول ۶-۵: بررسی درصد دقت صحت کلمات اسناد فارسی ۹۶

فصل اول

مقدمه

۱-۱- سند چیست؟

این واژه کاربردی وسیع و همگانی یافته و همه‌ی افراد جامعه به‌طور روزمره با نوعی از اسناد سروکار دارند. از این‌رو، برداشتی از مفهوم سند در ذهنشان شکل گرفته است، علاوه بر کاربرد عام سند در گفتگوهای روزانه این واژه در متون تخصصی و در ادبیات اداری در معانی متعددی به‌کاررفته است.

در حالت کلی منظور از سند کلیه‌ی ارتباطات نوشتاری مانند گزارش‌های فنی، فایل‌های اداری، کتاب‌ها، روزنامه، مقالات، مجلات، نامه‌ها، چک‌های بانکی و هر چیزی که حاوی اطلاعات بوده و روشی را برای انتقال اطلاعات فراهم می‌کند، است.

امروزه با پیشرفت علم و فناوری و استفاده از دستگاه‌های دیجیتال، به‌جای اسناد چاپی از اسناد الکترونیکی استفاده می‌شود. در واقع سند الکترونیکی تصویری، از سندهای چاپی است که این تصویر به کمک دوربین-های تصویربرداری یا اسکنرها ایجاد شده است. اسناد الکترونیکی بسیاری از مشکلات سندهای معمولی از جمله فرسودگی، نیاز به فضای زیاد و هزینه‌ی زیاد را ندارند؛ به همین منظور استفاده از این اسناد روزبه‌روز در حال افزایش است.

۱-۲- نویسه‌خوان نوری

نویسه‌خوان نوری^۲ یا بازشناسی نوری نویسه‌ها^۳ که به‌صورت مختصر OCR نامیده می‌شود، ابتدا تنها در مورد بازشناسی ارقام و حروف چاپی به کار گرفته می‌شد. باگذشت زمان و پیشرفت قابل توجه در این زمینه، روش‌های بازشناسی دست‌نوشته و متون چاپی مطرح شدند که دامنه‌ی کاربرد را به کلمه و عبارات رساندند. با توجه به عدم تطبیق OCR با این موارد این نام برای این روش‌ها نیز استفاده شد و رواج پیدا کرد. امروزه از OCR بیشتر برای بازشناسی مستندات چاپی، مثل صفحات کتاب، مجله‌ها و نامه‌ها استفاده می‌کنند.

یک سامانه‌ی OCR مثل یک ماشین‌نویس متن را خوانده و آن را به قالب مناسب برای ذخیره در رایانه تبدیل می‌کند. این سامانه اشیاء موجود در تصویر را که ارقام، حروف و کلمات هستند؛ بازشناسی کرده و در قالب مناسب ذخیره می‌کند.

¹ Document

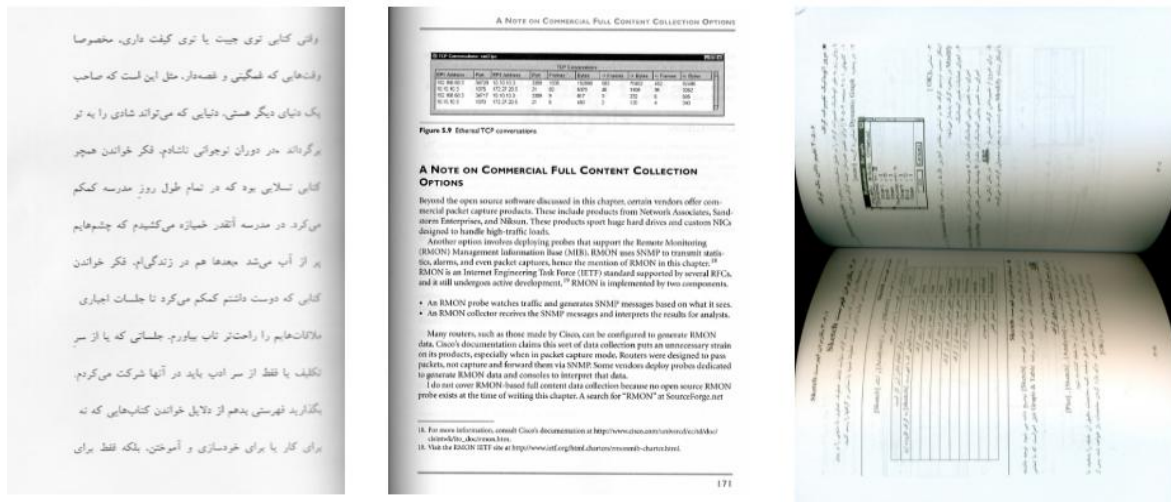
² Optical Character Reader (OCR)

³ Optical Character Recognition (OCR)

۳-۱- تعریف مسئله

معمولاً تصاویر ورودی که به عنوان داده‌ی اولیه برای سیستم OCR مورداستفاده قرار می‌گیرند، توسط اسکنرها و دوربین‌های دیجیتال تهیه می‌شوند. علاوه بر آن، با پیشرفت علم و فناوری و بیشتر در دسترس قرار گرفتن دوربین‌های دیجیتال، استفاده از آن‌ها به منظور عکس برداری افزایش یافته و با توجه به مشکلات عکس برداری نیاز به بازسازی تصاویر بیش از پیش احساس می‌شود.

در شکل ۱-۱، سه نمونه از اسناد الکترونیکی که نیاز به بازسازی دارند را مشاهده می‌کنید. بیشتر این مشکلات توسط چشم انسان قابل چشم‌پوشی بوده ولی برای پردازش‌های رایانه‌ای از جمله تشخیص^۱ و تحلیل^۲ و بازشناسی^۳ مشکل ایجاد می‌کنند.



شکل ۱-۱: اسناد الکترونیکی مشکل‌دار

به همین منظور بیان مشکلات تصویربرداری اسناد به منظور بررسی و رفع آن‌ها الزامی است. مهم‌ترین علت‌های این امر را می‌توان به صورت زیر بیان نمود:

- ۱- استفاده‌ی نادرست و دقت پایین برخی از دوربین‌های دیجیتال و اسکنرها
- ۲- شرایط محیطی نامطلوب مثل نورپردازی نامناسب
- ۳- کیفیت پایین اسناد چاپی برای مثال اسناد قدیمی، اسناد تاخوردیده یا اسناد مچاله شده

¹ Understanding

² Analysis

³ Recognition

۴- چرخش یا کجی سند در طی فرایند اسکن گرفتن یا عکس برداری

۵- انحنای ناشی از کتابهای قطور

علت‌های ذکرشده به صورت کلی باعث ایجاد دو نوع مشکل در تصویربرداری می‌شوند:

۱- اعوجاج فتومتریکی^۱

۲- اعوجاج هندسی^۲

به سایه‌ی ایجادشده در امتداد تصویر که با توجه به نورپردازی و قطر کتاب و نوع اسکنر یا دوربین دیجیتال متفاوت است در اصطلاح اعوجاج فتومتریکی گفته می‌شود و به خمیدگی ایجادشده در سطح کتاب یا کجی سند که هنگام اسکن کردن یا عکس برداری ایجاد می‌شود و علت اصلی آن انحنای برگه‌های سند ناشی از قطر کتاب یا کج قرار گرفتن سند در زمان اسکن یا عکس برداری است، اعوجاج هندسی گفته می‌شود. این دو نوع اعوجاج باعث می‌شوند تا نرم‌افزارهای رایج^۳ قادر به تشخیص صحیح کلمات نباشند، به همین منظور در این پایان‌نامه سعی بر این است که این دو نوع اعوجاج تا حد امکان برطرف شود.

۱-۴- اهداف پایان‌نامه

هدف از انجام این پایان‌نامه ارائه‌ی روشی مناسب به منظور رفع دو نوع اعوجاج فتومتریکی و اعوجاج هندسی از اسناد است. در این روش با استفاده از اطلاعات روشنایی پس‌زمینه و اطلاعات هندسی خطوط به رفع این دو نوع اعوجاج می‌پردازیم.

با توجه به این نکته که سایه در لایه‌ی پس‌زمینه‌ی^۴ تصویر قرار دارد، ابتدا سعی می‌شود لایه‌ی پس‌زمینه از لایه‌ی پیش‌زمینه^۵ جداشده و با اعمال روش درون‌نگاری، سایه از لایه‌ی پس‌زمینه حذف شود. در نهایت با ترکیب لایه‌ها تصویر نهایی بدون اعوجاج فتومتریکی تولید می‌شود. همچنین به منظور حذف اعوجاج هندسی از تصاویر از اطلاعات هندسی خطوط استفاده کرده و اعوجاج هر خط بر اساس میزان انحنای همان خط برطرف می‌گردد.

¹ Photometric distortion

² Geometric distortion

³ Omni Page, Read Iris, پرشبانگار

⁴ BackGround

⁵ Forground

۱-۵- ساختار پایان نامه

در فصل دوم به معرفی و بررسی روش‌های موجود و کارهای گذشتگان پیرامون رفع اعوجاج فتومتریک و اعوجاج هندسی پرداخته می‌شود. همچنین روش‌های موجود دسته‌بندی شده و شرح مختصری از هر کدام از این روش‌ها بیان می‌شوند.

در فصل سوم روش‌های تئوری استفاده شده در پایان نامه با ذکر جزئیات بیان می‌شوند.

در فصل چهارم پس از معرفی پایگاه‌های داده‌ی مورد استفاده به معرفی روش پیشنهادی به منظور رفع اعوجاج فتومتریک از تصاویر می‌پردازیم و در بخش آخر روش پیشنهادی به منظور رفع اعوجاج هندسی معرفی می‌شود.

فصل پنجم به بررسی نتایج روش پیشنهادی می‌پردازد. در بخش اول پس از مقداردهی به پارامترهای روش پیشنهادی، نتایج تجربی به منظور رفع اعوجاج فتومتریک نشان داده می‌شوند و این نتایج با روش‌های موجود رفع اعوجاج فتومتریک مقایسه می‌شوند. در بخش آخر این فصل نیز به بررسی نتایج روش‌های پیشنهادی به منظور رفع اعوجاج هندسی پرداخته شده و نتایج با روش‌های موجود به صورت کمی و کیفی مقایسه می‌شوند. برای مقایسه‌ی نتایج به صورت کمی از درصد دقت صحت به عنوان معیار سنجش استفاده شده است.

و در فصل ششم به نتیجه‌گیری و بیان یافته‌های پایان نامه می‌پردازیم. همچنین پیشنهادهایی به منظور افزایش کارایی روش‌های پیشنهادی ارائه می‌شوند که در آینده انجام داده خواهند شد.

فصل دوم

بررسی روش‌های موجود

۱-۲- مقدمه

در این فصل ابتدا به معرفی روش‌های موجود پیرامون رفع اعوجاج فتومتریک می‌پردازیم و در بخش سوم روش‌های موجود پیرامون رفع اعوجاج هندسی را بیان می‌کنیم.

۲-۲- اعوجاج فتومتریک

در دو دهه‌ی گذشته روش‌های زیادی به‌منظور رفع اعوجاج فتومتریک بیان شده است. همان‌طور که در جدول ۱-۲ مشاهده می‌کنید شش دسته کلی از روش‌های رایج برای حذف اعوجاج فتومتریک معرفی شده است. در ادامه به معرفی روش‌های ذکر شده پرداخته می‌شود.

جدول ۱-۲: جدول روش‌های رایج رفع اعوجاج فتومتریک از تصویر

نام	دسته
روش‌های حذف اعوجاج فتومتریک مبتنی بر شکل هیستوگرام	یک
روش‌های حذف اعوجاج فتومتریک مبتنی بر خوشه‌بندی	دو
روش‌های حذف اعوجاج فتومتریک مبتنی بر آنتروپی	سه
روش‌های حذف اعوجاج فتومتریک مبتنی بر تشابه نسبی	چهار
روش‌های حذف اعوجاج فتومتریک مبتنی بر اطلاعات فضایی	پنج
روش‌های حذف اعوجاج فتومتریک مبتنی بر اطلاعات محلی	شش

۲-۲-۱- روش‌های حذف اعوجاج فتومتریکی مبتنی بر شکل هیستوگرام

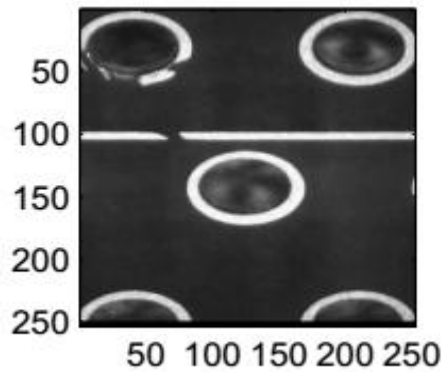
این دسته از روش‌ها آستانه گذاری را بر مبنای شکل هیستوگرام انجام می‌دهند. به‌طور کلی در این رویکرد اطلاعات مربوط به قله‌ها، دره‌ها و انحنای هیستوگرام نرم شده، مورد تحلیل قرار گرفته و با انتخاب یک آستانه‌ی مناسب اعوجاج فتومتریکی رفع می‌گردد.

روزن فیلد و دلاتوره در [۱] از تحلیل تقعر هیستوگرام به‌عنوان رویکردی در انتخاب آستانه بهره گرفتند و عملکرد آن را بر روی مجموعه‌ای از هیستوگرام‌های تصاویر مادون قرمز مورد بررسی قرار دادند که البته این روش از دقت خوبی برخوردار نبود.

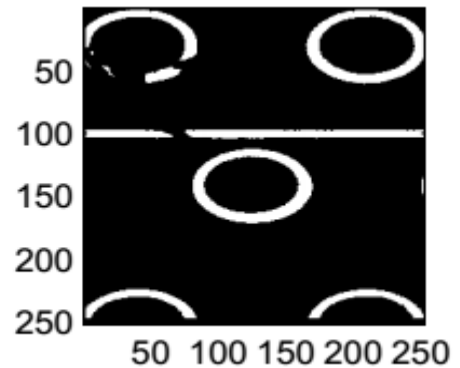
سیزن در [۲] به کمک تحلیل قله با کانالو کردن تابع هیستوگرام با یک کرنل نرم و مشتق گیر مقدار آستانه محاسبه می‌شود. در این روش به‌وسیله تنظیم کردن دهانه نرم‌کننده کرنل و ترکیب قله‌ها، هیستوگرام به یک تابع دوسطحی تبدیل شده و در نهایت اعوجاج فتومتریکی تا حدودی برطرف می‌شود. رامش و همکارانش در [۳] از دو روش برای تخمین هیستوگرام به یک تابع دوسطحی و انتخاب مقدار آستانه استفاده کرده‌اند. در روش اول مقدار آستانه بر اساس کمینه کردن مجموع مربعات خطا میان تابع دوسطحی تخمین زده شده و هیستوگرام، محاسبه شده و در روش دوم از کمینه کردن مقدار واریانس تابع دوسطحی تخمین زده شده، مقدار آستانه به دست می‌آید.

کمکا و کوپر در [۴] با تعمیم ایده‌ی تخمین شکل و با استفاده از مراحل شامل کوانتیزه سازی، روشی برای محاسبه مقدار آستانه مناسب معرفی کردند.

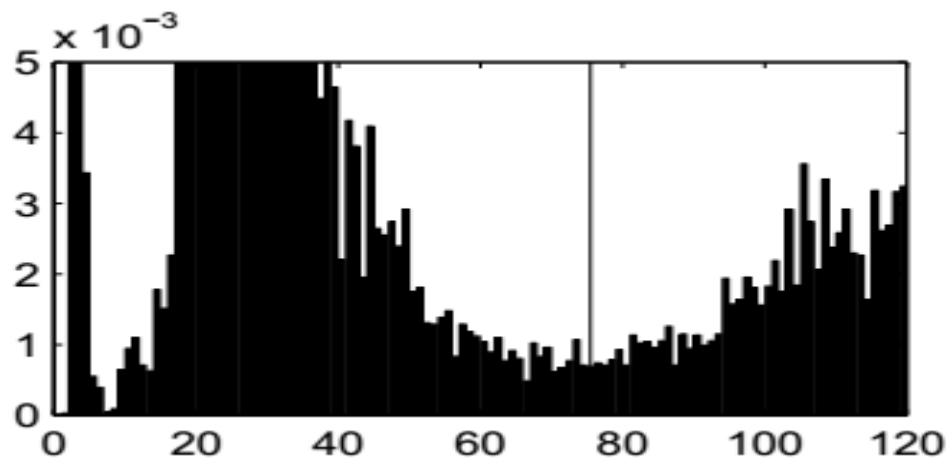
در شکل ۱-۲ نتیجه‌ی حاصل از اعمال روش [۴] را مشاهده می‌کنید. در شکل ۱-۲ الف یک تصویر ورودی سطح خاکستری نشان داده شده است. همچنین هیستوگرام این تصویر و مقدار آستانه‌ی محاسبه شده توسط این روش در شکل ۱-۲ ج رسم شده است. تصویر دوسطحی شده نهایی بر اساس مقدار آستانه‌ی به‌دست آمده، در شکل ۱-۲ ب قابل مشاهده است.



الف



ب



ج

شکل ۲-۱: نتیجه‌ی اعمال روش [۴] الف) تصویر ورودی ب) نتیجه‌ی روش [۴] ج) هیستوگرام تصویر ورودی و مقدار آستانه

۲-۲-۲- روش‌های حذف اعوجاج فتومتریک مبتنی بر خوشه‌بندی

در این دسته از روش‌ها سطوح خاکستری موجود در تصاویر در دو گروه پیش‌زمینه و پس‌زمینه خوشه‌بندی شده و یا به عبارتی به صورت تلفیقی از دو توزیع گاوسی مدل می‌شوند. ریدلر و کلورد در [۵] یکی از اولین روش‌های مبتنی بر مدل‌های گاوسی دو کلاسه را بهبود دادند. با تکرار مراحل، این الگوریتم یک آستانه جدید به دست می‌آید و تا زمان همگرایی این مراحل تکرار می‌شوند؛ و در نهایت مقدار آستانه با میانگین گرفتن از متوسط پیکسل‌های پیش‌زمینه و پس‌زمینه محاسبه می‌شود.

اتسودر [۶]، با استفاده از مینیمم کردن مجموع وزن گذاری شده‌ی واریانس‌های بین کلاسی پیکسل‌های پس‌زمینه و پیش‌زمینه مقدار آستانه‌ی بهینه را محاسبه کرده است. این روش هنگامی که تعداد پیکسل‌ها در هر کلاس نزدیک به هم باشد، نتایج رضایت بخشی می‌دهد. روش اتسو همچنان یکی از پرکاربردترین روش‌های مورد استفاده در آستانه گذاری است.

لوید در [۷] با در نظر گرفتن توابع چگالی گوسی هم واریانس و با یک جستجوی مکرر، خطا در طبقه‌بندی را کمینه می‌کند. پس از لوید، کیتلر و الینگورس در [۸]، فرض واریانس‌های برابر را برداشته و با مطرح کردن یک مسئله انطباق کمترین خطای چگالی گوسی به بهبود این روش پرداختند.

چو و همکارانش در [۹] با مشاهده‌ی متوسط‌ها و واریانس‌های تخمین زده شده، خطای ناشی از آستانه گذاری را کاهش داده و به نتایج قابل قبولی در آستانه گذاری دست یافتند. البته بایاس زمانی اضافه می‌شود که دو قله هیستوگرام قابل تشخیص نباشند.

۲-۲-۳- روش‌های حذف اعوجاج فتومتريک مبتنی بر آنتروپی

در این گروه از روش‌ها با استفاده از آنتروپی نواحی پیش‌زمینه و پس‌زمینه و در نظر گرفتن اختلاف مقدار آنتروپی بین تصویر اصلی و تصویر دوسطحی شده به محاسبه‌ی یک مقدار آستانه بهینه پرداخته می‌شود.

کاپور و همکارانش در [۱۰] با در نظر گرفتن پیش‌زمینه و پس‌زمینه‌ی تصویر، به‌عنوان دو منبع سیگنال متفاوت و با بیشینه کردن حاصل جمع آنتروپی‌های دو کلاس، به بهینه‌سازی آستانه گذاری پرداخته‌اند.

لی و همکارانش در [۱۱] و [۱۲] با فرموله کردن مسئله‌ی آستانه گذاری به کمینه کردن یک فاصله‌ی تئوریک اطلاعات رسیدند که این فاصله، فاصله‌ی کولبک^۱ نام دارد. از این معیار فاصله در کمینه کردن مقدار آنتروپی متقاطع^۲ استفاده می‌شود. طبق این روش مقدار آستانه زمانی انتخاب می‌شود که مقدار آنتروپی متقاطع کمینه شود.

¹ Kullback–Leibler

² Cross Entropy

۲-۲-۴- روش‌های حذف اعوجاج فتومتریکی مبتنی بر تشابه نسبی

این الگوریتم‌ها مقدار آستانه را بر مبنای بعضی کمیت‌های نسبی یا معیارهای تشابه میان تصویر اصلی و تصویر دوسطحی شده انتخاب می‌کنند. این نسبت‌ها می‌توانند شکل انطباق لبه، فشردگی شکل، گشتاورهای سطح خاکستری، اتصال، بافت یا پایداری اشیاء مجزا را به خود بگیرند. بعضی از دیگر الگوریتم‌ها به‌طور مستقیم، شباهت تصویر ورودی سطح خاکستری را با تصویر دودویی شده با استفاده از یک معیار فازی، ارزیابی می‌کند. برای مثال به تشابه‌اندازه‌ی فازی^۱ می‌توان اشاره کرد [۱۳].

تسای در [۱۴] تصویر سطح خاکستری را به‌عنوان نسخه مات شده از تصویر دودویی ایده آل فرض می‌کند و آستانه گذاری طوری پایه‌ریزی می‌شود که سه گشتاور اول سطح خاکستری، با سه گشتاور اول تصویر دودویی انطباق داشته باشند.

هرتز و شفر در [۱۵] یک تکنیک چند آستانه‌ای^۲ بر اساس تطبیق لبه^۳ را معرفی کردند. در این روش تصویر بر اساس یک مقدار آستانه اولیه دوسطحی شده و در هر مرحله بر اساس تطبیق لبه‌ی تصویر دوسطحی شده و لبه‌ی تصویر خاکستری به تغییر مقدار آستانه اولیه پرداخته می‌شود. این کار تا زمانی که لبه‌های دو تصویر تشابه نسبی داشته باشند، ادامه پیدا می‌کند.

۲-۲-۵- روش‌های حذف اعوجاج فتومتریکی مبتنی بر اطلاعات فضایی

این دسته از الگوریتم‌ها، نه‌تنها توزیع‌هایی باارزش سطح خاکستری را استفاده می‌کنند، بلکه وابستگی پیکسل‌ها در یک همسایگی را نیز به کار می‌برند. برای مثال احتمالات متن، توابع همبستگی، احتمالات هم رخداد، مدل‌های وابستگی خطی پیکسل‌ها اشاره کرد

کربی و روزنفیلد در [۱۶] از میانگین محلی سطوح خاکستری برای آستانه گذاری استفاده نموده‌اند. چانگ و همکارانش در [۱۷] دو رویکرد برای استفاده از احتمالات هم رخداد پیشنهاد داده‌اند. در این روش در اولین رویکرد تصویر دوسطحی شده باید تا حد ممکن گذرهای پیش‌زمینه به پس‌زمینه و بالعکس داشته باشد و در رویکرد دوم احتمال پیکسل‌های همسایگی که در همان کلاس باقی‌مانده‌اند، تقویت می‌شوند.

¹ Measures of Fuzziness

² Multilevel Thresholding

³ Edge Matching

۲-۲-۶- روش‌های حذف اعوجاج فتومتریکی مبتنی بر اطلاعات محلی

در دوسطحی سازی به روش محلی برای هر کدام از پیکسل‌های تصویر یک آستانه اختیار می‌گردد. روش‌های گسترده‌ای در این حوزه وجود دارد به‌عنوان مثال روش‌های [۱۸-۱۹-۲۰-۲۱] بر پایه‌ی استفاده از واریانس و میانگین محلی نقاط اطراف هر پیکسل و روش [۲۲] بر پایه‌ی مقایسه‌ی سطح خاکستری هر پیکسل با سطح خاکستری پیکسل‌های همسایه‌ی آن است.

نیبلک در [۱۸] روشی جدید در دوسطحی سازی محلی پایه‌گذاری کرد. در روش نیبلک مقدار آستانه از جمع میانگین با مقدار ضریب منفی از انحراف معیار طبق معادله ۲-۱ محاسبه می‌شود.

$$T(x, y) = \mu(x, y) + k \cdot \sigma(x, y) \quad (1-2)$$

که در آن $\mu(x, y)$ و $\sigma(x, y)$ به ترتیب انحراف معیار محلی و میانگین محلی در پنجره‌های $n \times n$ بوده و k یک پارامتر ثابت با مقدار منفی است.

ساولا در [۲۱] روش نیبلک را با بیشتر کردن اثر انحراف معیار ارتقاء داد. در این روش انحراف معیار با تقسیم شدن بر یک عدد ثابت به نام R که بیشینه مقدار انحراف معیار برای تصویر سطح خاکستری است، نرمالیزه می‌شود. این مقدار برای سطوح خاکستری برابر ۱۲۸ است.

بعد از ساولا محققین در [۲۳-۲۴] با انجام یک سری از پیش‌پردازش‌ها و همچنین به‌کارگیری از یک سری فیلترهایی مثل فیلتر وینر، نوفه^۱ تصویر را کاهش داده و با اعمال روش ساولا به تخمین پس‌زمینه پرداخته و با درون‌یابی پیکسل‌های پیش‌زمینه، پیش‌زمینه را تخمین زده و به نتایج قابل‌قبولی دست یافتند. ولی از معایب این روش دقت پایین آن در رفع اعوجاج فتومتریکی تصاویر با درجه تفکیک‌پذیری^۲ پایین است.

از جمله روش‌هایی که با تخمین پس‌زمینه به رفع اعوجاج فتومتریکی می‌پردازد می‌توان به روش درون‌نگاری^۳ اشاره کرد. روش درون‌نگاری اولین بار توسط برتالمیو [۲۵] در سال ۲۰۰۰ مطرح شد. برتالمیو از روش درون‌نگاری به‌منظور رفع مشکلات تصاویر قدیمی و آسیب‌دیده استفاده کرد. همان‌طور که در شکل ۲-۲ مشاهده می‌کنید این روش به‌خوبی اثرات ناشی از تاخوردگی و فرسودگی تصویر را از بین می‌برد. این روش همچنین قادر به حذف متن از تصویر است.

¹ Noise

² Resolution

³ Inpainting

کریمینسی و همکارانش [۲۶] در سال ۲۰۰۴ از روش درون‌نگاری به منظور حذف اشیا از تصویر و ترمیم ناحیه‌ی مربوط به شیء حذف‌شده استفاده کردند.



الف

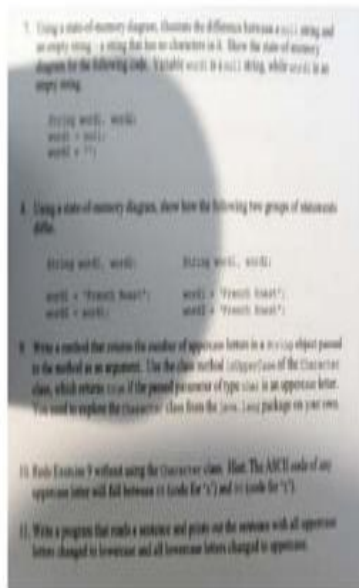


ب

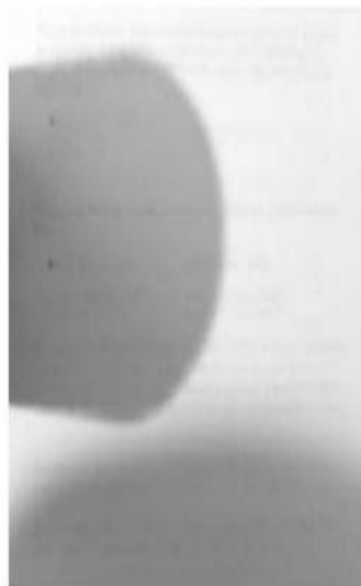
شکل ۲-۲: استفاده از روش درون‌نگاری به منظور بازسازی تصاویر آسیب‌دیده (الف) تصویر ورودی (ب) نتیجه‌ی روش [۲۵]

تیموتی در [۲۷] به بررسی روش‌های درون‌نگاری به منظور بازسازی تصاویر چندلایه و همچنین به بیان مفاهیم درون‌نگاری ویدئو پرداخته و یک روش سریع پیرامون این موضوع ارائه کرده است.

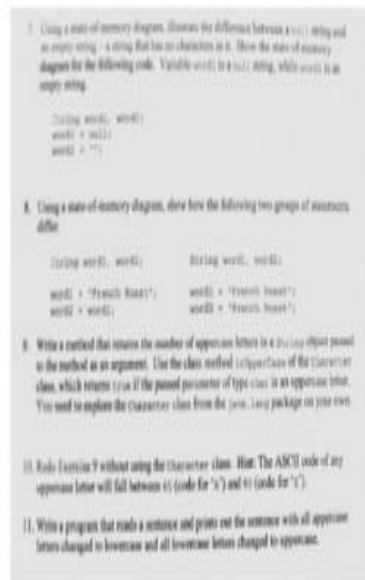
پس از آن ژانگ و همکارانش [۲۸ و ۲۹] از این روش به منظور رفع اعوجاج فتومتریک استفاده کردند. در این روش برای اصلاح سایه، در ابتدا تصویر پس‌زمینه را بر اساس روش درون‌نگاری جدا کرده و پس از آن تصویر پیش‌زمینه به منظور اصلاح بازتاب مورد بررسی قرار می‌گیرد. در شکل ۲-۳ نتیجه‌ی اعمال این روش را مشاهده می‌کنید.



الف



ب



ج

شکل ۲-۳: رفع اعوجاج فتومتریک به کمک روش [۲۹] الف) تصویر ورودی دارای اعوجاج فتومتریک
 ب) استخراج لایه پیش‌زمینه ج) تصویر خروجی بدون اعوجاج

جنارتانن در سال ۲۰۱۲ [۳۰] و راوی و همکارانش در سال ۲۰۱۳ در [۳۱] به بررسی روش‌های رایج درون‌نگاری پرداخته و مزیت‌ها و معایب هر روش به‌منظور درون‌نگاری بر روی تصاویر و ویدئو را بیان نموده‌اند.

۲-۳- اوجاج هندسی

در سال‌های اخیر روش‌های متعددی در زمینه‌ی رفع اوجاج هندسی در تصاویر متنی لاتین پیشنهاد شده است. این تکنیک‌ها به دودسته کلی تقسیم می‌شوند. دسته اول تکنیک‌های مبتنی بر بازسازی سه‌بعدی تصاویر متن هستند و دسته دوم تکنیک‌هایی که با استفاده از شناسایی خطوط متن، اوجاج هندسی را برطرف می‌کنند.

۲-۳-۱- تکنیک‌های مبتنی بر بازسازی سه‌بعدی تصاویر

در دو دهه‌ی گذشته کارهای زیادی در زمینه‌ی رفع اوجاج هندسی بر اساس بازسازی سه‌بعدی تصاویر انجام شده است، در برخی از این روش‌ها دانستن اطلاعات اولیه در مورد نوع اسکنر و همچنین سیستم نورپردازی موردنیاز است؛ که البته این اطلاعات همیشه در دسترس نیست. در برخی دیگر از این روش‌ها با مدل کردن سطح صفحه به سطحی استوانه‌ای شکل به رفع این اوجاج پرداخته می‌شود که البته این روش‌ها، با توجه به این نکته که نوع پیچش صفحه همیشه استوانه‌ای شکل نیست؛ دارای جامعیت کامل نمی‌باشند. در ادامه برخی از مهم‌ترین این روش‌ها را بیان می‌کنیم.

پیلو در [۳۲] روشی بر مبنای مدل فیزیکی تغییر شکل کاغذ ارائه داد. در این روش سطح به‌وسیله‌ی یک مش چندوجهی با ابعاد مناسب و مشخص بودن فاصله‌ی بین گره‌ها شناخته شده است. به‌منظور هموارسازی این سطح از الگوریتم Relaxation استفاده شده است. البته نتایج حاصل از روش حتی برای چشم انسان نیز ناخوانا است.

اولگس و همکارانش در [۳۳] به رفع اوجاج هندسی از تصاویر گرفته شده توسط دوربین‌های دیجیتال پرداخته است. در این روش با توجه به مدل‌سازی سه‌بعدی از صفحه به رفع اوجاج هندسی پرداخته می‌شود. در شکل ۲-۴ نتیجه حاصل روش اولگس را مشاهده می‌کنید که تا حدود زیادی اوجاج هندسی را برطرف نموده است.

The problem with this approach is that we need to find p and q at the new point in order to continue to develop equations for the changes δp and δq in p as we move. Before we address this issue, let us look at the image of a planar surface patch g in the image plane $(E_x, E_y)^T$. We know that a planar surface patch has uniform brightness in the image. Thus a nonzero change in brightness occurs only where the surface is curved. To find the image irradiance equation

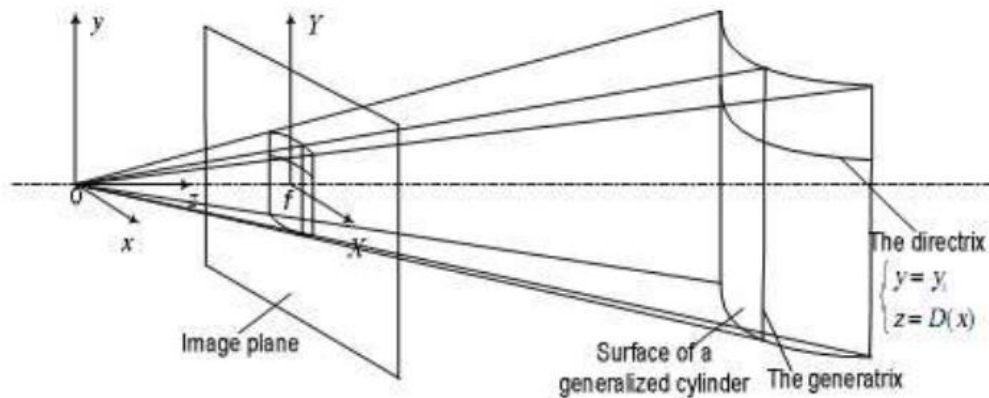
الف

The problem with this approach is that we need to find p and q at the new point in order to continue to develop equations for the changes δp and δq in p as we move. Before we address this issue, let us look at the image of a planar surface patch g in the image plane $(E_x, E_y)^T$. We know that a planar surface patch has uniform brightness in the image. Thus a nonzero change in brightness occurs only where the surface is curved. To find the image irradiance equation

ب

شکل ۲-۴: نتیجه‌ی حاصل از روش اولگس [۳۳] (الف) تصویر دارای اعوجاج هندسی (ب) تصویر خروجی بدون اعوجاج

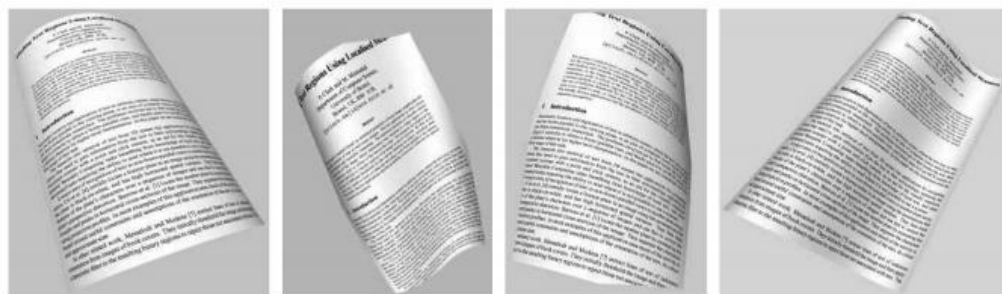
کثو و همکارانش در [۳۴] روشی برای یک‌سوسازی اعوجاج‌های ناخواسته معرفی کرده‌اند که در آن با به وجود آوردن یک مدل استوانه‌ای سعی در تخمین شکل سه‌بعدی صفحه‌ی متن شده است. یک محدودیت این روش فرض شکل استوانه‌ای است؛ چون در برخی از موارد شکل استوانه‌ای جامعیت کامل را دارا نیست. محدودیت دیگر این روش زمانی است که تعداد خطوط در صفحه کم باشد که در این صورت روش قادر به رفع اعوجاج هندسی به‌درستی نیست. شکل ۲-۵ مدل تشکیل تصویر سطح استوانه‌ای سند را نشان می‌دهد.



شکل ۲-۵: تشکیل سطح استوانه‌ای تصویر سند [۳۴]

لیانگ و همکارانش در [۳۵] روشی معرفی کرده‌اند که در آن سطح صفحه با سطوح پیچش دار مدل می‌شود تا شکل سه‌بعدی صفحه را با استفاده از میدان‌های فلوی^۱ تخمین بزنند. فرض این روش ملایم بودن پیچش صفحه و این که دوربین یک دوربین استاندارد روزنه‌ای^۲ می‌باشد، است. شکل ۲-۶ چند نمونه از اسناد پیچش دار استفاده‌شده در این مقاله را نشان می‌دهد.

وادا و همکارانش در [۳۶] روش دیگری را معرفی کردند که در آن از تکنیک شکل به کمک سایه^۳ استفاده می‌شود تا شکل سه‌بعدی سطح متن را بازسازی کند. از معایب این روش نیاز به مشخص بودن اطلاعاتی در مورد نور محیط است که این اطلاعات، معمولاً در دسترس نیست.



شکل ۲-۶: اسناد پیچش دار [۳۵]

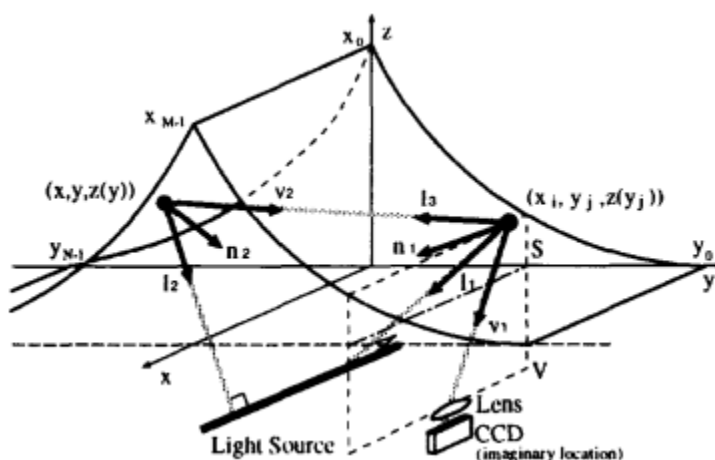
برون و سیلس در [۳۷] یک روش عمومی ضد کجی برای تصاویر پیچش دار دلخواه، بر مبنای شکل سه‌بعدی معرفی کردند. در این روش ابتدا عمق هر نقطه را در تصویر با روش‌های بینایی سلول به دست می‌آورند تا یک نگاشت عمق بسازند، سپس تصویر را با توجه به آن نگاشت رفع پیچش می‌کنند. از جمله مهم‌ترین مشکلات این روش می‌توان به نیازمند بودن این روش به یک سیستم نورپردازی خاص اشاره کرد.

وادا و همکارانش در [۳۸] روشی برای بازیابی سطح سه‌بعدی کتاب با ترکیب بازتاب‌های داخلی انجام‌شده ارائه دادند که فرض‌های این پیشنهاد موازی بودن تیر کتاب با نور اسکن‌کننده و نیز استوانه‌ای بودن سطح کتاب است. عیب این روش حجم بالای محاسبات است که حتی پس از استفاده از روش موزاییک‌سازی نیز کاملاً برطرف نشد. شکل ۲-۷ مدل‌سازی بازتاب داخلی سطوح کتاب در این روش را نشان می‌دهد.

¹ field flow

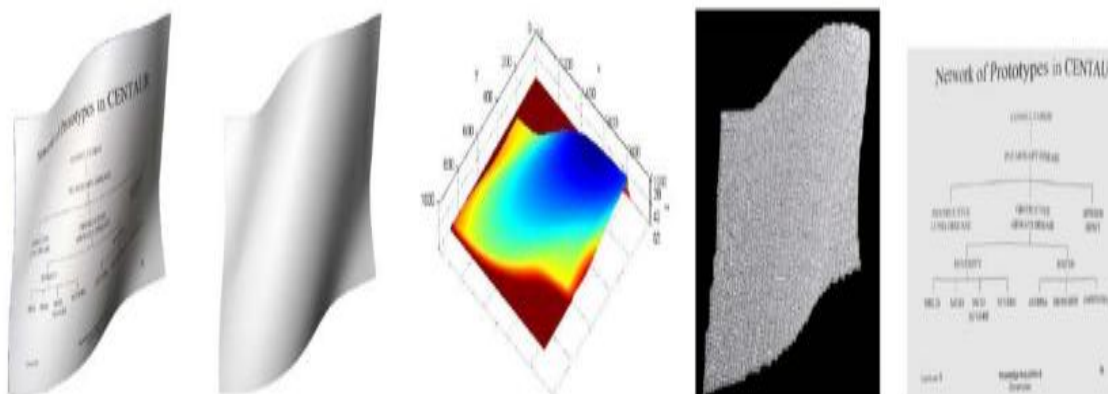
² Pinhole

³ Shape from Shading



شکل ۲-۷: مدل سازی بازتاب داخلی سطوح کتاب [۳۸]

ژانگ و همکارانش در [۲۹] روشی را برای بازیابی اسناد خمیده با استفاده از تکنیک تخمین شکل به وسیله‌ی سایه معرفی کردند. در این روش بر اساس توزیع بازتاب مدلی برای از بین بردن سایه بیان شده و پس از حذف اعوجاج فتومتریک، اعوجاج هندسی را هم از بین می‌برد. در این روش پس از حذف اعوجاج فتومتریک با استفاده از روش درون‌نگاری به رفع اعوجاج هندسی پرداخته می‌شود. همان‌طور که در شکل ۲-۸ مشاهده می‌کنید پس از پیدا کردن تصویر پیش‌زمینه به کمک درون‌نگاری و استخراج مدل سایه به بازسازی سطح صفحه پرداخته و تصویر نهایی بدون اعوجاج ساخته می‌شود؛ اما مشکل اصلی این روش استفاده از مشخصات اسکنر بوده که همیشه در دسترس نیست.



شکل ۲-۸: نتایج حاصل از روش ژانگ [۲۹]

۲-۳-۲ - تکنیک‌های مبتنی بر شناسایی خطوط متن

دسته دوم، تکنیک‌های مبتنی بر شناسایی خطوط متن هستند که در این روش‌ها انحناى خطوط سند مشخص‌کننده‌ی اعوجاج هندسی صفحه بوده و به‌وسیله تقریب زدن یک تابع به هر خط، متن دارای اعوجاج را می‌توان صاف کرد. برای مثال لاوایلی و همکارانش در [۳۹] از یک مدل تحلیلی، برای شناسایی خطوط متن استفاده کرده‌اند؛ اما این روش در صورت ناهمگون بودن فاصله بین خطوط ناکارآمد است. شکل ۲-۹ نتایج حاصل از این الگوریتم را نشان می‌دهد.

people are lesse Philofophers and more foolish, then *Platoes* Phylodoxes, or lovers of their owne opinions. We must knowe whether fire be hot, whether snowe be white, whether in our knowledge there be any thing hard or soft. And touching the answeres, wherefore they tell old tales, as to him who made a doubt of heate, to whom one replied, that to try he should cast himselfe into the fire; to him that denied the yfe to be cold, that he should put some in his bosome; they are most vnworthie the profession of a Philosopher. If they had leaft vs in our owne naturall estate, admitting of strange appa-ances, as they present themselves vnto vs by our senses, and had suffred vs to follow our naturall appetites, directed by the condition of our birth, they should then have reason to speake so. But from them it is, that we have learnt to become judges of the world; it is from them we hold this conceit, that mans reason is the generall controuller of all that is, both without and within heavens-vault; which embraceth al, and can do all by meanes whereof, all things are knowne and discerned. This answer were good among the Caniballs, who without any of *Aristotles* precepts, or so much as knowing the name of Phisike, enjoy most happily, a long, a quiet, and a peaceable life. This answer might happily avails more, and be of more force, then all those they can
E c borrow

الف

people are lesse Philofophers and more foolish, then *Platoes* Phylodoxes, or lovers of their owne opinions. We must knowe whether fire be hot, whether snowe be white, whether in our knowledge there be any thing hard or soft. And touching the answeres, wherefore they tell old tales, as to him who made a doubt of heate, to whom one replied, that to try he should cast himselfe into the fire; to him that denied the yfe to be cold, that he should put some in his bosome; they are most vnworthie the profession of a Philosopher. If they had leaft vs in our owne naturall estate, admitting of strange appa-ances, as they present themselves vnto vs by our senses, and had suffred vs to follow our naturall appetites, directed by the condition of our birth, they should then have reason to speake so. But from them it is, that we have learnt to become judges of the world; it is from them we hold this conceit, that mans reason is the generall controuller of all that is, both without and within heavens-vault; which embraceth al, and can do all by meanes whereof, all things are knowne and discerned. This answer were good among the Caniballs, who without any of *Aristotles* precepts, or so much as knowing the name of Phisike, enjoy most happily, a long, a quiet, and a peaceable life. This answer might happily avails more, and be of more force, then all those they can
E c borrow

ب

شکل ۲-۹: نتیجه‌ی اعمال روش لاوایلی [۳۹] (الف) تصویر ورودی دارای اعوجاج هندسی (ب) تصویر خروجی بدون اعوجاج هندسی

ماسالووویچ در [۴۰] روشی به نام^۱ (SKEL) را معرفی کرده است. ایده‌ی اصلی این روش بر اساس استخراج اسکلت بیرونی تصاویر متنی است که در آن شاخه‌های به‌هم‌پیوسته توسط منحنی‌هایی تقریب زده می‌شوند و فضای درون خطوط متن را تعیین می‌کنند. وی همچنین روشی برای تخمین بدفرمی کل تصویر با استفاده از فضاهاى بین خطی ارائه کرده است. شکل ۲-۱۰ تخمین بدفرمی تصویر و نتیجه‌ی بازسازی تصویر به کمک این روش را نشان می‌دهد.

¹ Skeletonization

HIGGINS (*brusquely, recognizing her with unconcealed disappointment, and at once, babylike, making an intolerable grievance of it*): Why, this is the girl I jotted down last night. Shes no use: I've got all the records I want of the Lisson Grove lingo; and I'm not going to waste another cylinder on it. (*To the girl*) Be off with you: I dont want you.

THE FLOWER GIRL: Dont you be so saucy. You aint heard what I come for yet. (*To MRS PEARCE, who is waiting at the door for further instructions*) Did you tell him I come in a taxi?

MRS PEARCE: Nonsense, girl! what do you think a gentleman like Mr Higgins cares what you came in?

THE FLOWER GIRL: Oh, we are proud! He aint above giving lessons, not him: I heard him say so. Well, I aint come here to ask for any compliment; and if my money's not good enough I can go elsewhere.

HIGGINS: Good enough for what?

THE FLOWER GIRL: Good enough for ye-oo. Now you know, dont you? I'm come to have lessons, I am. And to pay for em too: make no mistake.

الف

HIGGINS (*brusquely, recognizing her with unconcealed disappointment, and at once, babylike, making an intolerable grievance of it*): Why, this is the girl I jotted down last night. Shes no use: I've got all the records I want of the Lisson Grove lingo; and I'm not going to waste another cylinder on it. (*To the girl*) Be off with you: I dont want you.

THE FLOWER GIRL: Dont you be so saucy. You aint heard what I come for yet. (*To MRS PEARCE, who is waiting at the door for further instructions*) Did you tell him I come in a taxi?

MRS PEARCE: Nonsense, girl! what do you think a gentleman like Mr Higgins cares what you came in?

THE FLOWER GIRL: Oh, we are proud! He aint above giving lessons, not him: I heard him say so. Well, I aint come here to ask for any compliment, and if my money's not good enough I can go elsewhere.

HIGGINS: Good enough for what?

THE FLOWER GIRL: Good enough for ye-oo. Now you know, dont you? I'm come to have lessons, I am. And to pay for em too: make no mistake.

ب

شکل ۲-۱: نتیجه‌ی روش ماسالوویچ [۴۰] الف (تخمین بدفرمی تصویرب) تصویر بازسازی شده

اولگس و همکارانش در [۴۱] الگوریتمی برای حذف اعوجاج پرسپکتیو و پیچش صفحه ارائه کرده‌اند. در این الگوریتم فرض بر این است که متن حاوی خطوط کاملاً موازی بافاصله خط یکسان است. در شکل ۲-۱۱ نتیجه اعمال این روش را مشاهده می‌کنید.

کاو در [۴۲] از تبدیلات فضایی بین تصویر پیچش دار و رفع پیچش شده استفاده کرده است. در این روش فرض بر این است که سطح عکس، استوانه است، بنابراین از یک تبدیل برای مسطح کردن تصویر متن استفاده می‌شود. این روش به محاسبات پیچیده نیاز دارد بنابراین زمان‌بر است. در شکل ۲-۱۲ نتیجه‌ی اعمال این روش نشان داده شده است.

mikrofonen nahm er den Gesang von Seepferdchen auf, mischte diesen mit dem Rhythmus von Gewitterdonner, dem Geheul von Moorhunden, dem unhörbaren Geschrei von Fledermäusen, dem Stöhnen von Friedhofswürmern und machte selber noch ein paar sehr eigenwillige Geräusche dazu. Dann ließ er das Ganze rückwärts mit doppelter Geschwindigkeit ablaufen. So ähnlich, bestätigte Qwert, höre sich die Musik in seiner Heimat an. Wir anderen gingen immer raus, wenn er sein Essen zu sich nahm.

الف

mikrofonen nahm er den Gesang von Seepferdchen auf, mischte diesen mit dem Rhythmus von Gewitterdonner, dem Geheul von Moorhunden, dem unhörbaren Geschrei von Fledermäusen, dem Stöhnen von Friedhofswürmern und machte selber noch ein paar sehr eigenwillige Geräusche dazu. Dann ließ er das Ganze rückwärts mit doppelter Geschwindigkeit ablaufen. So ähnlich, bestätigte Qwert, höre sich die Musik in seiner Heimat an. Wir anderen gingen immer raus, wenn er sein Essen zu sich nahm

ب

شکل ۲-۱۱: نتیجه‌ی روش اولگس [۴۱] الف (تصویر دارای اعوجاج هندسی ب) نتیجه‌ی روش اولگس



الف

ب

شکل ۲-۱۲: نتیجه‌ی روش کاو [۴۲] الف) تصویر ورودی با اعوجاج هندسی ب) نتیجه‌ی روش اولگس

در روش گتوس [۴۳] که به روش (SEG) نیز معروف است؛ خمیدگی هر کلمه به تنهایی توسط شیب آن کلمه برطرف می‌گردد. سپس با توجه به چپ‌ترین کلمه‌ی موجود در آن خط هم‌ترازی عمودی صورت می‌گیرد.

در این روش به خاطر چرخش و جابجایی صورت گرفته برای قرار گرفتن در مکان نهایی شکل حروف تا حدود زیادی تخریب خواهد شد. این روش تنها بر روی اسناد بدون اجزای گرافیکی کارایی دارد ولی قادر به حذف اجزای گرافیکی با اعمال فیلتری که اجزای گرافیکی را بر اساس سائزشان حذف می‌کند، می‌باشد.

از میان روش‌های گفته شده در دسته دوم دو روش عمده جبران اعوجاج تصویر با نام‌های ارائه‌ی تصویر اسکلتی پیوسته برای جبران اعوجاج تصویر متن (SKEL) [۴۰] و بخش بندی برای جبران اعوجاج تصویر متن (SEG) [۴۳] بیشتر مورد توجه هستند.

ایده‌ی اصلی روش SKEL بر اساس استخراج اسکلت بیرونی تصاویر متنی است که در آن شاخه‌های به هم پیوسته که توسط منحنی‌هایی تقریب زده می‌شوند، فضای درون خطوط متن را تعیین می‌کنند. همچنین روش SEG بر اساس بخش‌بندی کلمات موجود در متن و تصحیح شیب این کلمات با استفاده از خطوط کرسی تعیین شده برای آن‌ها و در نهایت انتقال عمودی این کلمات چرخیده شده است. این دو روش، در کاهش اعوجاج چرخش سه بعدی متون لاتین کاربرد دارند، اما هیچ‌کدام از این روش‌ها نمی‌توانند به‌طور

¹ Segmentation

کامل، اعوجاج‌های مربوط به این متون را برطرف کنند. روش SKEL اعوجاج‌های انحنای صفحه را می‌تواند برطرف کند ولی اعوجاج‌های پرسپکتیو را برطرف نمی‌کند. همچنین اعوجاج‌های مربوط به فرمول‌ها را نیز برطرف نمی‌کند. روش SEG، اعوجاج مربوط به فرمول‌ها و انحنای صفحه را برطرف می‌کند، اما برای از بین بردن اعوجاج پرسپکتیو ناتوان است.

در تمامی مقالات بررسی شده تاکنون الگوریتم‌های ارائه‌شده بر روی تصاویر متنی لاتین است و کارهای انجام‌شده در زمینه‌ی رفع اعوجاج‌های هندسی تصاویر متن فارسی محدود بوده است.

ده بوید و همکارانش در [۴۴] برای کاهش چرخش سه‌بعدی در تصاویر متنی فارسی تصویربرداری شده توسط دوربین الگوریتمی جدید معرفی کرده‌اند. هدف از جبران چرخش سه‌بعدی جبران اعوجاج صفحه و مسطح کردن متون تصویربرداری شده توسط دوربین می‌باشد به گونه‌ای که این متون توسط سیستم OCR خوانا باشند.

ایده اصلی الگوریتم پیشنهادی آنها بر اساس بخش بندی عناصر تشکیل دهنده متن می‌باشد در این الگوریتم روشی مقاوم برای تعیین خطوط کرسی بالا و پایین کلمات پیشنهاد شده است که از آن برای تعیین شیب کلمات استفاده می‌شود. همچنین انتقال کلمات چرخیده با توجه به برازش یک منحنی درجه دوم به مراکز کلمات یک خط نسبت به خط افق صورت می‌پذیرد. تصویر نمونه ورودی و خروجی برای این الگوریتم در شکل ۲-۱۳ آمده است.

گل پرداز در [۴۵] الگوریتمی برای کاهش اعوجاج چرخش سه‌بعدی روی تصاویر متنی حاصل شده از دوربین دیجیتال پیشنهاد نموده است. ایده اصلی الگوریتم، بر اساس بخش‌بندی عناصر تشکیل دهنده متن است. در این الگوریتم، خطوط کرسی بالا و پایین کلمات مشخص می‌شود و از آن برای تعیین شیب کلمات استفاده می‌گردد. همچنین انتقال کلمات چرخیده شده با توجه به برازش یک منحنی درجه دوم به مراکز کلمات یک خط نسبت به خط افق صورت می‌پذیرد. تصویر ورودی و تصویر حاصل از پیاده‌سازی این الگوریتم در شکل ۲-۱۴ نشان داده شده است.

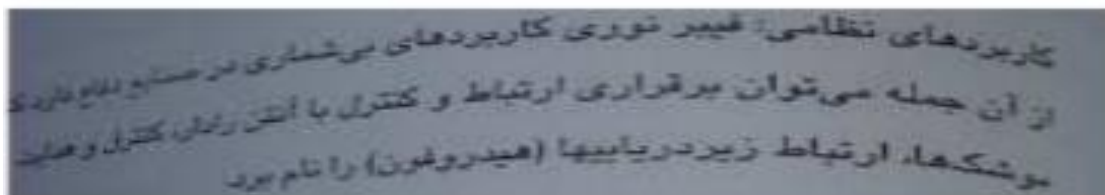
شکل ۸-۱۲ فرآیند رمزگذاری دوبعدی را برای پیمایش یک خط نشان می‌دهد. توجه کنید که مراحل اولیه این روش، به یافتن چندین عنصر متغیر^۳ مهم هدایت شده‌اند: a_0, a_1, a_2, b_1 و b_2 . عنصر متغیر، توسط استاندارد، به عنوان پیکسلی تعریف می‌شود که مقدارش متفاوت از پیکسل قبلی در همان خط است. مهمترین عنصر متغیر a_0 (عنصر مرجع) است، که یا برابر با مکان عنصر متغیر سفید فرضی در سمت چپ اولین پیکسل هر خط رمزگذاری می‌شود یا از حالت رمزگذاری قبلی تعیین می‌شود. حالت‌های رمزگذاری در ادامه بحث می‌شود. پس از تعیین مکان a_0, a_1 به عنوان مکان عنصر متغیر بعدی در سمت راست a_0 روی خط رمزگذاری فعلی، a_2 به عنوان عنصر متغیر بعدی در سمت راست a_1 روی خط رمزگذاری فعلی، b_1 به عنوان عنصر متغیر یا مقدار مخالف (با a_0) و در سمت راست a_0 در خط مرجع (یا خط قبلی)، و b_2 به عنوان عنصر متغیر بعدی در سمت راست b_1 روی خط مرجع مشخص می‌شود. اگر هر کدام از این عناصر متغیر تشخیص داده نشوند، مکان آن، نسبت به یک فرضی در سمت راست آخرین پیکسل، روی خط مناسب تعیین می‌شود. شکل ۸-۱۵ دو نمونه از روابط کلی بین عناصر متغیر مختلف را نشان می‌دهد.

الف

شکل ۸-۱۴ فرآیند رمزگذاری دوبعدی را برای سیمایس یک خط نشان می‌دهد. توجه کنید که مراحل اولیه این روش، به یافتن چندین عنصر متغیر^۳ مهم هدایت شده‌اند: a_0, a_1, a_2, b_1 و b_2 . عنصر متغیر، توسط استاندارد، به عنوان پیکسلی تعریف می‌شود که مقدارش متفاوت از پیکسل قبلی در همان خط است. مهمترین عنصر متغیر a_0 (عنصر مرجع) است، که یا برابر با مکان عنصر متغیر سفید فرضی در سمت چپ اولین پیکسل هر خط رمزگذاری می‌شود یا از حالت رمزگذاری قبلی تعیین می‌شود. حالت‌های رمزگذاری در ادامه بحث می‌شود. پس از تعیین مکان a_0, a_1 به عنوان مکان عنصر متغیر بعدی در سمت راست a_0 روی خط رمزگذاری فعلی، a_2 به عنوان عنصر متغیر بعدی در سمت راست a_1 روی خط رمزگذاری فعلی، b_1 به عنوان عنصر متغیر یا مقدار مخالف (با a_0) و در سمت راست a_0 در خط مرجع (یا خط قبلی)، و b_2 به عنوان عنصر متغیر بعدی در سمت راست b_1 روی خط مرجع مشخص می‌شود. اگر هر کدام از این عناصر متغیر تشخیص داده نشوند، مکان آن، نسبت به یک فرضی در سمت راست آخرین پیکسل، روی خط مناسب تعیین می‌شود. شکل ۸-۱۵ دو نمونه از روابط کلی بین عناصر متغیر مختلف را نشان می‌دهد.

ب

شکل ۲-۱۳: نتیجه روش ده بوید [۴۴] (الف) تصویر ورودی با اعوجاج هندسی (ب) تصویر خروجی بدون اعوجاج



الف

کاربردهای نظامی: عبور نوری کاربردهای بی‌شماری در صنایع دفاع دارد که از آن جمله می‌توان برقراری ارتباط و کنترل یا آنتن رادار، کنترل و هدایت موشک‌ها، ارتباط زیردریاییها (هیدروگون) را نام برد.

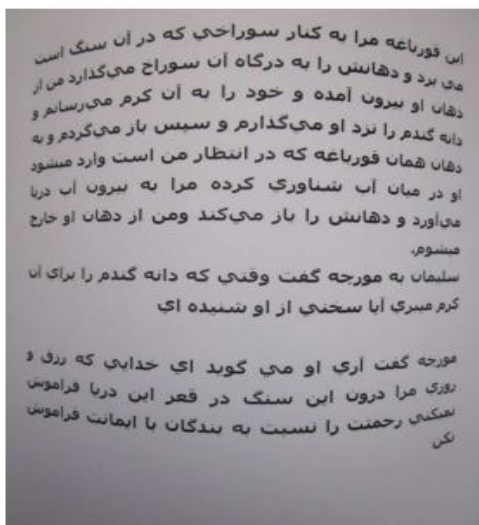
ب

شکل ۲-۱۴: نتیجه‌ی روش گل پرداز [۴۵] (الف) تصویر ورودی با اعوجاج هندسی (ب) تصویر خروجی بدون

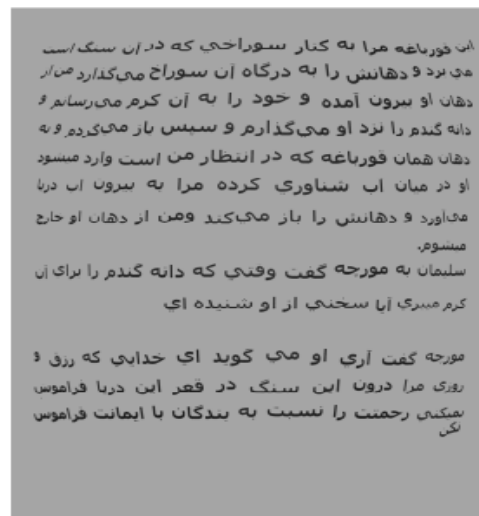
اعوجاج هندسی

خسروی راد در [۴۶] الگوریتمی جدید جهت رفع اعوجاج چرخش سه‌بعدی در تصاویر اسناد فارسی پیشنهاد داده است. در الگوریتم پیشنهادی او ابتدا پیش‌پردازش‌های لازم از جمله دوسطحی سازی، آغشته سازی و تشخیص خطوط زمینه بر روی تصویر انجام شده سپس کلمات خطوط رفع اعوجاج شده بر روی خطوط زمینه قرار می‌گیرند. همان‌طور که در شکل ۲-۱۵ مشاهده می‌کنید؛ مشکلاتی از جمله تغییر اندازه‌ی کلمات و دقت پایین در رفع اعوجاج هندسی این روش مشهود است؛ و در ضمن این روش در حذف اعوجاج تصاویر دوستونه و دارای اشکال گرافیکی ناتوان است.

شامقلی در [۴۷] روشی برای رفع اعوجاج هندسی بیان نموده است. در این روش با استفاده از تخمین خمیدگی خطوط در بالا و پایین صفحه و با انتقال سطح خمیده‌ی سند به سطح مستطیلی شکل اعوجاج هندسی تصویر برطرف می‌شود. البته این روش توانایی رفع اعوجاج اسناد دو ستونه را ندارد و در ضمن به علت انتقال تمام نقاط تصویر (حتی نواحی سفید بدون اطلاعات) روشی زمان‌بر می‌باشد. در شکل ۲-۱۶ نتایج اعمال این روش نشان داده شده است.



الف



ب

شکل ۲-۱۵: نتیجه‌ی روش خسروی راد [۴۶] (الف) تصویر ورودی با اعوجاج هندسی (ب) تصویر خروجی بدون اعوجاج هندسی



«تو که بیشتر از من اونو می‌بینی»
 سؤال من آن‌طور که انتظار داشتم، قفل دهان او را باز نکرده بود.
 زیر لب گفتم: «توی محفل تاز، نمی‌شه راحت حرف زده بعد دوباره سعی کردم او
 را به حرف بیارم: «تازکی» یا کسی بیرون ترفتن؟»
 ومنتش نه، گاهی با کافر بیرون می‌رم. دو هفته پیش هم با اریک به گردش رفتیم.
 جسیکا بعد از گفتن این حرف پنجم‌هاش را چرتخی داد و من آغاز داستان بلند
 او را حس کردم نمی‌تونستم این فرصت را نادیده بگیرم.
 پرسیدم: «اریک یورکی؟» کسی از کی دعوت کرده؟
 او ناله‌ای کرد و بیشتر سر شوق آمد. گفتم: «البته که اون از من دعوت کرده من هم
 هر کاری کردم، نتونستم بهش جواب رد بدم»
 با افسردگی پرسیدم: «چه وقت با هم بیرون رفتین؟» می‌دانستم که او اشتیاق من را به
 حساب علاقه‌مندی‌ام خوانده گذاشته گفتم: «همه چیزو برام تعریف کن.»
 او قصه‌اش را شروع کرد و من روی صندلی‌ام آرام گرفتم. حالا احساس راحتی
 بیشتری می‌کردم. در ظاهر توجه زیادی به حرف‌های او نشان می‌دادم؛ گاهی زیربسی
 چیزی برای همراهی او می‌گفتم و در صورت نیاز، واکنش می‌کردم که از تروس به
 نفس نفس افتادم! وقتی داستان او درباره اریک به آخر رسید، بدون مقدمه داستان
 مربوط به کافر را شروع کرد.
 فیلم کمی زودتر از حد انتظار ما شروع می‌شد، بنابراین جسیکا تصمیم گرفت که
 سانس غروب را تماشا کنیم و بعد از فیلم شام بخوریم. از اینکه با همه پیشنهادهای او
 موافقت کرده بودم، احساس خوشحالی می‌کردم؛ در هر حال من به چیزی که
 می‌خوانشتم، رسیدم و مردم و چارلی تا مدتی دست از سر من برمی‌داشتند.
 کاری کردم که جسی در حین پخش پیش‌بردهای فیلم هم، به وراجی خودش
 ادامه دهد تا به این ترتیب مجبور نباشم به آنها توجه کنم. اما وقتی که فیلم شروع شد
 من عصبی شدم. زوج جوانی کنار ساحل قدم می‌زدند، دست‌هایشان را تکان می‌دادند
 و با حالت احساسی کاذبی در مورد محبت دوجانبه خودشان، داد سخن می‌دادند. در

1. Eric Yeckie

الف

ب

شکل ۲-۱۶: نتیجه‌ی روش شاملی [۴۷] الف) تصویر ورودی با اعوجاج هندسی ب) تصویر خروجی بدون اعوجاج هندسی

فصل سوم

تئوری

۱-۳- مقدمه

در این فصل به بررسی تئوری روش‌هایی پرداخته می‌شود که در این پایان‌نامه مورد استفاده قرار گرفته است.

۲-۳- پردازش تصویر مبتنی بر ریخت‌شناسی^۱

زبان ریخت‌شناسی مبتنی بر ریاضی، تئوری مجموعه‌هاست و ریخت‌شناسی روشی قدرتمند برای برخورد با مسائل پردازش تصویر است. در ریخت‌شناسی مبتنی بر ریاضی مجموعه‌ها بیانگر اشیاء موجود در یک تصویر هستند. به‌عنوان مثال، مجموعه‌ی تمام پیکسل‌های سفید در تصویر، یک توصیف مبتنی بر ریخت‌شناسی از تصویر است. در تصاویر دودویی، مجموعه‌های موردنظر، اعضای مجموعه اعداد صحیح دوبعدی هستند که در آن هر عضو یک مجموعه، یک چندتایی (بردار دوبعدی) با مختصات (X, Y) از یک پیکسل سفید (سیاه، بسته به قرارداد) در تصویر می‌باشد. [۴۸]

بحث در مورد ریخت‌شناسی را با معرفی عملگرهای فرسایش^۲، گسترش^۳، بستن^۴ ادامه می‌دهیم.

۱-۲-۳- فرسایش

اگر A و B مجموعه‌هایی در Z باشند فرسایش A و B که با $A \ominus B$ نمایش داده می‌شود و در اینجا B جزء ساختاری بوده، به‌صورت رابطه‌ی ۱-۳ تعریف می‌گردد.

$$A \ominus B = \{z \mid (B)_z \subseteq A\} \quad (1-3)$$

این رابطه بیان می‌کند که فرسایش A و B ، مجموعه تمام نقاط Z به نحوی است که اگر B ، به‌اندازه‌ی Z انتقال یابد، زیرمجموعه‌ای از A باشد. به بیان ساده‌تر، همان‌طور که از اسم فرسایش مشخص است، این عملگر ریخت‌شناسی باعث کاهش اندازه‌ی ناحیه‌ی A می‌شود. شکل ۱-۳ ب نتیجه‌ی فرسایش با جزء ساختاری مربعی شکل 3×3 را نشان می‌دهد.

¹ Morphological Image Processing

² Erosion

³ Dilation

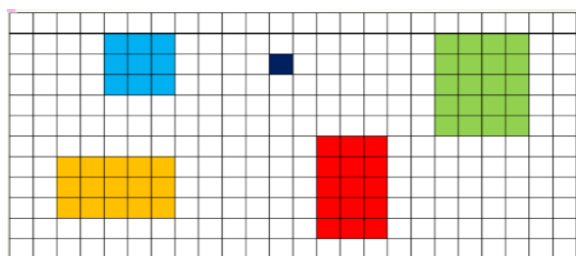
⁴ Closing

گسترش ۲-۲-۳

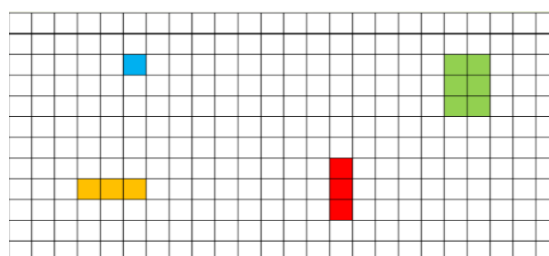
اگر A و B مجموعه‌هایی در Z باشند گسترش A و B که با $A \oplus B$ نمایش داده می‌شود و در اینجا B ، جزء ساختاری بوده، به صورت رابطه‌ی ۲-۳ تعریف می‌گردد.

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (2-3)$$

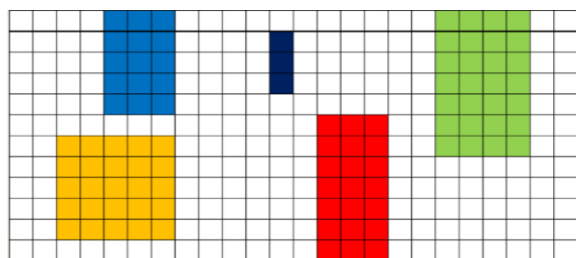
این رابطه بر اساس بازتاب دادن B حول مبدأ مختصات و انتقال دادن این بازتاب به اندازه‌ی Z است؛ بنابراین گسترش A و B ، مجموعه‌ی تمام جابجایی‌های Z است به نحوی که B با A جزء مشترک داشته باشد. به بیان ساده‌تر، همان‌طور که از اسم گسترش مشخص است، این عملگر ریخت‌شناسی باعث افزایش اندازه‌ی ناحیه‌ی A می‌شود. در شکل ۱-۳ ج نتیجه‌ی گسترش با جزء ساختاری ستونی به طول سه پیکسل قابل مشاهده است.



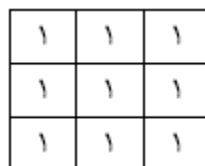
الف



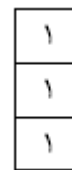
ب



ج



د



ه

شکل ۱-۳: الف) تصویر ورودی (ب) نتیجه‌ی فرسایش با جزء ساختاری مربعی 3×3 (ج) نتیجه‌ی گسترش با جزء ساختاری ستونی به طول ۳ پیکسل (د) جزء ساختاری مربعی 3×3 (ه) جزء ساختاری ستونی به طول ۳ پیکسل

۳-۲-۳ بستن

بستن مجموعه‌ی A با جزء ساختاری B به صورت $A \bullet B$ نمایش داده شده و طبق رابطه‌ی ۳-۳ تعریف می‌گردد.

$$A \bullet B = (A \oplus B) \ominus B \quad (۳-۳)$$

طبق رابطه‌ی ۳-۳ بستن A و B، ترکیبی از گسترش و سپس فرسایش است. از جمله کاربردهایی که این عملگر دارد می‌توان به نرم کردن زاویه‌های مرز خارجی یک‌شکل و چسباندن المان‌هایی در تصویر که بافاصله‌ی کمی از هم قرار دارند، نام برد.

۳-۳ درون نگاری

از لحاظ ریاضی مسئله‌ی درون نگاری یک مسئله‌ی درون‌یابی است [۴۹]، اگر تصویر I_0 را با یک حفره Ω_h درون آن در نظر بگیریم، هدف پیدا کردن تصویر I است که در نواحی بیرون حفره مشابه I_0 بوده و نواحی داخل حفره را نیز پر کند، بنابراین برای پیدا کردن I باید انرژی را طبق رابطه‌ی ۳-۴ کمینه کرد؛ که در آن λ پارامتر نرم‌کننده و χ تابع مشخصه طبق رابطه‌ی ۳-۵ می‌باشد.

$$E(I) = \int_{\Omega} \chi \cdot (I - I_0)^2 dx + \lambda \int_{\Omega} |\nabla I|^2 dx \quad (۴-۳)$$

$$\chi(x) = \begin{cases} 1 & x \in \Omega \setminus \Omega_h \\ 0 & x \geq 0 \end{cases} \quad (۵-۳)$$

بنابراین لازم است رابطه‌ی ۳-۴ را بر اساس لاگرانژ و روش گرادیان نزولی حل کرد که در نهایت از رابطه‌ی ۳-۶ به رابطه‌ی ۳-۷ می‌رسیم.

$$\frac{\partial E}{\partial I} = \frac{\partial F}{\partial I} - \frac{\partial}{\partial u} \frac{\partial F}{\partial I_u} - \frac{\partial}{\partial v} \frac{\partial F}{\partial I_v} = 2[\chi(I - I_0) - \lambda(I_{uu} - I_{vv})] = 0 \quad (۵-۳)$$

$$I_{i,j}^{n+1} = I_{i,j}^n + \Delta t \left(\frac{\lambda}{h^2} (I_{i+1,j}^n + I_{i-1,j}^n + I_{i,j+1}^n + I_{i,j-1}^n - 4I_{i,j}^n) - \chi_{i,j} (I_{i,j}^n - I_{0i,j}) \right) \quad (۶-۳)$$

در این روابط n نشان دهنده‌ی مرحله‌ی تکرار، h پارامتر اندازه‌ی مش^۱، λ پارامتر نرم‌کننده و Δt گام زمانی است که این پارامترها به صورت سعی و خطا انتخاب می‌شوند.

۴-۳- فیلترگذاری مکانی

یک فیلتر مکانی، متشکل از یک منطقه‌ی همسایگی (نوعاً یک مستطیل کوچک) و یک عملیات از پیش تعریف‌شده که روی پیکسل‌های تصویر محصور درون منطقه‌ی همسایگی انجام می‌شود، است. فیلترگذاری یک پیکسل جدید با مختصات برابر با مرکز منطقه‌ی همسایگی و با مقداری که حاصل از عمل فیلترگذاری است، تولید می‌کند. تصویر پردازش‌شده همزمان با قرار گرفتن مرکز فیلتر بر روی تصویر در حال پردازش شدن (فیلترگذاری شدن) به روی هر پیکسل ایجاد می‌شود [۴۸].

فیلترگذاری مکانی انواع متفاوتی دارد که در این بخش به فیلترهای ملایم‌کننده و فیلترهای تیزکننده پرداخته می‌شود.

۳-۴-۱- فیلترهای مکانی ملایم‌کننده

فیلترهای ملایم‌کننده جهت مات کردن و کاهش نوفه به کار می‌روند. مات کردن در اعمال پیش‌پردازش، مانند حذف جزئیات کوچک از یک تصویر قبل از استخراج اجزای بزرگ‌تر آن، متصل کردن نواحی خالی در خطوط و منحنی‌ها و کاهش نوفه کاربرد دارد. از جمله مهم‌ترین فیلترهای مکانی ملایم‌کننده می‌توان به فیلتر ملایم‌کننده‌ی میانگین‌گیر و فیلتر ملایم‌کننده‌ی گوسی که در ادامه معرفی می‌شوند، اشاره کرد.

۳-۴-۱-۱- فیلتر ملایم‌کننده‌ی میانگین‌گیر

هر عضو نقاب یک فیلتر ملایم‌کننده‌ی میانگین‌گیر m در n به صورت رابطه (۳-۸) تعریف شده و عبارت است از میانگین پیکسل‌های موجود در همسایگی نقاب فیلتر می‌باشد.

$$R = \frac{1}{mn} \sum_{i=1}^{mn} z_i \quad (۳-۷)$$

که در رابطه فوق m و n به ترتیب تعداد سطرها و ستون‌های همسایگی نقاب فیلتر و z_i روشنایی پیکسل i ام در این پنجره است. در شکل ۳-۲ دو نمونه از فیلترهای ملایم‌کننده‌ی میانگین‌گیر نشان داده شده است.

¹ Grid Size

$$\frac{1}{9} \times \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

الف

$$\frac{1}{5} \times \begin{array}{|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

ب

شکل ۳-۲: فیلتر ملایم‌کننده‌ی میانگین‌گیر الف) فیلتر ملایم‌کننده‌ی میانگین‌گیر 3×3 ب) فیلتر ملایم‌کننده‌ی میانگین‌گیر سطری به طول ۵ پیکسل

۳-۴-۱-۲- فیلتر ملایم‌کننده‌ی گوسی

فیلتر ملایم‌کننده‌ی گوسی فیلتری است که مقادیر نقاب فیلتر بر اساس رابطه‌ی گوسین به دست می‌آید. معادلات مربوط به این فیلتر را در رابطه‌های ۳-۹ و ۳-۱۰ مشاهده می‌کنید.

$$h_g(n_1, n_2) = e^{-\frac{(n_1^2 + n_2^2)}{2\sigma^2}} \quad (۳-۸)$$

$$h(n_1, n_2) = \frac{h_g(n_1, n_2)}{\sum_{n_1} \sum_{n_2} h_g} \quad (۳-۹)$$

که در این روابط σ مقدار واریانس تابع گوسی بوده و مقادیر n_1 و n_2 به‌اندازه‌ی نقاب مربوط می‌شوند به‌گونه‌ای که برای یک فیلتر 3×3 در 3 مقادیر n_1 و n_2 در مرکز برابر صفر و در بقیه نقاط بسته به محل بین $\{1, 0, -1\}$ است.

دو نمونه از فیلترهای ملایم‌کننده‌ی گوسی را در شکل ۳-۳ مشاهده می‌کنید. شکل ۳-۳ الف یک فیلتر 3×3 گوسی با واریانس ۰,۵ را نشان می‌دهد. در شکل ۳-۳ ب نیز یک نمونه فیلتر سطری گوسی با واریانس ۰,۵ به طول ۵ پیکسل را مشاهده می‌کنید.

0.0113	0.0838	0.0113
0.0838	0.6193	0.0838
0.0113	0.0838	0.0113

الف

0.0003	0.1065	0.7866	0.1065	0.0003
--------	--------	--------	--------	--------

ب

شکل ۳-۳: نقاب‌های فیلتر ملایم‌کننده‌ی گوسی الف) فیلتر ملایم‌کننده‌ی گوسی 3×3 با واریانس ۰,۵

ب) فیلتر ملایم‌کننده‌ی گوسی سطری به طول ۵ پیکسل با واریانس ۰,۵

۳-۴-۲- فیلترهای تیز کننده

هدف اصلی تیزسازی مشخص کردن بهتر پرش‌های شدت روشنایی است. تیزسازی از طریق مشتق‌گیری مکانی امکان‌پذیر می‌باشد و اساساً قدرت پاسخ عملگر مشتق‌گیری با گسستگی شدت روشنایی تصویر در نقطه‌ی اعمال عملگر نسبت دارد؛ بنابراین مشتق‌گیری از تصویر لبه‌ها و سایر گسستگی‌ها را بهبود داده و نقاطی که شدت روشنایی آن‌ها تغییرات آهسته‌تری دارد را تضعیف می‌کند. این گروه از فیلترها به دودسته‌ی خطی و غیرخطی تقسیم می‌شوند. لاپلاسین و گرادیان دو نمونه از این فیلترها می‌باشند.

۳-۴-۲-۱- لاپلاسین

لاپلاسین ساده‌ترین عملگر مشتق‌گیری ایزوتروپیک بوده که برای یک تابع با دو متغیر طبق رابطه‌ی ۳-۱۱ تعریف می‌شود.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (۳-۱۰)$$

از آنجایی که مشتقات با هر مرتبه‌ای، عملیاتی خطی هستند. لذا لاپلاسین عملگری خطی است. جهت بیان این معادله به شکل گسسته، به رابطه‌های ۳-۱۲ و ۳-۱۳ می‌رسیم.

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y) \quad (۳-۱۱)$$

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y) \quad (۳-۱۲)$$

-1	-2	-1
0	0	0
1	2	1

الف

-1	0	1
-2	0	2
-1	0	1

ب

شکل ۳-۵: عملگرهای سوبل الف) عملگر سوبل در راستای x ب) عملگر سوبل در راستای y

۳-۵- روش‌های مبتنی بر افکنش^۱

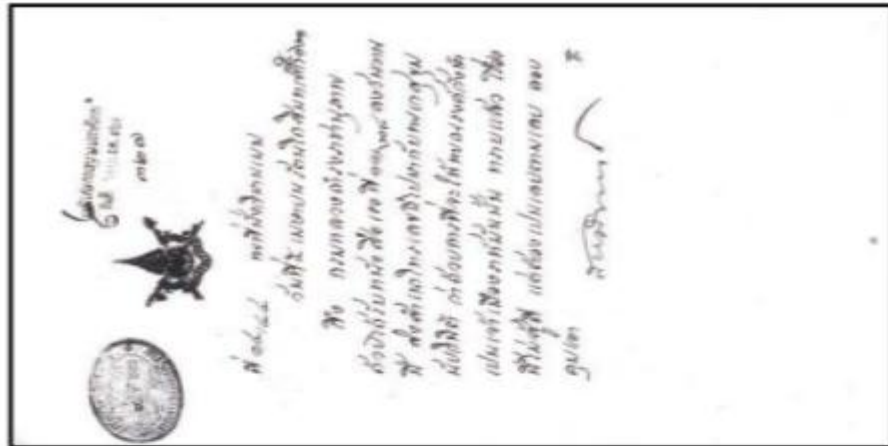
نمودارهای افکنش عموماً از جمع کردن مقادیر تابعی از یک تصویر در راستای افقی و عمودی ایجاد می‌شوند که از روی آن اطلاعاتی در مورد تراکم پیکسل‌های سفید و سیاه تصویر، تغییرات شدت روشنایی و... قابل دستیابی می‌باشد. از این اطلاعات در پردازش تصویر می‌توان جهت تشخیص خطوط یا نواحی سفید بین خطوط نیز استفاده کرد. رابطه‌ی ۳-۱۷ بیان ریاضی این روش را نشان می‌دهد.

$$P(y) = \sum_{1 \leq x \leq M} f(x, y) \quad (۳-۱۶)$$

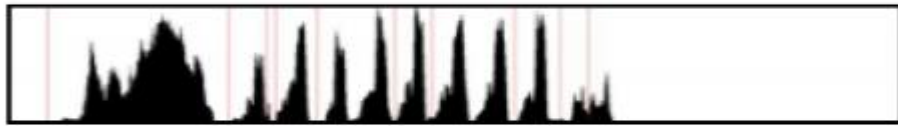
رابطه ۳-۱۷ افکنش افقی را بیان می‌کند که طبق این رابطه مقادیر تابع f در هر سطر جمع می‌شود. بسته به اینکه تابع f چه نوع تابعی باشد، کاربرد این روش متفاوت است. به‌عنوان مثال اگر تابع f شدت روشنایی یک تصویر دوسطحی باشد از روی آن می‌توان جهت تشخیص خط یا نواحی سفید بین خطوط استفاده کرد [۵۰ و ۵۱ و ۵۲].

شکل ۳-۶ نتیجه‌ی افکنش افقی مبتنی بر شدت روشنایی را نشان می‌دهد [۵۳]. بر طبق این تصویر نقاطی که در آن نمودار افکنش بیشترین مقدار را دارد، بیانگر مختصات وقوع خطوط و نقاطی که در آن نمودار کمترین مقدار را دارد، بیانگر نواحی خالی بین خطوط است. همچنین اگر تابع f تابع گرادیان در راستای y باشد، نمودار افکنش افقی اطلاعاتی در مورد چرخش و تغییرات شدت روشنایی در اختیار ما قرار می‌دهد [۵۴].

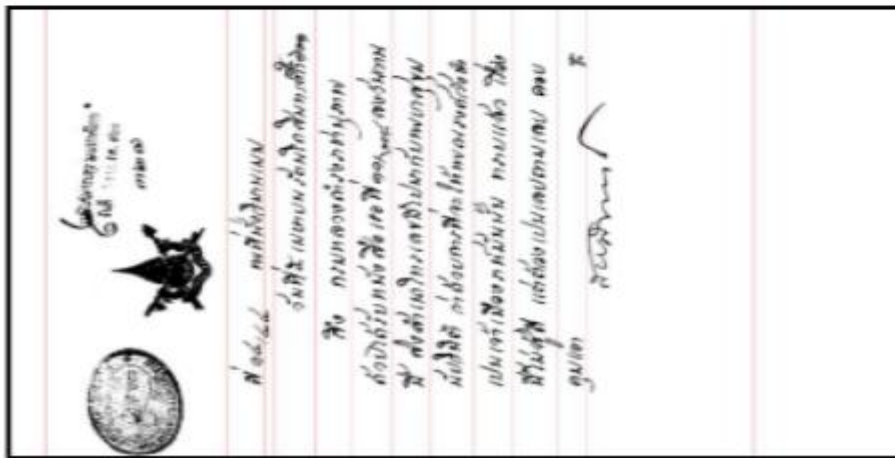
¹ Projection Based Methods



الف



ب

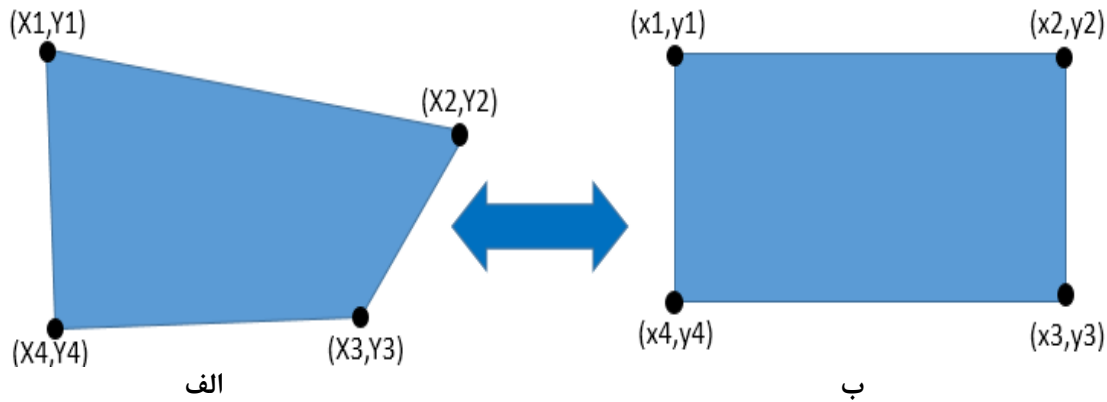


ج

شکل ۳-۶: نتیجه‌ی افکنش افقی مبتنی بر شدت روشنایی [۵۳] الف) تصویر ورودی ب) نمودار افکنش افقی ج) خطوط یافته شده بر اساس نمودار افکنش افقی (تصاویر برای فهم بهتر ۹۰ درجه چرخانده شده‌اند)

۳-۶- تخمین تبدیل پرسپکتیو^۱

یک انتقال هندسی پرسپکتیو که بسیاری از فرآیندهای تصویربرداری را می‌توان به کمک آن مدل کرد، توسط ماتریس انتقال 3×3 با ۸ درجه آزادی تعریف می‌گردد. برای یافتن پارامترهای آن حداقل به چهار نقطه خارج یک خط در تصاویر مبدأ و نتیجه نیاز می‌باشد. در شکل ۳-۷ این انتقال نشان داده شده است.



شکل ۳-۷: انتقال نقاط از چهارضلعی به مستطیل و بالعکس (الف) چهارضلعی (ب) مستطیل

بر طبق [۵۵ و ۵۶] این انتقال را می‌توان به صورت رابطه ۳-۱۸ مدل کرد.

$$X = Hx \quad (۱۷-۳)$$

که در این رابطه X برداری^۲ شامل مختصات نقاط چهارضلعی و x برداری^۳ شامل مختصات نقاط مستطیل و H ماتریس انتقال ۳ در ۳ می‌باشد. فرم باز شده معادله‌ی فوق به صورت رابطه‌ی ۳-۱۹ می‌باشد.

$$\begin{bmatrix} XW \\ YW \\ W \end{bmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (۱۸-۳)$$

طبق این رابطه w که پارامتری مربوط به زاویه‌ی دید می‌باشد، برابر است با:

$$W = gx + hy + 1 \quad (۱۹-۳)$$

¹ Perspective Transform Estimation

² Vector of world plane

³ Vector of Image plane

بنابراین با تقسیم طرفین رابطه‌ی ۳-۱۹ بر W داریم:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}{gx + hy + 1} \quad (۳-۲۰)$$

و پس از ساده‌سازی رابطه‌ی ۳-۲۱ داریم:

$$X = \frac{ax + by + c}{gx + hy + 1} \quad (۳-۲۱)$$

$$Y = \frac{dx + ey + f}{gx + hy + 1} \quad (۳-۲۲)$$

و با بازنویسی روابط ۳-۲۲ و ۳-۲۳ به روابط زیر می‌رسیم:

$$X = ax + by + c - gXx - hYy \quad (۳-۲۳)$$

$$Y = dx + ey + f - gYx - hYy \quad (۳-۲۴)$$

حال با اضافه کردن المان‌هایی با ضرایب صفر در ۳-۲۴ و ۳-۲۵ داریم:

$$X = ax + by + c - 0d + 0e + 0f - gXx - hYy \quad (۳-۲۵)$$

$$Y = 0a + 0b + 0c + dx + ey + f - gYx - hYy \quad (۳-۲۶)$$

حال اگر مختصات چهار نقطه‌ی متناظر از چهارضلعی و مستطیل را داشته باشیم به رابطه‌ی ۳-۲۸ می‌رسیم.

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -X_3x_3 & -X_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -Y_3x_3 & -Y_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -X_4x_4 & -X_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -Y_4x_4 & -Y_4y_4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ X_3 \\ Y_3 \\ X_4 \\ Y_4 \end{bmatrix} \quad (۳-۲۷)$$

با توجه به این نکته که در رابطه‌ی ۲۸-۳ مقادیر $X1$ تا $X4$ و $Y1$ تا $Y4$ و همچنین مقادیر $x1$ تا $x4$ و $y1$ تا $y4$ مشخص است پس با ۸ معادله و ۸ مجهول مواجه شده که به راحتی قابل حل می‌باشد. پس از حل مقادیر ماتریس H به دست می‌آید.

بنابراین با توجه به اینکه هدف انتقال نقاط از چهارضلعی به مستطیل می‌باشد لذا طبق رابطه‌ی ۱۹-۳ مقادیر X و Y مجهول بوده و با ضرب کردن طرفین رابطه‌ی ۱۹-۳ در H^{-1} و با در نظر گرفتن داریم: $H^*H^{-1}=1$

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix}^{-1} \begin{bmatrix} XW \\ YW \\ W \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (28-3)$$

در نهایت با تقسیم کردن طرفین بر W و ساده‌سازی به رابطه‌ی ۳۰-۳ می‌رسیم.

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix}^{-1} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} K1 \\ K2 \\ K3 \end{bmatrix} \quad (29-3)$$

که در این رابطه $K1 = \frac{x}{W}$ و $K2 = \frac{y}{W}$ و $K3 = \frac{1}{W}$ بوده و در هر مرحله با محاسبه‌ی W از روی $K3$ مقادیر X و Y به صورت رابطه‌های ۳۱-۳ و ۳۲-۳ محاسبه می‌شوند:

$$x = k1 * W \quad (30-3)$$

$$y = k2 * W \quad (31-3)$$

فصل چهارم

روش پیشنهادی

۱-۴ - مقدمه

در این فصل ابتدا به معرفی پایگاه‌های داده مورد استفاده پرداخته می‌شود؛ سپس روش ژانگ و همکارانش [۲۹] که به منظور حذف اعوجاج فتومتریک بوده، معرفی می‌شود و پس از آن روش پیشنهادی که بهبود یافته‌ی روش ژانگ بوده، بیان می‌شود. در بخش آخر این فصل روش پیشنهادی پیرامون رفع اعوجاج هندسی معرفی می‌شود.

۲-۴ - معرفی پایگاه‌های داده‌ی مورد استفاده

در این پایان‌نامه از سه پایگاه داده استفاده شده است، پایگاه داده‌ی خسروی راد [۴۶]، پایگاه داده‌ی شامقلی [۴۷] و پایگاه داده‌ی DFKI [۵۷]. در این بخش به توضیح مختصری در مورد هر پایگاه داده می‌پردازیم.

۱-۲-۴ - پایگاه داده‌ی خسروی راد

این پایگاه داده [۴۶] شامل ۱۲۰ تصویر از متون فارسی می‌باشد. بیشتر این تصاویر توسط دوربین دیجیتال Canon IXY 200F, 12.1 Mega Pixel، عکس‌برداری شده است. تعدادی دیگر از این تصاویر توسط اسکنر Canon LIDE 110، تهیه شده است. در تولید نمونه‌های ایجاد شده، از ۱۰ قلم فارسی معروف که توسط نرم‌افزار پرشیا نگار قابل شناسایی هستند، استفاده شده است. این قلم‌ها عبارت‌اند از:

B-lotus, Bnazanin, Mitra, Yagut, Zar, Traffic, Titr, Tahoma, Homa, Times New Romans

در تولید نمونه‌های این پایگاه داده سعی شده است تا انواع اعوجاج‌های سه‌بعدی ممکن ایجاد شوند. در شکل ۱-۴ چند نمونه از این تصاویر را مشاهده می‌کنید.

۲-۲-۴ - پایگاه داده‌ی شامقلی

پایگاه داده‌ی شامقلی [۴۷] شامل ۶۰۰ تصویر اسکن شده از کتب فارسی می‌باشد. این تصاویر توسط اسکنر HP Scanjet G3110 تهیه شده‌اند. در این تصاویر هیچ نورپردازی ویژه‌ای به کار گرفته نشده است. در این پایگاه داده شامل تصاویری به صورت تک‌صفحه‌ای، دو صفحه‌ای، ترکیبی از کلمات انگلیسی و فارسی، اجزای گرافیکی و فرمول می‌باشد. در شکل ۲-۴ چند نمونه از این تصاویر را مشاهده می‌کنید.

صنایع ایران که طبق اساستنامه آن، موسسه‌ای است علمی، پژوهشی، آموزشی و غیر انتفاعی، در همان از وزارت علوم، تحقیقات و فناوری مجوز دریافت نمود. مهندسی صنایع رشته‌ای است که با طراحی، بهبود، و پیاده‌سازی سیستم‌های یکپارچه از افراد، مواد، اطلاعات، تجهیزات و انرژی مرتبط می‌باشد. این رشته بر پایه دانش تخصصی و تبحر در رشته ریاضی و طبیعی، اجتماعی و طراحی بنا شده است تا به کمک این علوم و قوانین به تعیین، پیش‌بینی و ارزیابی نتایج حاصل از سیستم‌های یکپارچه بپردازد.

مهندسی صنایع آن‌گونه که در ایران تدریس می‌شود در واقع مهندسی صنایع کلاسیک است که مبنی بر بهینه‌سازی، بهینه‌سازی خطی یا تحقیق در عملیات و کارسنجی و زمانسنجی می‌باشد. اما مهندسی صنایع نوین مباحثی چون برنامه‌ریزی تعاملی، آمار مهندسی، تریز و تجزیه و تحلیل سیستم‌ها را نیز در بر می‌گیرد. در حال حاضر، در مقطع تحصیلی کارشناسی ارشد، گرایش

لذت بردن می‌خوانم. جذاب‌ترین همان دنیای دیگری که پیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. مین شانزده تا بیست و شش سالگی خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم. برای خودم کسی بشوم. آگاهی‌هایم را بیشتر کنم و بدین‌ترتیب به روحم شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکرم را، علاقم را، رویاهایم را و افق‌هایم را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، رویاها و تأملات دیگران را در داستان‌ها یا نوشته‌ها

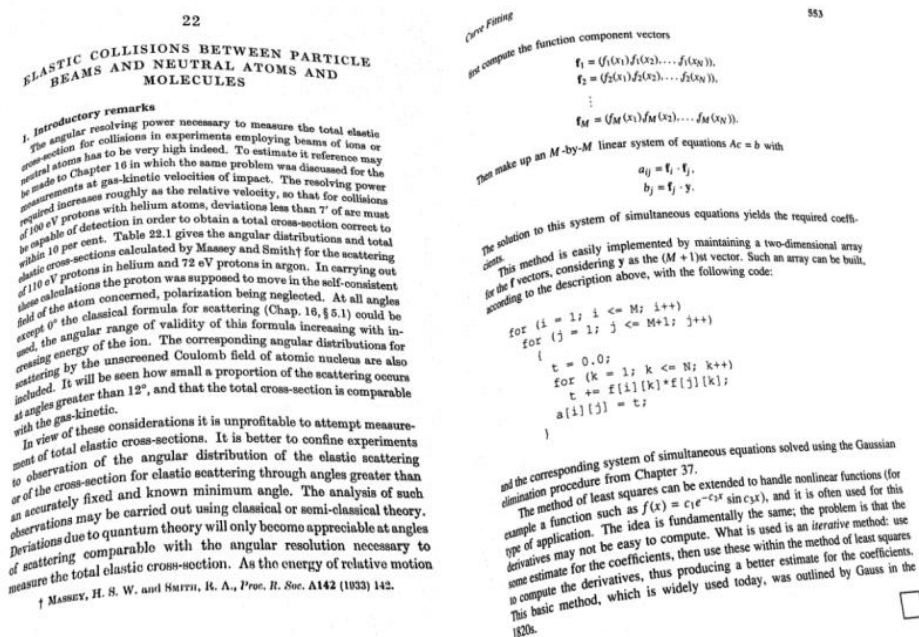
وقتی کتابی توی جیبیت یا توی کیفیت داری، مخصوصاً وقت‌هایی که غمگینی و غصه‌دار. مثل این است که صاحب یک دنیای دیگر هستی، دنیایی که می‌تواند شادی را به تو برگرداند. در دوران نوجوانی ناشادم، فکر خواندن همچو کتابی تسلایی بود که در تمام طول روز مدرسه کمکم می‌کرد. در مدرسه آنقدر خمیازه می‌کشیدم که چشم‌هایم پر از آب می‌شد. بعدها هم در زندگی‌ام، فکر خواندن کتابی که دوست داشتم کمکم می‌کرد تا جلسات اجباری ملاقات‌هایم را راحت‌تر تاب بیاورم. جلساتی که یا از سر تکلیف یا فقط از سر ادب باید در آنها شرکت می‌کردم. بگذارید فهرستی بدهم از دلایل خواندن کتاب‌هایی که نه برای کار یا برای خودسازی و آموختن، بلکه فقط برای

برای کار یا برای خودسازی و آموختن، بلکه فقط برای لذت بردن می‌خوانم. جذاب‌ترین همان دنیای دیگری که پیشتر یاد کردم، شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. مین شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم، برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین‌ترتیب به روحم شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکرم را، علاقم را، رویاهایم را و افق‌هایم را که

شکل ۴-۱: نمونه تصاویر پایگاه داده‌ی خسروی‌راد [۴۶]

۴-۲-۳ - پایگاه داده‌ی DFKI

این پایگاه داده [۵۷] متشکل از ۱۰۲ تصویر از متون انگلیسی می‌باشد. این تصاویر توسط دوربین دیجیتال عکس‌برداری شده‌اند. در تهیه‌ی این تصاویر از هیچ‌گونه نورپردازی خاصی استفاده نشده است. این پایگاه داده به‌عنوان پایگاه داده‌ی ارائه‌شده در کارگاه آموزشی^۱ CBDR2007 نیز معروف می‌باشد. در شکل ۴-۳ دو نمونه از تصاویر این پایگاه داده را مشاهده می‌کنید.

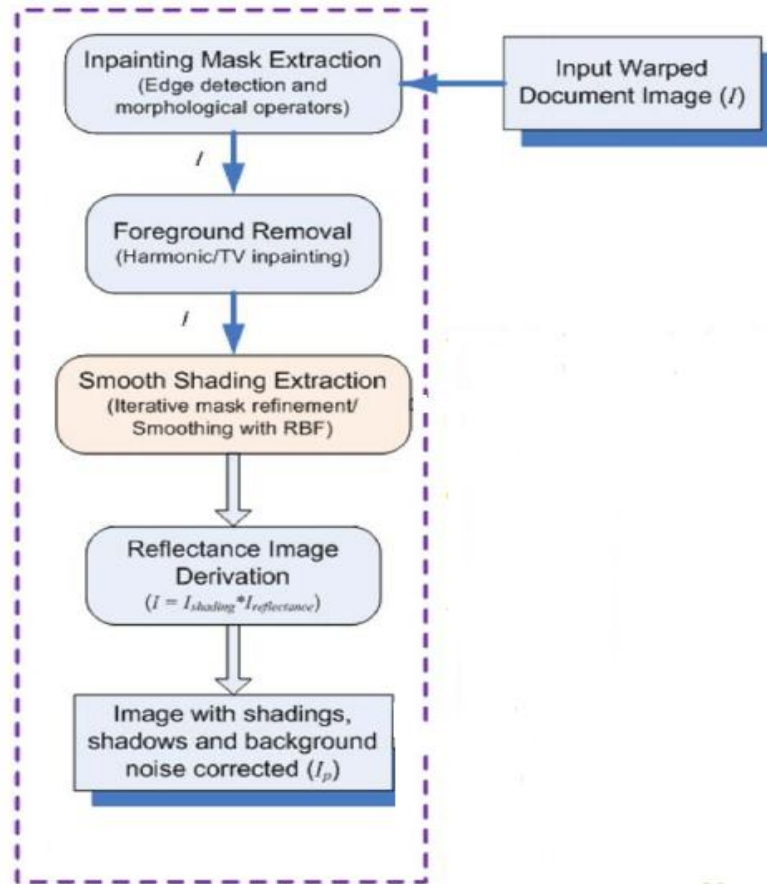


شکل ۴-۳: دو نمونه از تصاویر پایگاه داده‌ی DFKI [۵۷]

¹ Second International Workshop on Camera-Based Document Analysis and Recognition

۳-۴- اعوجاج فتومتریکی

ژانگ و همکارانش در [۲۹] با معرفی یک روش مبتنی بر درون‌نگاری به رفع اعوجاج فتومتریکی پرداخته‌اند. در این بخش روش ژانگ معرفی می‌گردد. شکل ۴-۴ فلوجارت مراحل انجام‌شده در این روش را نشان می‌دهد. در ادامه به توضیح هر مرحله می‌پردازیم.



شکل ۴-۴: فلوجارت حذف اعوجاج فتومتریکی در روش [۲۹]

۳-۴-۱- ساختن ماسک درون‌نگاری^۱

در این مرحله از روی تصویر سند ورودی یک ماسک ساخته می‌شود که در آن در نقاط تیره (قلم) مقدار صفر و پس‌زمینه مقدار یک را دارد. به این منظور مراحل زیر انجام می‌شود:

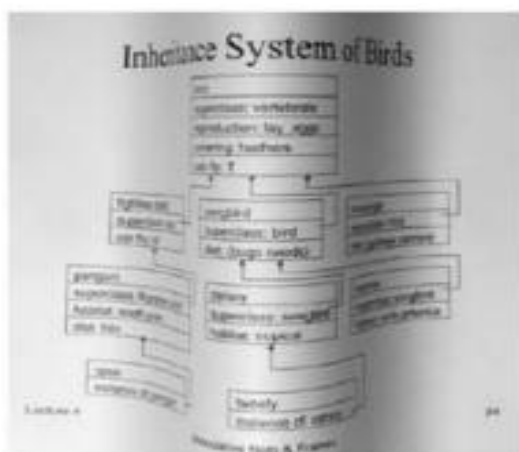
¹ Inpainting Mask Extraction

الف) استخراج کانال شدت روشنایی تصاویر

ب) مشخص کردن لبه‌ها به کمک روش آشکارساز کنی^۱

ج) انجام عملیات مورفولوژیک گسترش^۲ همراه با عملیات بستن^۳ (هدف از انجام این مرحله این است که تمام پیکسل‌های پیش‌زمینه پوشانده شود در ضمن المان ساختاری هم می‌تواند دستی انتخاب شود و هم می‌تواند به صورت خودکار بر اساس میانگین ارتفاع حروف انتخاب شود)

در شکل ۴-۵ نتیجه‌ی حاصل از این مرحله نشان داده شده است. همان‌طور که مشاهده می‌کنید ماسک تولیدی تمام حروف تصویر را پوشانده است.



الف

ب

شکل ۴-۵: نتیجه‌ی اعمال روش ژانگ [۲۹] به منظور تولید ماسک الف (تصویر ورودی

ب) ماسک تولیدی تصویر ۴-۵ الف

۴-۳-۲- حذف پیش‌زمینه^۴

در این مرحله با استفاده از الگوریتم‌های درون‌نگاری در محل پیکسل‌های قلم، روشنایی با مقدار پیکسل‌های پس‌زمینه همسایه جایگزین می‌شود. این مرحله برای به دست آوردن نتیجه‌ی بهتر چندین بار تکرار می‌شود.

¹ Canny

² Dilation

³ Closing

⁴ Foreground Removal

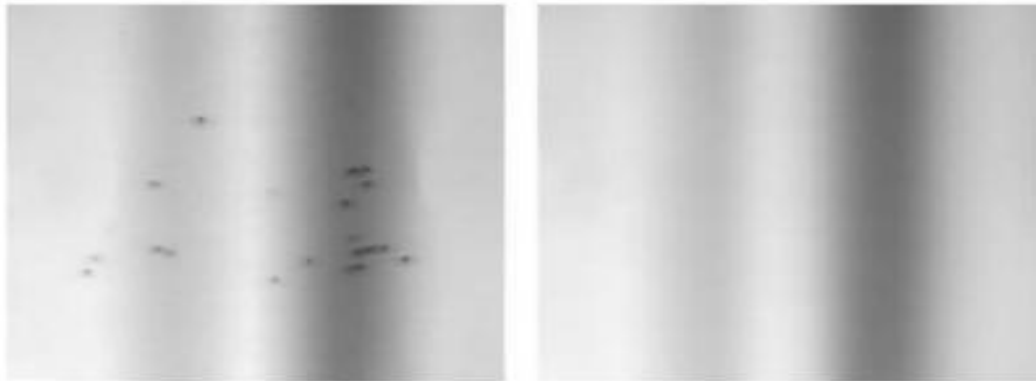
در این روش می‌توان از الگوریتم‌های درون‌نگاری هارمونیک^۱ و درون‌نگاری تغییرات کل^۲ که در [۴۹] اشاره شده است، استفاده کرد.

در روش ژانگ با استفاده از درون‌نگاری هارمونیک به رفع اعوجاج فتومتریک پرداخته می‌شود. همان‌طور که از فصل سوم به خاطر دارید رابطه‌ی درون‌نگاری هارمونیک به صورت رابطه‌ی ۴-۱ بیان شده است.

$$I_{i,j}^{n+1} = I_{i,j}^n + \Delta t \left(\frac{\lambda}{h^2} (I_{i+1,j}^n + I_{i-1,j}^n + I_{i,j+1}^n + I_{i,j-1}^n - 4I_{i,j}^n) - \chi_{i,j} (I_{i,j}^n - I_{0i,j}) \right) \quad (۱-۴)$$

۴-۳-۳ نرم کردن سایه‌ی استخراج شده^۳

در این مرحله می‌توان با استفاده از عملیات ملایم کردن به کمک روش RBF تصویر را نرم کرد [۵۸] همچنین به منظور بهبود عملکرد می‌توان ماسک تولیدی را بعد از هر تکرار دوباره ساخت. نتیجه‌ی اعمال این مرحله در شکل ۴-۶ ب قابل مشاهده می‌باشد.



الف

ب

شکل ۴-۶ نتیجه‌ی اعمال روش ژانگ [۲۹] به منظور حذف پیش‌زمینه و نرم کردن سایه‌ی استخراج شده شکل

۴-۵ الف) نتیجه‌ی حذف پیش‌زمینه به روش درون‌نگاری ب) نتیجه‌ی نرم کردن سایه به روش RBF

¹ Harmonic Inpainting

⁵ Total Variation(TV) Inpainting

³ Smooth Shading Extraction

۴-۳-۴ بازتاب تصویر استخراج شده^۱

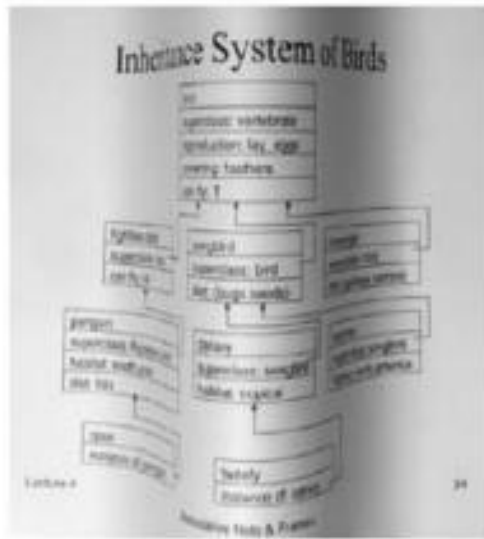
پس از محاسبه‌ی تابع روشنایی در لایه‌ی پس‌زمینه طبق رابطه‌ی ۲-۴ و رابطه‌ی ۳-۴ می‌توان تصویر لایه‌ی پیش‌زمینه را به دست آورد که در این روابط I_S تصویر پس‌زمینه و I_r تصویر پیش‌زمینه و I کانال شدت روشنایی در مدل HSV می‌باشد [۵۹].

$$I = I_r \cdot I_s \quad (۲-۴)$$

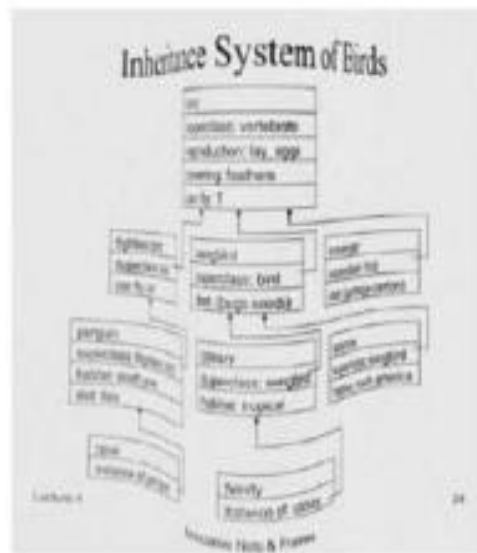
$$I_r = e^{\text{Log}I - \text{Log}I_s} \quad (۳-۴)$$

پس‌ازانجام این مرحله تصویر خروجی بدون اعوجاج فتومتریکی طبق رابطه‌ی ۴-۴ به دست خواهد آمد. نتیجه‌ی اعمال این مرحله در شکل ۷-۴ نشان داده‌شده است.

$$I_p = k \cdot I_r \quad k \in [0,1] \quad (۴-۴)$$



الف



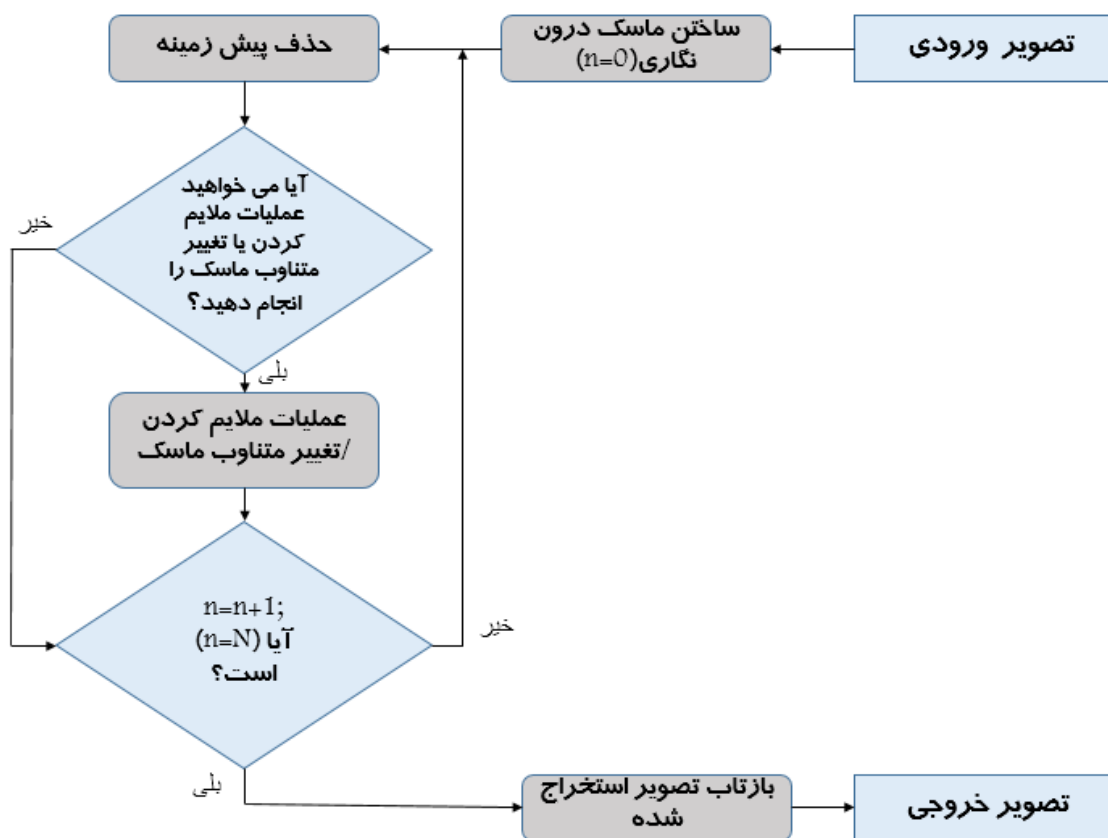
ب

شکل ۷-۴: نتیجه‌ی نهایی روش ژانگ [۲۹] (الف) تصویر ورودی (ب) تصویر نهایی بدون اعوجاج فتومتریکی

¹ Reflection Image Derivation

۴-۴- روش پیشنهادی به منظور رفع اعوجاج فتومتریکی

ابتدا فلوجارت روش پیشنهادی را رسم کرده و در نهایت به بررسی تغییرات انجام شده در این روش نسبت به روش ژانگ پرداخته می شود. پس از انجام این تغییرات سرعت اجرای الگوریتم تغییر محسوسی پیدا خواهد کرد و دقت روش پیشنهاد شده نیز قابل قبول خواهد بود. فلوجارت روش پیشنهادی را در شکل ۴-۸ مشاهده می کنید. این فلوجارت تا حدودی شبیه فلوجارت شکل ۴-۳ بوده اما در ادامه تغییرات پیشنهادی ارائه می گردند.

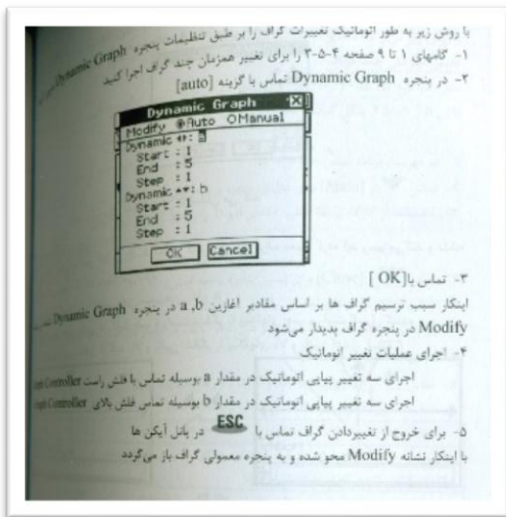


شکل ۴-۸: فلوجارت روش پیشنهادی

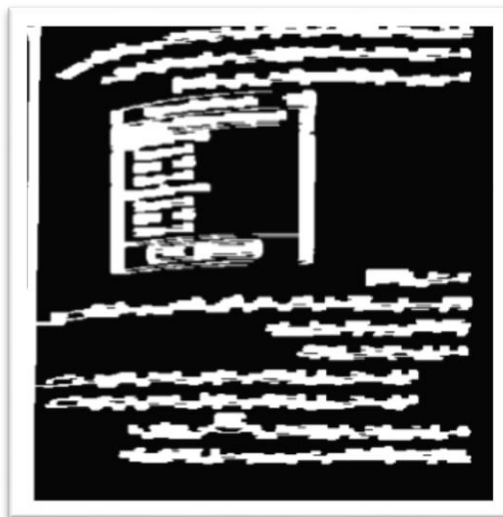
۴-۴-۱- ساختن ماسک درون نگاری

با توجه به فلوجارت ارائه شده در شکل ۴-۸ ابتدا از تصویر ورودی ماسک درون نگاری تولید می شود. نحوه تولید ماسک طبق روش ژانگ که در بخش ۴-۳-۱ اشاره شده است، می باشد. همان طور که در آن بخش اشاره شد با استفاده از کانال شدت روشنایی تصویر و استفاده از عملگر گسترش همراه با عملگر بستن این

ماسک تولید می‌شود. در شکل ۹-۴ نتیجه‌ی اعمال این مرحله بر روی یک نمونه از مجموعه تصاویر پایگاه داده معرفی شده در مرجع [۴۷] را مشاهده می‌کنید.



الف



ب

شکل ۹-۴: نتیجه‌ی تولید ماسک درون‌نگاری (الف) تصویر اولیه با اعوجاج فتومتریک (ب) ماسک روش درون‌نگاری

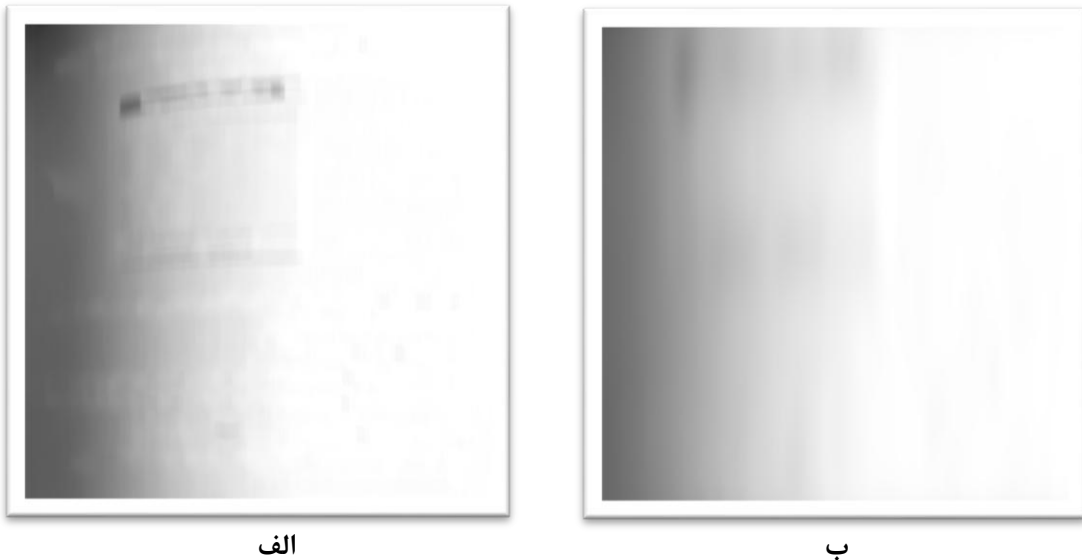
۴-۴-۲- حذف پیش‌زمینه با استفاده از درون‌نگاری هارمونیک بهبودیافته

با توجه به رابطه‌ی ۴-۱ که در بخش قبلی اشاره شده است، در این فرمول فقط از همسایگی‌های اصلی تصویر استفاده شده است که اگر ما بخواهیم این رابطه نسبت به چرخش حساس نباشد؛ می‌توان از همسایگی هشت‌تایی هر پیکسل از تصویر استفاده نماییم. البته اگر در این رابطه مستقیماً همسایگی هشت‌تایی را جاگذاری کنیم زمان اجرا تقریباً دو برابر می‌شود که بدین جهت روش مطلوبی نیست. به همین دلیل به پیشنهاد نموده‌ایم تا به صورت ترکیبی، از همسایه‌های چهارتایی اصلی و همسایه‌های چهارتایی قطری به صورت تناوبی در مراحل تکرار الگوریتم استفاده شود. بنابراین در هر تکرار و به صورت متناوب یک‌بار از همسایگی‌های اصلی طبق رابطه‌ی ۴-۱ و یک‌بار از همسایگی‌های قطری طبق رابطه‌ی ۴-۵ استفاده می‌کنیم.

$$I_{i,j}^{n+1} = I_{i,j}^n + \Delta t \left(\frac{\lambda}{h^2} (I_{i+1,j+1}^n + I_{i-1,j-1}^n + I_{i-1,j+1}^n + I_{i+1,j-1}^n - 4I_{i,j}^n) - \chi_{i,j} (I_{i,j}^n - I_{0i,j}) \right) \quad (5-4)$$

در این رابطه n نشان‌دهنده‌ی مرحله‌ی تکرار، h پارامتر اندازه‌ی مش^۱، λ پارامتر نرم‌کننده، Δt گام زمانی بوده که این پارامترها به‌صورت سعی و خطا انتخاب می‌شوند. همچنین λ نشان‌دهنده‌ی تابع مشخصه می‌باشد که در فصل سوم معرفی شده است.

بنابراین به‌منظور محاسبه‌ی مقدار درون‌نگاری شده‌ی هر پیکسل در مرحله‌ی $n+1$ ام، با توجه به مقادیر درون‌نگاری شده‌ی پیکسل‌ها در مرحله‌ی n ، مقدار پیکسل موردنظر تغییر پیدا می‌کند. نتیجه‌ی اعمال این مرحله را در شکل ۴-۱۰ مشاهده می‌کنید.



شکل ۴-۱۰: نتیجه‌ی روش درون‌نگاری هارمونیک بهبودیافته (الف) بعد از اولین تکرار (ب) بعد از ۳۰ تکرار

همچنین با توجه به این نکته که به‌منظور پیاده‌سازی روش اشاره‌شده نیاز به استفاده از سه حلقه‌ی تکرار تودرتو می‌باشد:

(۱) حلقه‌ی تکرار روش درون‌نگاری

(۲) حلقه‌ی جاروب پیکسل‌ها بر روی سطرها

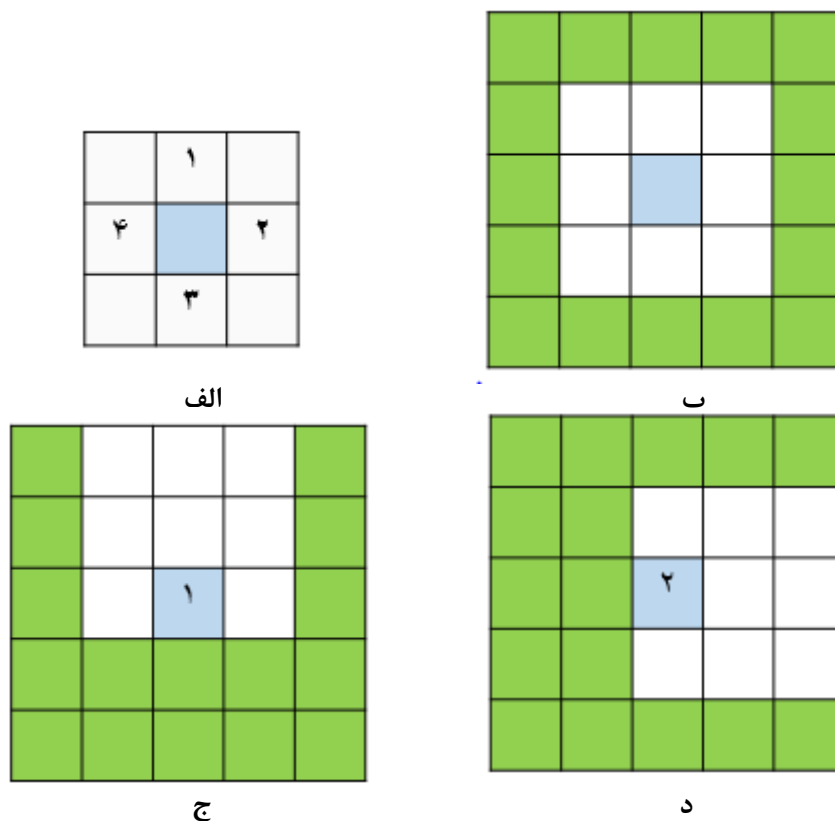
(۳) حلقه‌ی جاروب پیکسل‌ها بر روی ستون‌ها

و این حلقه‌ها باعث افزایش زمان اجرای برنامه می‌شود، می‌توان تعداد این حلقه‌های تودرتو را کاهش داد. با توجه به رابطه‌های ۴-۱ و ۴-۵ مشاهده می‌کنیم که در هر مرحله، هر پیکسل با چهار پیکسل همسایه‌ی خود جمع می‌شود.

بنابراین اگر سطر و ستون ماتریس I را از هر طرف یک واحد اضافه کرده و آن‌ها را با پیکسل همسایه پرکنیم (شکل ۴-۱۱ ب) و با استفاده از عملیات شیفت هر بار یکی از همسایه‌ها را به مرکز انتقال بدهیم

¹ Grid Size

(شکل ۴-۱۲ ج، د) و در نهایت ماتریس‌های حاصل را باهم جمع کنیم؛ نتیجه‌ی مطلوب حاصل می‌شود. با این روش دو حلقه‌ی تکرار ۱ و ۲ حذف‌شده و سرعت اجرای برنامه به‌صورت چشمگیری افزایش می‌یابد. در فصل پنجم سرعت این روش بررسی شده است. در شکل ۴-۱۱ نحوه‌ی این کار برای همسایه‌های چهارتایی اصلی به نمایش آمده است. برای همسایه‌های قطری نیز به‌صورت مشابه، ولی در هر مرحله با دو بار شیفت با جهت‌های مختلف صورت می‌پذیرد.



شکل ۴-۱۱: روش پیشنهادی جهت کاهش حلقه‌های تودرتو (الف) تصویر ورودی (ب) پر کردن سطر و ستون اضافه‌شده با پیکسل‌های همسایه (ج) شیفت به بالا (د) شیفت به راست

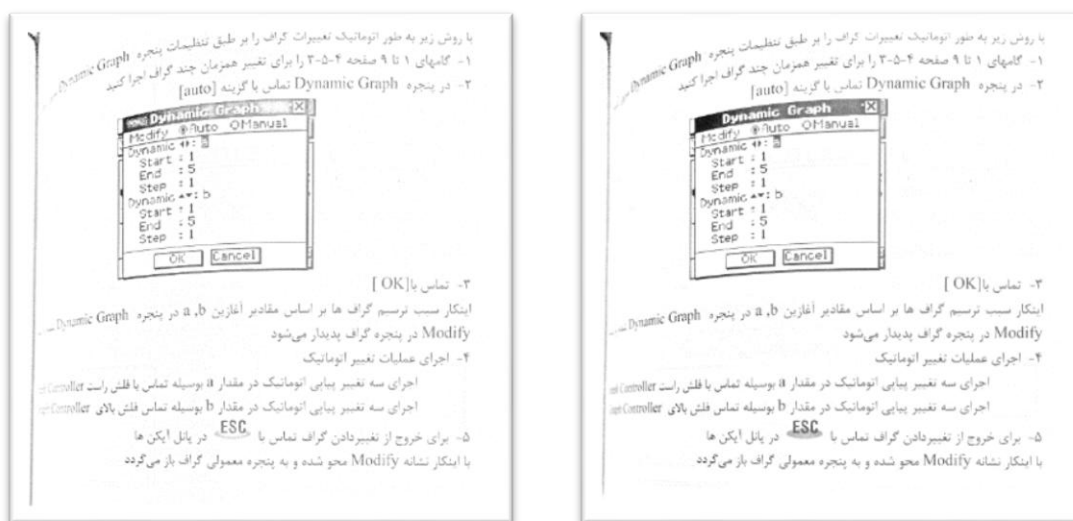
۴-۴-۳ - عملیات ملایم کردن / تغییر متناوب ماسک

با توجه به این نکته که روش RBF ذکرشده در روش ژانگ از لحاظ محاسباتی پیچیده و زمان‌بر است، مخصوصاً زمانی که تعداد پیکسل‌های تصویر زیاد باشد، چندین دقیقه زمان اجرا طول می‌کشد. به همین دلیل پیشنهاد می‌شود، از فیلترهای ملایم کننده گوسی یا میانگین‌گیر که خواسته‌ی ما را برآورده کرده و از لحاظ سرعت اجرا سریع‌تر می‌باشد، استفاده کرد. همچنین به‌منظور بهبود عملکرد و همین‌طور افزایش

سرعت اجرا می‌توان از کرنل‌های ستونی یا عمودی میانگین‌گیر یا گوسی استفاده کرد. به منظور کاهش پیچیدگی محاسبات این گام را می‌توان چند مرحله یکبار تکرار کرد.

۴-۴-۴ بازتاب تصویر استخراج شده

این مرحله نیز مشابه روش ژانگ بوده و پس از محاسبه‌ی تابع روشنایی در لایه‌ی پس‌زمینه طبق رابطه‌ی ۴-۲ و رابطه‌ی ۴-۳ می‌توان تصویر لایه‌ی پیش‌زمینه را به دست آورد و طبق رابطه‌ی ۴-۴ تصویر نهایی تولید می‌گردد. در شکل ۴-۱۲ یک نمونه از نتیجه‌ی نهایی روش پیشنهادی را مشاهده می‌کنید. همان‌طور که در این تصاویر مشاهده می‌کنید؛ تصاویر خروجی به‌صورت خاکستری می‌باشند.



الف

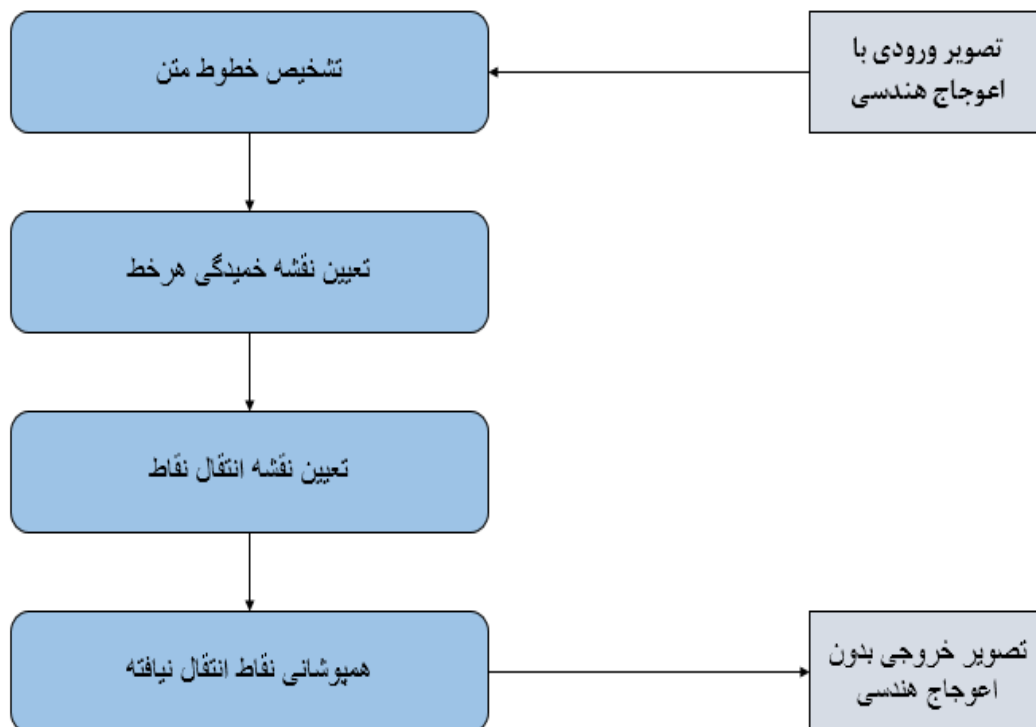
ب

شکل ۴-۱۲: نتیجه‌ی نهایی روش پیشنهادی (الف) به کمک ملایم‌کننده‌ی میانگین‌گیر 3×3 (ب) به کمک ملایم‌کننده‌ی میانگین‌گیر سطری به طول ۲۰ پیکسل

۴-۵-۴ اعوجاج هندسی

به‌منظور رفع اعوجاج هندسی، تنها به اطلاعات دوبعدی موجود در سند اکتفا کرده و به هیچ نوع اطلاعاتی در مورد مشخصات دوربین دیجیتال یا اسکنر و طریقه‌ی قرار گرفتن سند، نیازی نداریم. روش پیشنهادی مبتنی بر استخراج خطوط می‌باشد به‌گونه‌ای که انحنای هر خط مشخص‌شده و پس از رفع اعوجاج هندسی خط اول این کار برای سایر خطوط نیز تکرار می‌شود.

در شکل ۴-۱۳ فلوجارت کلی روش پیشنهادی به منظور رفع اعوجاج هندسی را مشاهده می‌کنید. طبق این فلوجارت در مرحله‌ی اول خطوط متن استخراج می‌شوند؛ در مرحله‌ی بعدی نقشه‌ی خمیدگی هر خط مشخص می‌شود؛ در مرحله‌ی سوم روشی جهت انتقال نقاط انجام شده و در مرحله‌ی آخر نقاط انتقال نیافته بر اساس نقاط همسایه تخمین زده می‌شود.



شکل ۴-۱۳: فلوجارت روش پیشنهادی رفع اعوجاج هندسی

۴-۵-۱- تشخیص خطوط متن

روش‌های زیادی به منظور تشخیص خطوط وجود دارد [۶۰ - ۶۳]. یکی از روش‌هایی که هم از لحاظ سرعت و هم از لحاظ دقت نتیجه‌ی قابل قبولی دارد و همچنین پیچیدگی محاسباتی کمی دارد؛ روشی است که در آن، از عملگر ریخت‌شناسی گسترش استفاده شده است.

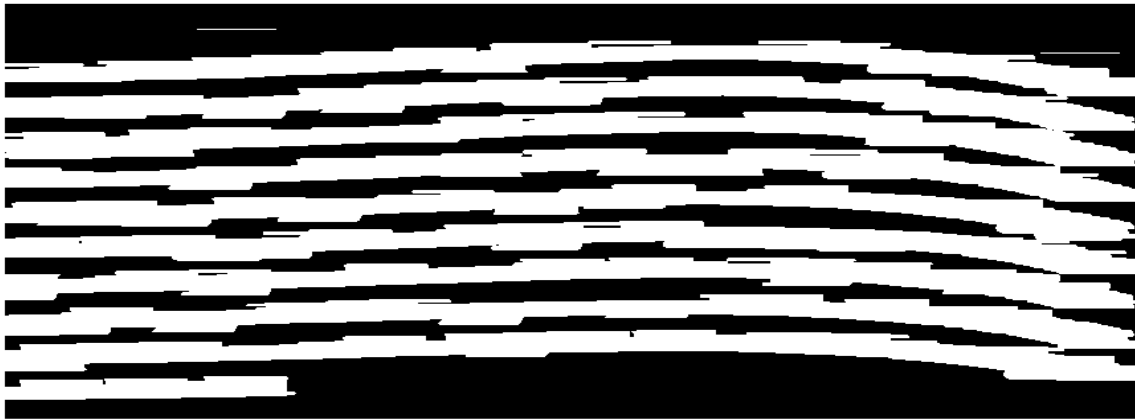
تصویر ورودی این روش می‌بایست بدون نوفه و به صورت دوسطحی باشد. الگوریتم کلی به صورت زیر بیان می‌شود:

- ۱- قرینه‌ی تصویر ورودی محاسبه می‌شود.
 - ۲- عملگر ریخت‌شناسی گسترش با یک جزء ساختاری مناسب به تصویر اعمال می‌گردد.
 - ۳- نقاط به هم چسبیده توسط روش اجزای به هم متصل تشخیص داده می‌شوند.
- با انجام دادن این مراحل، ماسک موردنیاز به‌منظور تشخیص دادن خطوط تولید می‌شود. در شکل ۴-۱۴ چند نمونه ماسک تولیدی توسط این روش را مشاهده می‌کنید.
- با مشاهده این تصاویر به دو نکته‌ی مهم پی می‌بریم:

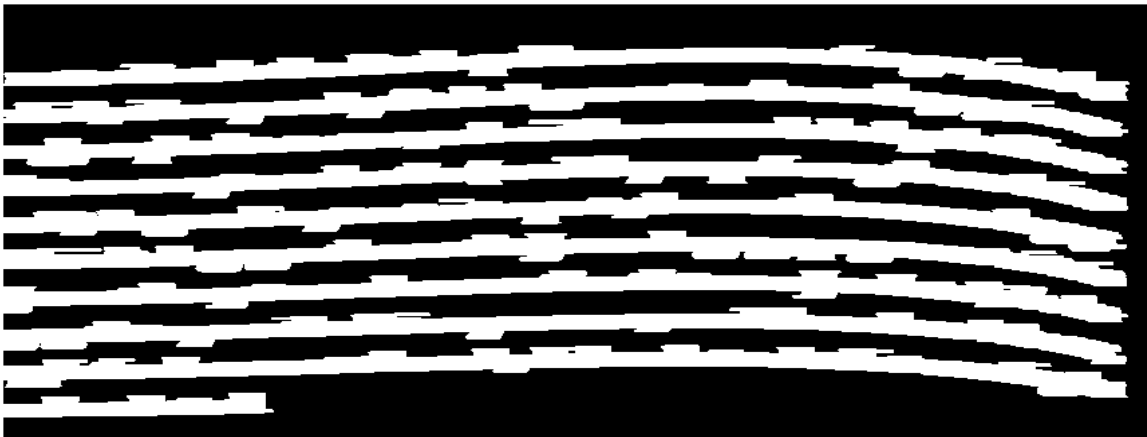
- ۱- انتخاب مناسب جزء ساختاری
 - ۲- حذف نقاط به‌هم‌پیوسته‌ی خیلی کوچک
- همان‌طور که در شکل ۴-۱۴ ب مشاهده می‌شود اگر جزء ساختاری بیش‌ازاندازه بزرگ باشد باعث می‌شود که در ماسک تولیدی خطوط به هم چسبیده و درنهایت این ماسک غیرقابل استفاده شود؛ به همین دلیل انتخاب اندازه‌ی جزء ساختاری بسیار مهم می‌باشد. این اندازه بسته به نوع پایگاه داده مورد استفاده متفاوت می‌باشد و به‌صورت تجربی برای هر یک از پایگاه‌های داده محاسبه می‌شوند.

nicate with each other. But the problems in communication grow, instead of shrink, with every new day of technology innovation. As more languages, increasingly complex data structures, and varied business needs develop, the difference between one system and another (even if they are performing the same business task!) also grows. For systems to interoperate, there must be a common language. This isn't a language in the sense that Java is a language; programs not written in Java must have just as much access as those that are in Java. Instead, it's a language that anyone with a dictionary can access; even if the words aren't the same, applications have access to the words, which can be translated into something that *can* be understood.

الف



ب



ج

شکل ۴-۱۴: ماسک تولیدی توسط عملگر گسترش و مشکلات رایج آن الف) تصویر ورودی ب) ماسک با جزء ساختاری نامناسب ج) ماسک با جزء ساختاری مناسب

همچنین با محاسبه‌ی طول اجزای به‌هم‌پیوسته و حذف اجزای کوچک مشکل بعدی هم حل خواهد شد. پس از رفع این مشکلات به تصویر ۴-۱۴ ج رسیده که نسبتاً مناسب می‌باشد؛ اما مشکل دیگری که به‌خصوص در اسناد فارسی با آن مواجه می‌شویم؛ مربوط به حذف خطوط کوچک و همچنین وجود حفره‌ها در ماسک تولیدی می‌باشد که این امر باعث حذف اجزایی از خط مثل نقاط، سرکش حروف و آکلاد می‌شود. راه‌حل این مشکل استفاده از عملگر بستن پس از انجام عملیات گسترش می‌باشد که البته طول جزء ساختاری برای عملیات بستن کوچک‌تر از طول جزء ساختاری عملیات گسترش بوده و می‌تواند این جزء ساختاری مستطیلی شکل با عرض کوچک نیز باشد.

این مشکل را در تصویر ۴-۱۵ ب مشاهده می‌کنید. در این تصویر، نقاط کلماتی مثل کشیده (نقاط مربوط به حرف ی) و شنیدم (نقاط مربوط به حرف ش) حذف شده‌اند؛ با توجه به روش پیشنهادی که نتیجه‌ی آن در شکل ۴-۱۵ ج نشان داده شده است، این مشکل نیز برطرف شده است.

اما اگر در این روش طول جزء ساختاری نامناسب باشد باعث پیوسته شدن ماسک خطوط و درنهایت تشخیص اشتباه جزءهای به‌هم‌پیوسته می‌شود؛ بنابراین این مشکل نیز باید برطرف شود.

به همین منظور و با توجه به این نکته که ما از ماسک تولیدی به‌صورت خط به خط استفاده می‌کنیم؛ یعنی ابتدا خط اول سپس خط دوم و به همین ترتیب تا خط آخر. می‌توان به‌جای استفاده کردن از عملگر بستن از عملگر گسترش با جزء ساختاری مستطیلی شکل با عرض و طول کوچک بر روی ماسک هر خط استفاده کرد؛ که در این روش مشکلی بابت پیوسته شدن ماسک‌های خطوط نداریم زیرا ماسک هر خط به‌صورت جدا بررسی می‌شود.

در شکل ۴-۱۶ ماسک تولیدی تصویر ۴-۱۵ الف را مشاهده می‌کنید. به‌منظور نمایش بهتر ماسک مربوط به هر خط با رنگی متفاوت نشان داده شده است. در شکل ۴-۱۷ نیز نمونه‌ای دیگر از ماسک تولیدی نشان داده شده است.

بنابراین به‌منظور استخراج خطوط به‌صورت مناسب ابتدا توسط عملگر گسترش اجزای هر خط به هم متصل شده و با حذف خطوط نازک و استفاده از روش اجزای به‌هم‌پیوسته ماسک هر خط مشخص می‌شود. درنهایت به‌منظور پوشش تمام حروف، یک‌بار دیگر ماسک هر خط با عملگر گسترش افزایش پیدا می‌کند.

نور کم‌رنگی که از بیرون به داخل گودال می‌تابید، به سرعت در تاریکی محو شده بود. صدای پاهای لرزان من، در فضای تاریک طنین‌انداز بود؛ به نظر می‌رسید مسیر ما جانی عریضی باشد، اما البته در این مورد مطمئن نبودم. صدایی به جز تپش دیوانه‌وار قلبم و کشیده شدن پاهایم روی زمین خیس، نمی‌شنیدم. فقط یک بار، صدای آهسی حاکی از بی‌قراری کسی را از پشت سر شنیدم.

الف

نور کم‌رنگی که از بیرون به داخل گودال می‌تابید، به سرعت در تاریکی محو شده بود. صدای پاهای لرزان من، در فضای تاریک طنین‌انداز بود؛ به نظر می‌رسید مسیر ما جانی عریضی باشد، اما البته در این مورد مطمئن نبودم. صدایی به جز تپش دیوانه‌وار قلبم و کشیده شدن پاهایم روی زمین خیس، نمی‌شنیدم. فقط یک بار، صدای آهسی حاکی از بی‌قراری کسی را از پشت سر شنیدم.

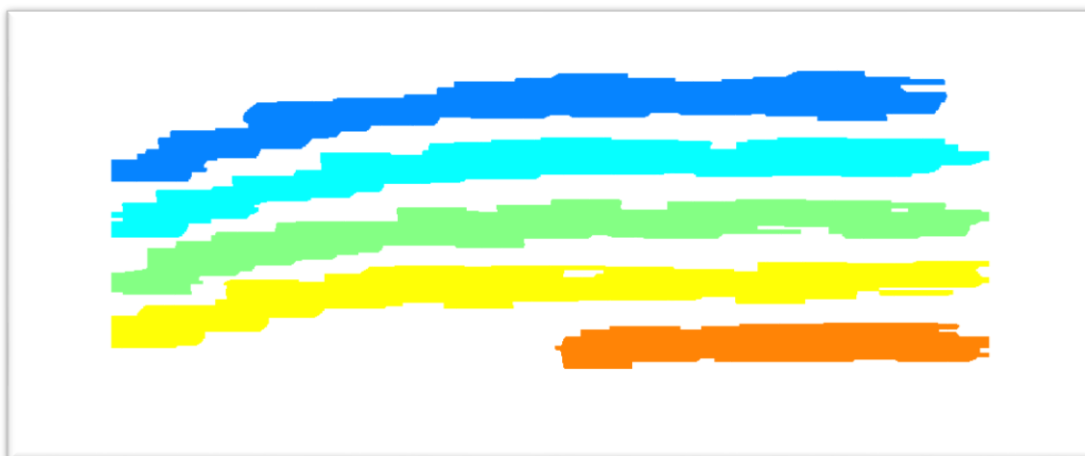
ب

نور کم‌رنگی که از بیرون به داخل گودال می‌تابید، به سرعت در تاریکی محو شده بود. صدای پاهای لرزان من، در فضای تاریک طنین‌انداز بود؛ به نظر می‌رسید مسیر ما جانی عریضی باشد، اما البته در این مورد مطمئن نبودم. صدایی به جز تپش دیوانه‌وار قلبم و کشیده شدن پاهایم روی زمین خیس، نمی‌شنیدم. فقط یک بار، صدای آهسی حاکی از بی‌قراری کسی را از پشت سر شنیدم.

ج

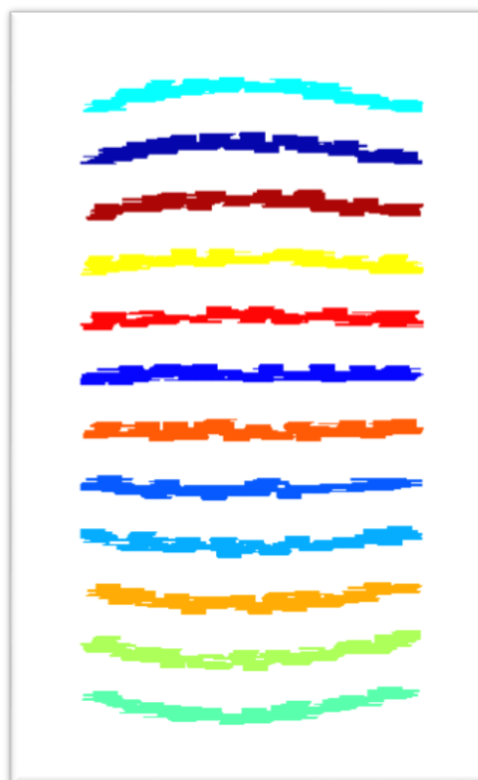
شکل ۴-۱۵: حذف نقاط از تصویر الف) تصویر ورودی ب) به اشتباه برخی از نقاط یا قسمتی از حروف

حذف‌شده‌اند ج) تصویر اصلاح‌شده با روش پیشنهادی



شکل ۴-۱۶: ماسک تولیدی تصویر ۴-۱۵ الف

بخت بردن می‌خوانم مجاذبه‌ی همان دنیای دیگری که بیشتر
 یاد کردم. شاید بتوان اسم این کار را قرار از واقعیت
 گزشت. آدم حتی اگر بتواند در عالم خیال از غشه‌های
 زندگی روزمره فرار کند و زمانی را در دنیای دیگر
 بگذراند خوب است. مین شانزده تا بیست و شش سالگی.
 خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم
 را بسازم. برای خودم کسی بشوم. آگاهی‌هایم را بیشتر کنم
 و بدین‌ترتیب به روحم شکل بدهم. در واقع، می‌خواستم
 بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا
 چیست؟ چقدر می‌توانم فکرم را، علاقم را، رویاها را و
 انتظاری را که در ذهن داشتم گسترش دهم؟ وقتی زندگی،
 رویاها و تأملات دیگران را در داستان‌ها یا نوشته‌ها و



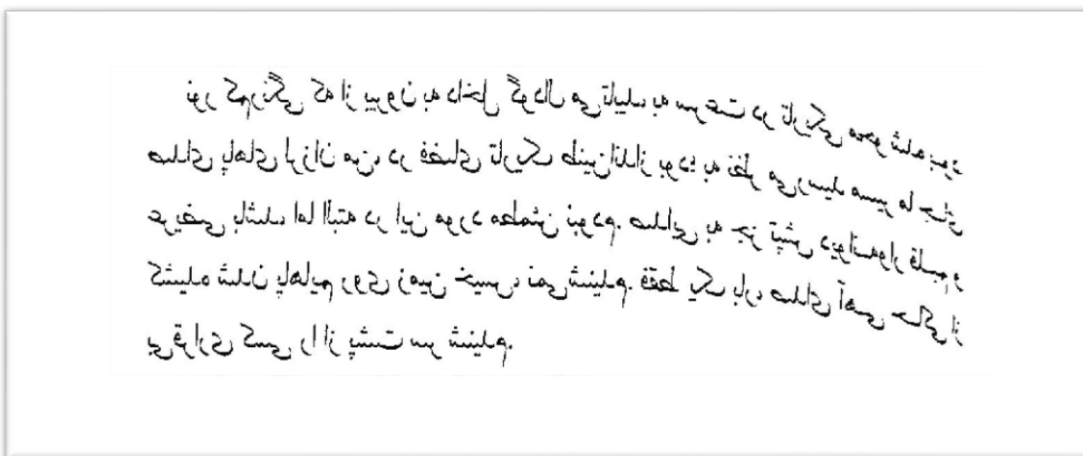
الف

ب

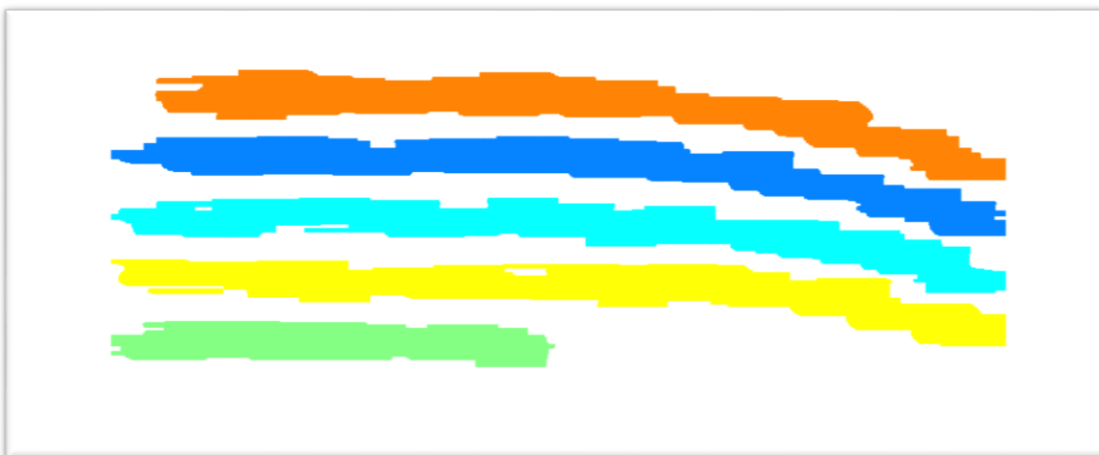
شکل ۴-۱۷: نمونه‌ای دیگر از نتیجه‌ی تولید ماسک الف (تصویر ورودی ب) ماسک تولیدی

همان طور که می‌دانید، حروف فارسی برخلاف حروف انگلیسی از سمت راست تصویر نوشته می‌شوند؛ و با توجه به این نکته که سایر پردازش‌هایی که در ادامه معرفی می‌شوند به نقطه‌ی شروع خط وابسته می‌باشند، برای متون فارسی از روش آئینه کردن افقی استفاده می‌شود و کلیدی مراحل برای اسناد فارسی بر روی تصویر آئینه شده انجام می‌پذیرد.

در شکل ۴-۱۸ نمونه‌ای از انجام عملیات آئینه کردن متن فارسی و ماسک تولیدی آن را مشاهده می‌کنید.



الف



ب

شکل ۴-۱۸: آئینه کردن افقی و تولید ماسک الف) آئینه‌ی افقی تصویر ۴-۱۵ الف ب) ماسک تولیدی

۴-۵-۲- تشخیص نقشه‌ی خمیدگی سند

یکی از اصلی‌ترین مرحله‌های رفع اعوجاج هندسی مرحله‌ی تشخیص نقشه‌ی خمیدگی سند می‌باشد. در این مرحله هدف مشخص کردن انحنای هر خط می‌باشد.

این مرحله خود از چند گام اساسی تشکیل شده است که در زیر مشاهده می‌کنید:

۱- تقسیم هر خط به نوارهای عمودی

۲- افکنش افقی با تابع گرادیان و یافتن مختصات بیشترین مقدار

۳- گذراندن منحنی درجه‌ی سه از نقاط مشخص شده

در ادامه این بخش به معرفی هر کدام از این گام‌ها می‌پردازیم.

۴-۵-۲-۱- تقسیم هر خط به نوارهای عمودی

در این گام هر خط استخراج شده را به نوارهایی بافاصله‌ی مساوی تقسیم می‌کنیم. این نوارها با شروع از سمت چپ تصویر آغاز شده و تا انتها ادامه پیدا می‌کند. اندازه‌ی عرض نوار بر روی هر مجموعه داده متفاوت می‌باشد، ولی به صورت تجربی مقداری بین ۱۵۰ تا ۲۵۰ پیکسل بر روی پایگاه‌های داده‌ی مورد استفاده، مناسب می‌باشد. همچنین با توجه به تعداد ستون‌های تصویر، عرض آخرین نوار ممکن است از بقیه نوارها کوچک‌تر باشد.

در شکل ۴-۱۹ دو نمونه از این تقسیم‌بندی نشان داده شده است.

between one system and another (even if they are performing the same business

الف

بین یک سیستم و سیستم دیگری (حتی اگر آنها در حال انجام یک کار مشابه باشند)

ب

شکل ۴-۱۹: تقسیم خطوط به نوارهای افقی الف) تقسیم خط چهارم شکل ۴-۹ الف به نوارهای افقی

ب) تقسیم خط دوم شکل ۴-۱۰ الف به نوارهای افقی

۴-۵-۲-۲- افکنش افقی با تابع گرادیان و یافتن مختصات بیشترین مقدار

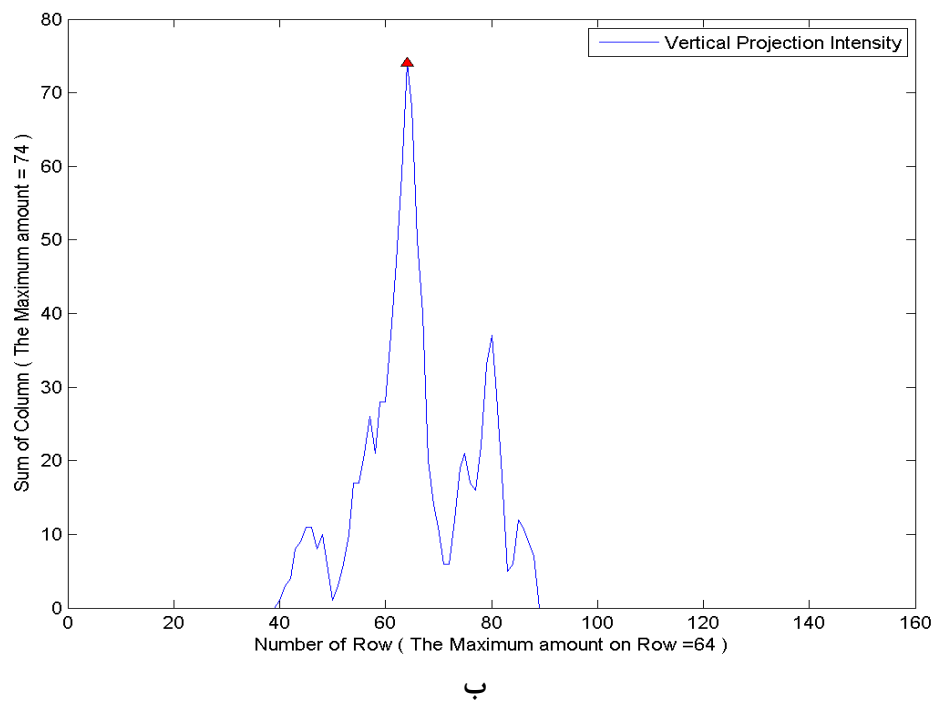
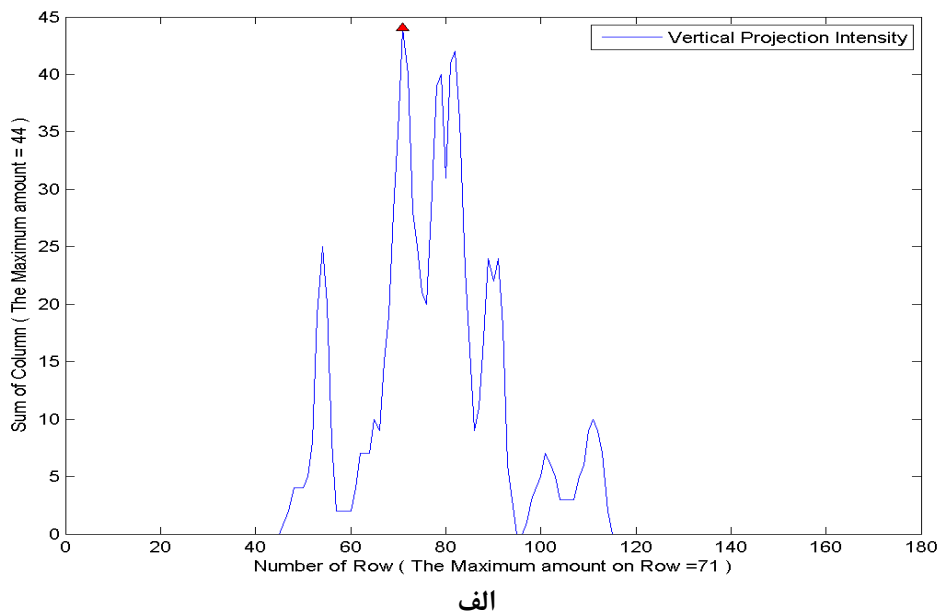
به منظور پیدا کردن نقاط جهت تخمین انحنای هر خط از بیشینه مقدار افکنش افقی با تابع گرادیان در راستای γ استفاده می‌کنیم. برای حصول نتیجه‌ی بهتر ابتدا هر نوار در بازه‌ی α تا $\alpha +$ چرخانده شده و زاویه‌ای که طبق آن حروف نوار تا حد امکان صاف باشد را به دست می‌آوریم. زاویه‌ی α محدوده‌ی مورد انتظار برای چرخش نوار می‌باشد.

با توجه به این نکته که برای نگارش حروف فارسی از یک خط پایه به‌عنوان مبنا استفاده می‌شود؛ بنابراین می‌توان نتیجه گرفت که در اثر چرخش نوار با زاویه‌های مختلف، زاویه‌ای که طبق آن بیشینه مقدار نمودار افکنش افقی با تابع شدت روشنایی از بقیه بزرگ‌تر باشد، زاویه‌ای است که طبق آن متن نوشته‌شده داخل نوار در صاف‌ترین حالت ممکن خود می‌باشد و مختصات این نقطه همان خط پایه می‌باشد؛ اما در اسناد انگلیسی با توجه به این نکته که حروف غالباً بین دو خط نوشته می‌شوند؛ بر اثر چرخش زمانی نوار در صاف‌ترین حالت قرار می‌گیرد که مختصات بیشینه‌ی نمودار افکنش افقی بر روی یکی از دو خطوط پایه قرار گیرد.

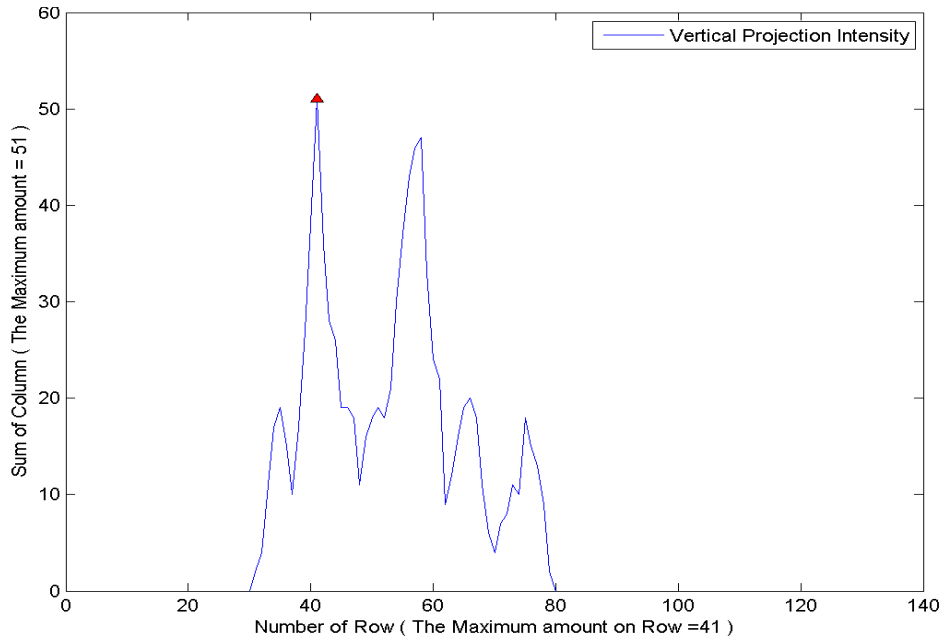
به منظور درک بهتر در شکل ۴-۲۰ تا ۴-۲۲ چند نمونه از چرخش‌های یک نوار و همچنین نمودار افکنش افقی نوار چرخانده شده نشان داده شده است. با مشاهده‌ی شکل‌های ۴-۲۰ متوجه می‌شویم که زاویه چرخش ۸- بهترین انتخاب می‌باشد؛ با توجه به بیشینه مقدار نمودارهای افکنش نیز به همین نتیجه خواهیم رسید؛ زیرا این مقدار برای زاویه‌ی ۸- برابر با ۷۴ بوده در حالیکه برای سایر زاویه‌ها از ۷۴ کوچک‌تر می‌باشد.



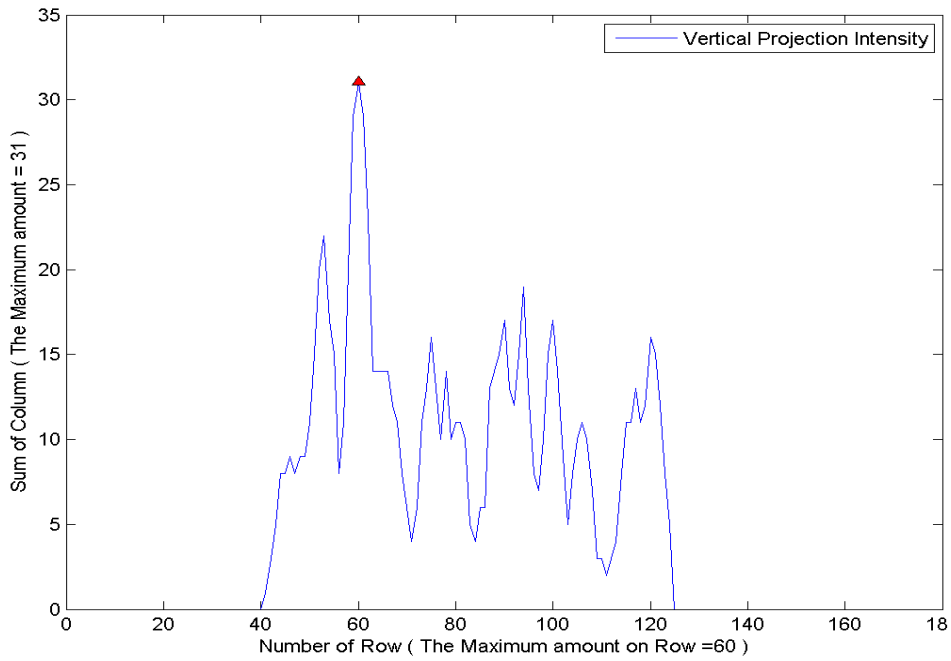
شکل ۴-۲۰: چرخش نوار بین زاویه‌های ۱۶- تا ۱۶ درجه الف) چرخش با زاویه ۱۶- درجه ب) چرخش با زاویه ۱۲- درجه پ) چرخش با زاویه ۸- درجه ت) چرخش با زاویه ۴- درجه ج) چرخش با زاویه ۰ درجه چ) چرخش با زاویه ۴ درجه ح) چرخش با زاویه ۸ درجه خ) چرخش با زاویه ۱۶ درجه



شکل ۴-۲۱: نمودار افکنش افقی الف) زاویه‌ی چرخش ۱۶- ب) زاویه‌ی چرخش ۸-



الف



ب

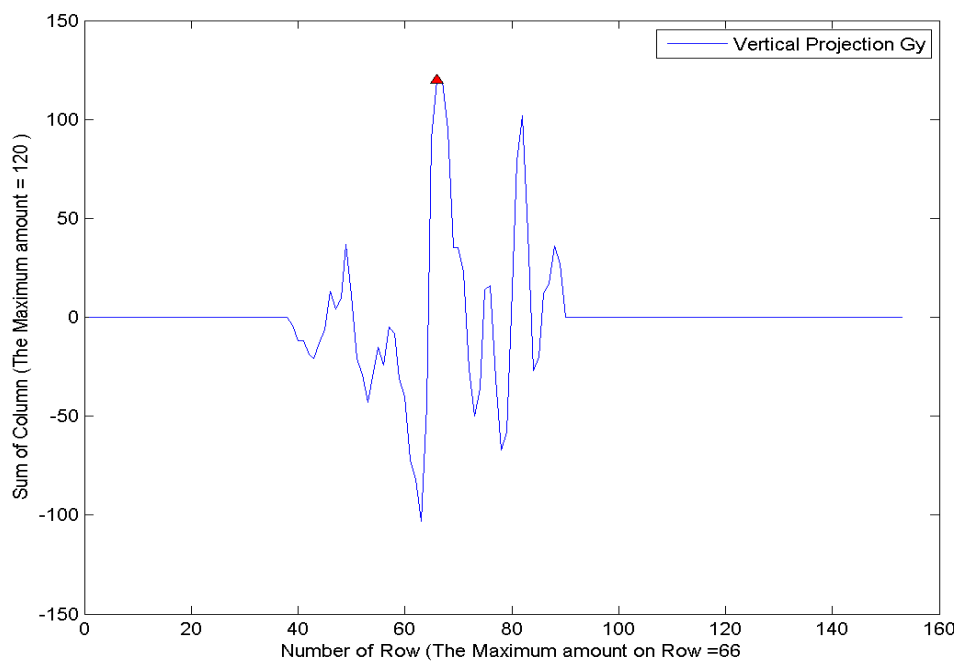
شکل ۴-۲۲: نمودار افکنش افقی الف) زاویه‌ی چرخش ۰ ب) زاویه‌ی چرخش ۱۶

اما مشکل اصلی این روش زمانی است که حروفی مثل ر، ز، و پشت سر هم قرار گرفته و یا تعداد حروف کمی در نوار قرار گیرد. در این صورت در تعیین زاویه‌ی دقیق با مشکل مواجه می‌شویم؛ که البته این مسئله را به‌عنوان خطا فعلاً قبول می‌کنیم.

پس از به دست آوردن بهترین زاویه، از تصویر چرخانده شده؛ گرادیان در راستای γ گرفته و این بار عملیات افکنش افقی را بر روی گرادیان تصویر انجام می‌دهیم و مختصاتی که نمودار افکنش در آن بیشینه مقدار را دارد، ذخیره می‌کنیم. در شکل ۴-۲۳ نتیجه‌ی این عملیات را مشاهده می‌کنید. بیشینه مقدار نمودار افکنش افقی در سطر ۶۶ ام اتفاق افتاده است، بنابراین این نقطه، نقطه‌ی مطلوب می‌باشد.



الف



ب

شکل ۴-۲۳: نمودار افکنش افقی Gy الف: گرادیان تصویر (ب) نمودار افکنش افقی

اما نکته‌ای که در این بین مهم می‌باشد این است که بر اثر چرخش اندازه‌ی تصویر تغییر پیدا می‌کند و تعداد پیکسل‌ها هم در راستای ستون‌ها و هم در راستای سطرها افزایش می‌یابد. این امر باعث می‌شود که مختصات به‌دست‌آمده در تصویر اصلی صدق نکند؛ بنابراین باید با توجه به زاویه‌ی چرخش مختصات این نقاط را به مختصات در نوار اصلی تبدیل کنیم. به همین منظور کارهای زیر پیشنهاد می‌شوند:

۱- جهت چرخش (علامت زاویه چرخش) تعیین شود.

۲- اندازه‌ی تعداد سطرها و تعداد ستون‌های نوار اصلی و نوار چرخانده شده تعیین شود.

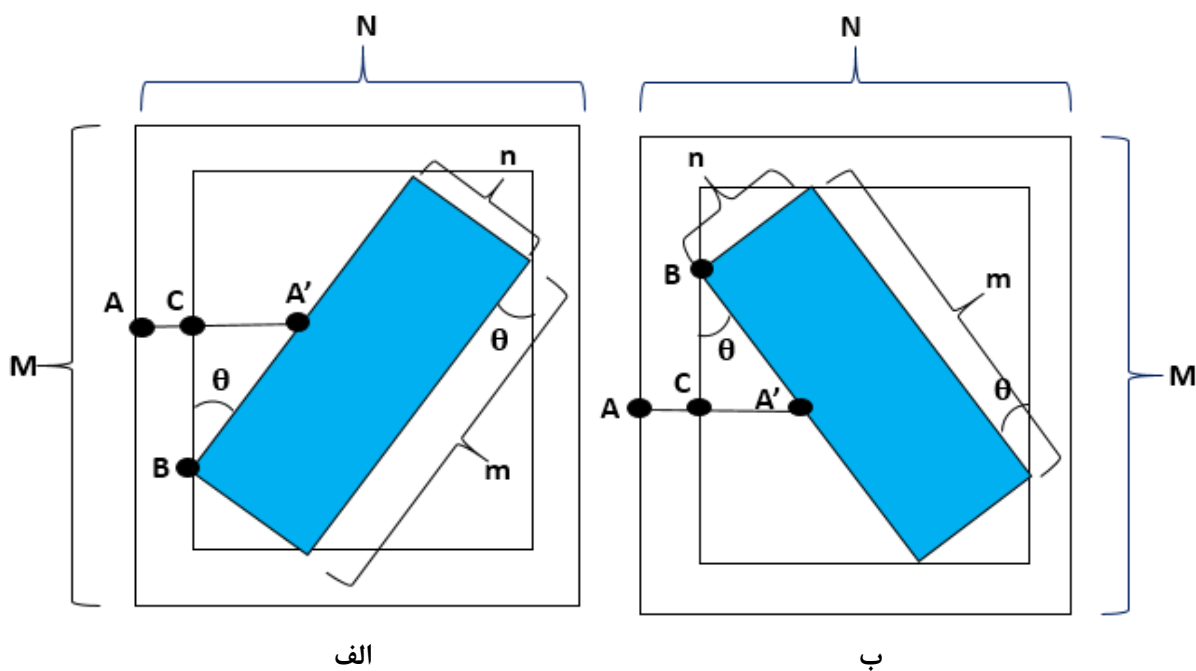
۳- بر اساس زاویه چرخش نقطه مذکور محاسبه شود.

با توجه به شکل ۴-۲۴، زاویه چرخش را با θ و تعداد سطرها و تعداد ستون‌های نوار اصلی را به ترتیب با m و n و تعداد سطرها و تعداد ستون‌های نوار چرخانده شده را به ترتیب با M و N نشان می‌دهیم ($N \geq n$ و $M \geq m$). هدف پیدا کردن تصویر نقطه A بر روی نوار چرخانده شده (A') و یافتن فاصله‌ی $|A'B'| = |BA'|$ می‌باشد.

برای این منظور یک ماتریس با ابعاد نوار اصلی فقط با مقادیر یک ایجاد کرده و به اندازه‌ی θ چرخانده می‌شود. بر اثر چرخش تعدادی سطر و ستون با پیکسل‌های مشکلی به تصویر اضافه خواهد. از سمت چپ تصویر شروع کرده و به ازای ستون‌های مختلف اولین سطری که دارای پیکسل سفیدرنگ است را مشخص می‌کنیم و آن نقطه را نقطه‌ی مرجع (B) نام‌گذاری می‌کنیم. فاصله‌ی عمودی نقطه‌ی A تا نقطه‌ی B را محاسبه کرده و نقطه‌ی C را با این فاصله از B ایجاد می‌کنیم. در این صورت مثلث $A'CB$ ، مثلثی قائم‌الزاویه بوده و اندازه‌ی وتر آن ($|A'B'| = |BA'|$) طبق رابطه‌ی ۴-۱۰ به دست می‌آید.

$$|A'B'| = \frac{|BC|}{\cos(\theta)} \quad (1-5)$$

و با توجه به این نکته که نقطه‌ی B برای چرخش به سمت راست در پایین صفحه و برای چرخش به سمت چپ در بالای صفحه در قرار دارد، مختصات نقطه‌ی جدید برای چرخش به سمت راست برابر با $(X(A), m - |A'B'|)$ و برای چرخش به سمت چپ برابر با $(X(A), |A'B'|)$ می‌باشد.



شکل ۴-۲۴: تغییر اندازه‌ی تصویر بر اثر چرخش (الف) چرخش به سمت راست ($\theta < 0$)
 (ب) چرخش به سمت چپ ($\theta > 0$)

بنابراین ما برای هر نوار می‌بایست مراحل ذکر شده را انجام داده تا به نقاط جهت تخمین انحنای هر خط برسیم. در شکل ۴-۲۵ نتیجه را مشاهده می‌کنید. به‌عنوان آخرین نکته در این بخش، به‌منظور به دست آوردن دقیق‌تر انحنای خطوط لازم است که نقطه‌ی انتهایی برای تشخیص انحنای نوار را نیز به دست بیاوریم؛ اما با توجه به تقسیم‌بندی نوارها که در شکل ۴-۲۵ الف نیز مشاهده می‌کنید؛ آخرین نقطه محاسبه‌شده تا انتهای خط فاصله‌ی قابل توجهی دارد.

به همین منظور پیشنهاد می‌شود یک نوار با طول برابر با سایر نوارها از سمت راست تصویر جدا کرده و تصویر آینه شده در آن قرار بگیرد و سایر کارها طبق روال گذشته انجام شود و نقطه‌ی نوار انتهایی در آخرین ستون به دست بیاید. نتیجه‌ی این کار را در شکل ۴-۲۵ ب مشاهده می‌کنید.

between one system and another (even if they are performing the same business

الف

between one system and another (even if they are performing the same business

ب

شکل ۴-۲۵: پیدا کردن نقاط جهت تشخیص انحنای هر خط الف) نتیجه‌ی پیدا کردن نقاط جهت تشخیص انحنای ب) نتیجه‌ی روش پیشنهادی جهت یافتن نقطه‌ی پایانی

۴-۵-۲-۳- گذراندن منحنی درجه‌ی سه از نقاط مشخص شده

برای تخمین خمیدگی هر خط از تقریب چند جمله‌ای درجه سوم استفاده می‌کنیم. اگر چند جمله‌ای درجه سوم را با رابطه‌ی ۴-۲ نشان دهیم و مقدار آن در نقطه‌ی x_i با y_i نشان دهیم؛ این تقریب طبق رابطه‌ی ۴-۳ محاسبه می‌شود.

$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0 \quad (۲-۵)$$

$$p(x) = \sum_{i=0}^n y_i \cdot \prod_{0 \leq j \leq n, j \neq i} \frac{x - x_j}{x_i - x_j} \quad (۳-۵)$$

همانطور که در رابطه ۴-۳ مشاهده می‌کنید، با توجه به تعداد داده‌های ورودی این تقریب زمان بر خواهد بود. به همین دلیل از ماتریس وندرموند^۱ [۶۴] که دترمینان آن برابر مخرج کسر رابطه‌ی ۴-۳ بوده، استفاده می‌کنیم. این ماتریس در رابطه‌ی ۴-۴ نشان داده شده است.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix} \quad (۴-۵)$$

¹ Vandermonde Matrix

به منظور تقریب به کمک ماتریس وندرموند به دستگاه معادلات $VA=Y$ می‌رسیم که در آن V ماتریس وندرموند، A برابر ماتریس ضرایب چند جمله‌ای درجه سوم و Y ماتریس مقادیر تابع چندجمله‌ای بوده که در رابطه‌ی ۴-۵ مشاهده می‌کنید.

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (5-5)$$

با توجه به مشخص بودن مقادیر ماتریس V و Y مقدار ماتریس A محاسبه می‌شود.

در شکل ۴-۲۶ چند نمونه از منحنی‌های تخمین زده شده به وسیله این روش نشان داده شده است. همان‌طور که در این شکل‌ها مشاهده می‌کنید، برای متون انگلیسی انحنا بر روی خط پایه پایینی و برای متون فارسی این انحنا بر روی خط پایه واقع می‌شود.

~~between one system and another (even if they are performing the same business.~~

الف

~~all because of web services. Of course I've skimmed over all the hard parts, but I'll~~

ب

~~مجموعه‌ای از سیستم‌ها را می‌بینیم که در حال انجام یک کار هستند و این کارها را با استفاده از سیستم‌های مختلف انجام می‌دهند.~~

ج

شکل ۴-۲۶: چند نمونه از تخمین انحنا (خطوط الف و ب) تشخیص انحنا (خطوط انگلیسی ج) تشخیص انحنا (خط فارسی) (از آیین‌های خط فارسی استفاده شده است).

اما نکته‌ی مهمی که در این قسمت قابل ذکر است این است که اگر طول هر خط از یک حدی کم‌تر باشد، در نتیجه تعداد نقاط کمی جهت گذراندن منحنی درجه سه ایجاد می‌شود که این مسئله به تشخیص اشتباه انحنای خطوط کوتاه منجر می‌شود.

به همین دلیل به منظور رفع این مشکل پیشنهاد می‌شود که در مواردی که طول خطوط از ۰,۳ بیشترین مقدار طول خط در تصویر کمتر باشد به جای انحنای به دست آمده، از انحنای خط بالایی و یا پایینی جهت رفع اعوجاج هندسی آن خط استفاده گردد.

۴-۵-۳- تعیین نقشه‌ی انتقال نقاط

پس از مشخص شدن انحنای هر خط، هدف برطرف کردن این انحنای می‌باشد. به همین منظور و با توجه به محل قرارگیری خطوط انحنای در اسناد فارسی و انگلیسی در گام اول باید هر خط را بین دو منحنی با انحنای یکسان قرار داد و با استفاده از تکنیک تخمین انحنای به خطوط تکه‌ای خطی به طول کوچک، ناحیه‌ی بین دو منحنی را به تعدادی چهارضلعی تبدیل کرده و با استفاده از تخمین تبدیل پرسپکتیو هر چهارضلعی را به یک مستطیل انتقال دهیم. این گام‌ها به صورت زیر بیان می‌شوند:

۱- قرار دادن هر خط بین دو منحنی

۲- تقسیم ناحیه‌ی بین دو منحنی به تعدادی چهارضلعی

۳- انتقال نقاط از چهارضلعی به مستطیل

در ادامه‌ی این بخش به توضیح هر کدام از این گام‌ها پرداخته می‌شود.

۴-۵-۳-۱- قرار دادن هر خط بین دو منحنی

در این گام با توجه به محل واقع شدن منحنی در متون فارسی و انگلیسی هر خط بین دو منحنی قرار می‌گیرد. برای این منظور کفایت طبق رابطه‌ی ۴-۲ دو منحنی درجه سه با ضرایب a_1 تا a_3 یکسان با ضرایب منحنی به دست آمده ایجاد کرده، ولی ضریب a_0 این دو منحنی با توجه به مکان قرارگیری دو منحنی متفاوت می‌باشد. فاصله‌ی بین دو منحنی بر اساس بیشترین ارتفاع حروف در هر خط انتخاب می‌شود.

۴-۵-۳-۲- تقسیم ناحیه‌ی بین دو منحنی به تعدادی چهارضلعی

در این گام ابتدا باید ناحیه‌ی بین دو منحنی را توسط خطوط عمودی بافاصله‌ی یکسان از هم به نوارهای عمودی شکل تبدیل کنیم. گام بعدی محاسبه‌ی نقاط تلاقی خطوط عمودی با دو منحنی می‌باشد. اگر فاصله‌ی بین خطوط عمودی کم باشد؛ می‌توان این نواحی را چهارضلعی فرض کرد. با توجه به این نکته که ما هر منحنی را با تعدادی خطوط با عرض کم تخمین زده‌ایم؛ بنابراین فرض چهارضلعی بودن این نواحی نیز تأیید می‌شود. با محاسبه‌ی نقاط تلاقی خطوط عمودی با دو منحنی درواقع مختصات چهار رأس هر چهارضلعی به دست خواهد آمد. در شکل ۴-۲۷ دو نمونه از اعمال این گام را مشاهده می‌کنید.

task!) also grows. For systems to interoperate, there must be a common language.

الف

فارسی: چهارضلعی‌ها (الف) متن انگلیسی (ب) متن فارسی

ب

شکل ۴-۲۷: نقاط رئوس چهارضلعی‌ها (الف) متن انگلیسی (ب) متن فارسی

۴-۵-۳-۳- انتقال نقاط از چهارضلعی به مستطیل

با به دست آمدن مختصات چهار رأس چهارضلعی‌ها، می‌توان با استفاده از روش تخمین تبدیل پرسپکتیو که در فصل سوم معرفی شده است، به انتقال نقاط از چهارضلعی‌ها به مستطیل‌هایی با مختصات چهار رأس مشخص شده پرداخت. با این کار اعوجاج هندسی خطوط برطرف می‌شود؛ اما نقاطی وجود دارند که انتقال پیدا نکرده‌اند. در شکل ۴-۲۸ یک نمونه از نتیجه‌ی این گام را مشاهده می‌کنید.

nicate with each other. But the problems in communication grow, instead of shrink, with every new day of technology innovation. As more languages, increasingly complex data structures, and varied business needs develop, the difference between one system and another (even if they are performing the same business task!) also grows. For systems to interoperate, there must be a common language. This isn't a language in the sense that Java is a language; programs not written in Java must have just as much access as those that are in Java. Instead, it's a language that anyone with a dictionary can access; even if the words aren't the same, applications have access to the words, which can be translated into something that *can* be understood.

الف

communication

ب

شکل ۴-۲۸: نتیجه‌ی تصویر رفع اعوجاج هندسی شده (الف) تصویر بدون اعوجاج هندسی (ب) نمایش نقاط انتقال نیافته قسمتی از شکل ۴-۲۸ الف

۴-۵-۴ - همپوشانی نقاط انتقال نیافته

در این گام با توجه به این نکته که تعداد پیکسل‌های چهارضلعی و مستطیل یکسان نیست؛ و همین‌طور احتمال انتقال دونقطه از چهارضلعی به یک نقطه از مستطیل، نقاطی در تصویر نهایی وجود دارد که انتقال نیافته‌اند هیچ مقداری ندارند. این نقاط با مقدار یک جایگزین می‌شوند؛ به همین دلیل تصاویری چون شکل ۴-۲۸ ب ایجاد می‌شود که برای حل این مشکل گام‌های زیر پیشنهاد می‌شود:

۱- ستون‌هایی در تصویر که تمام مقادیر آن یک است را پیدا کرده، در این ستون‌ها اگر پیکسلی بود

که دو همسایه‌ی سمت چپ و راست آن مقدار صفر داشت، مقدار پیکسل مذکور برابر با صفر می‌شود.

۲- در هر سطر اگر پیکسلی با مقدار یک وجود داشت که دو همسایه‌ی بالایی و پایینی آن مقدار صفر

داشتند، مقدار آن پیکسل نیز صفر می‌شود.

با اعمال گام‌های فوق تصویر خروجی بدون نقاط سفیدرنگ زائد ایجاد می‌شود. نتیجه‌ی اعمال این مرحله را در شکل ۴-۲۹ مشاهده می‌کنید.

nicate with each other. But the problems in communication grow, instead of shrink, with every new day of technology innovation. As more languages, increasingly complex data structures, and varied business needs develop, the difference between one system and another (even if they are performing the same business task!) also grows. For systems to interoperate, there must be a common language. This isn't a language in the sense that Java is a language; programs not written in Java must have just as much access as those that are in Java. Instead, it's a language that anyone with a dictionary can access; even if the words aren't the same, applications have access to the words, which can be translated into something that *can* be understood.

الف

communication

ب

شکل ۴-۲۹: همپوشانی نقاط انتقال نیافته الف) تصویر خروجی همپوشانی شده
ب) همپوشانی نقاط انتقال نیافته قسمتی از تصویر ۴-۲۹ الف

فصل پنجم

بررسی نتایج روش پیشنهادی

۵-۱- مقدمه

در این فصل به مقایسه‌ی نتایج روش‌های پیشنهادی با سایر روش‌های رایج پرداخته می‌شود. در بخش اول از این فصل به مقایسه‌ی نتایج روش پیشنهادی به‌منظور رفع اعوجاج فتومتری با روش‌های، نیبلک، اتسو و ژانگ پرداخته می‌شود. در بخش دوم به مقایسه‌ی نتایج روش پیشنهادی به‌منظور رفع اعوجاج هندسی با روش‌های خسروی راد و شامقلی پرداخته می‌شود.

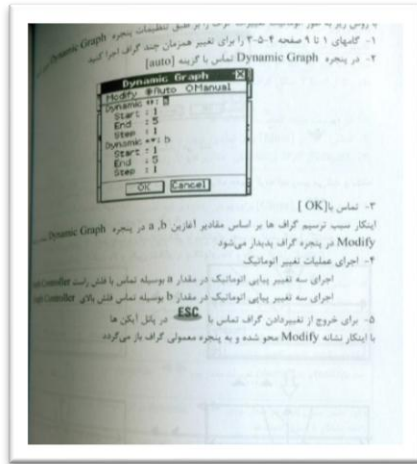
به‌منظور پیاده‌سازی روش‌ها از رایانه‌ای با مشخصات CPU Core(TM) 2 Duo، Ram 4 GB، win7، 64 bit و نرم افزار Matlab 2013 استفاده شده است. همان‌طور که مشخص است سیستم مورد استفاده یکی از سیستم‌های متوسط حال حاضر می‌باشد.

۵-۲- نتایج روش پیشنهادی به‌منظور رفع اعوجاج فتومتری

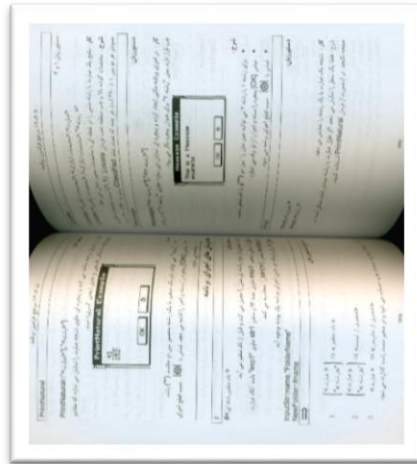
در این قسمت روش پیشنهادی برای رفع اعوجاج فتومتری با روش‌های ژانگ، نیبلک و اتسو مقایسه می‌شود. همان‌طور که در فصل چهارم ذکر شد در روش پیشنهادی که بهبود یافته‌ی روش ژانگ نیز می‌باشد از درون‌نگاری هشت‌تایی استفاده شده و با انتخاب ملایم کننده‌های مناسب علاوه بر کاهش چشم‌گیر زمان اجرا، دقت نیز افزایش می‌یابد. در پیاده‌سازی روش ژانگ این نکته مهم می‌باشد که برای تصاویر با درجه‌ی تفکیک‌پذیری بالا با توجه به زمان بالای اجرای عملیات ملایم کردن با تابع RBF این عملیات انجام نمی‌شود.

برای مقایسه، چهار تصویر از پایگاه داده‌ی شامقلی و خسروی راد و در نظر گرفته شده است و نتایج روش‌های ذکر شده باهم مقایسه می‌شوند. دو شکل ۵-۱ الف و ب از پایگاه داده‌ی شامقلی و تصاویر ۵-۱ ج و د هم از پایگاه‌های داده خسروی راد گرفته شده است. ابعاد هر کدام از این تصاویر نیز در جدول ۵-۱ آورده شده است.

نتایج چهار روش اتسو، نیبلک، ژانگ و روش پیشنهادی را در تصاویر ۵-۲ تا ۵-۵ نشان داده شده است سپس به بررسی نتایج هر روش می‌پردازیم. در بررسی این روش‌ها توجه به این نکته مهم می‌باشد که خروجی روش ژانگ و روش پیشنهادی برخلاف دو روش اتسو و نیبلک خاکستری می‌باشد.



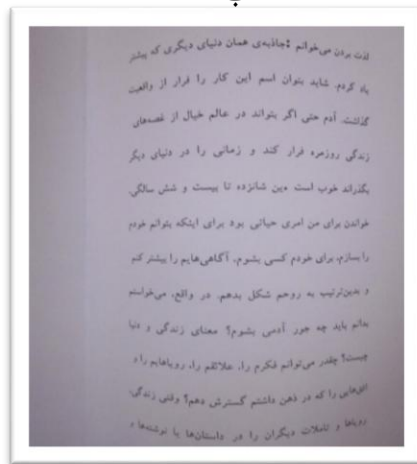
الف



ب



ج



د

شکل ۵-۱: نمونه تصاویر از پایگاه‌های داده‌ی مختلف الف و ب) پایگاه داده‌ی شامقلی ج و د) پایگاه داده‌ی خسروی راد

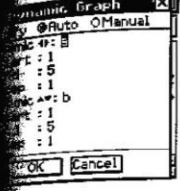
جدول ۵-۱: مشخصات تصاویر نمونه

ابعاد (پیکسل)		پایگاه داده	شماره‌ی تصاویر
عرض	طول		
۶۲۷	۷۴۹	شامقلی	شکل ۵-۱ الف
۲۴۵۷	۲۷۸۵	شامقلی	شکل ۵-۱ ب
۱۵۱۶	۲۴۰۸	خسروی راد	شکل ۵-۱ ج
۳۰۰۰	۴۰۰۰	خسروی راد	شکل ۵-۱ د

با روش زیر به طور انوماتیک تعبیرات کراف را بر طبق تنظیمات پنجره Dynamic Graph انجام دهید

۱- گامهای ۱ تا ۹ صفحه ۴-۵-۳ را برای تغییر همزمان چند گراف انجام دهید

۲- در پنجره Dynamic Graph تماش با گزینه [auto]



۳- تماش با [OK]

اینکار سبب ترسیم گراف ها بر اساس مقادیر آغازین a, b در پنجره Modify در پنجره گراف پدیدار می شود

۴- اجرای عملیات تغییر انوماتیک

اجرای سه تغییر پیاپی انوماتیک در مقدار a بوسیله فلش راست Controller

اجرای سه تغییر پیاپی انوماتیک در مقدار b بوسیله فلش بالای Controller

۵- برای خروج از تغییر دادن گراف تماش با **ESC** در پائل ایکن ها

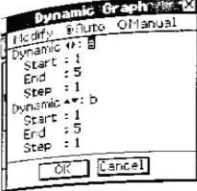
با اینکار نشانه Modify محو شده و به پنجره معمولی گراف باز می گردد

الف

با روش زیر به طور انوماتیک تعبیرات کراف را بر طبق تنظیمات پنجره Dynamic Graph انجام دهید

۱- گامهای ۱ تا ۹ صفحه ۴-۵-۳ را برای تغییر همزمان چند گراف اجرا کنید

۲- در پنجره Dynamic Graph تماش با گزینه [auto]



۳- تماش با [OK]

اینکار سبب ترسیم گراف ها بر اساس مقادیر آغازین a, b در پنجره Dynamic Graph در پنجره Modify در پنجره گراف پدیدار می شود

۴- اجرای عملیات تغییر انوماتیک

اجرای سه تغییر پیاپی انوماتیک در مقدار a بوسیله تماش با فلش راست Controller

اجرای سه تغییر پیاپی انوماتیک در مقدار b بوسیله تماش با فلش بالای Controller

۵- برای خروج از تغییر دادن گراف تماش با **ESC** در پائل ایکن ها


با اینکار نشانه Modify محو شده و به پنجره معمولی گراف باز می گردد

ب

با روش زیر به طور انوماتیک تعبیرات کراف را بر طبق تنظیمات پنجره Dynamic Graph انجام دهید

۱- گامهای ۱ تا ۹ صفحه ۴-۵-۳ را برای تغییر همزمان چند گراف اجرا کنید

۲- در پنجره Dynamic Graph تماش با گزینه [auto]



۳- تماش با [OK]

اینکار سبب ترسیم گراف ها بر اساس مقادیر آغازین a, b در پنجره Dynamic Graph در پنجره Modify در پنجره گراف پدیدار می شود

۴- اجرای عملیات تغییر انوماتیک

اجرای سه تغییر پیاپی انوماتیک در مقدار a بوسیله تماش با فلش راست Controller

اجرای سه تغییر پیاپی انوماتیک در مقدار b بوسیله تماش با فلش بالای Controller

۵- برای خروج از تغییر دادن گراف تماش با **ESC** در پائل ایکن ها


با اینکار نشانه Modify محو شده و به پنجره معمولی گراف باز می گردد

ج

با روش زیر به طور انوماتیک تعبیرات کراف را بر طبق تنظیمات پنجره Dynamic Graph انجام دهید

۱- گامهای ۱ تا ۹ صفحه ۴-۵-۳ را برای تغییر همزمان چند گراف اجرا کنید

۲- در پنجره Dynamic Graph تماش با گزینه [auto]



۳- تماش با [OK]

اینکار سبب ترسیم گراف ها بر اساس مقادیر آغازین a, b در پنجره Dynamic Graph در پنجره Modify در پنجره گراف پدیدار می شود

۴- اجرای عملیات تغییر انوماتیک

اجرای سه تغییر پیاپی انوماتیک در مقدار a بوسیله تماش با فلش راست Controller

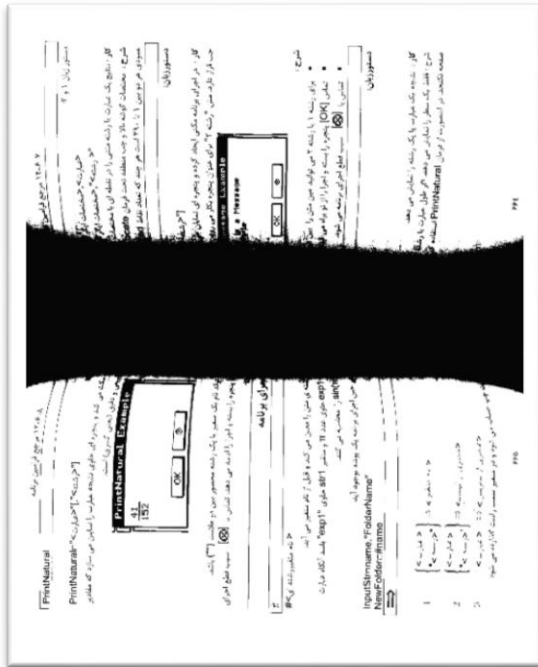
اجرای سه تغییر پیاپی انوماتیک در مقدار b بوسیله تماش با فلش بالای Controller

۵- برای خروج از تغییر دادن گراف تماش با **ESC** در پائل ایکن ها

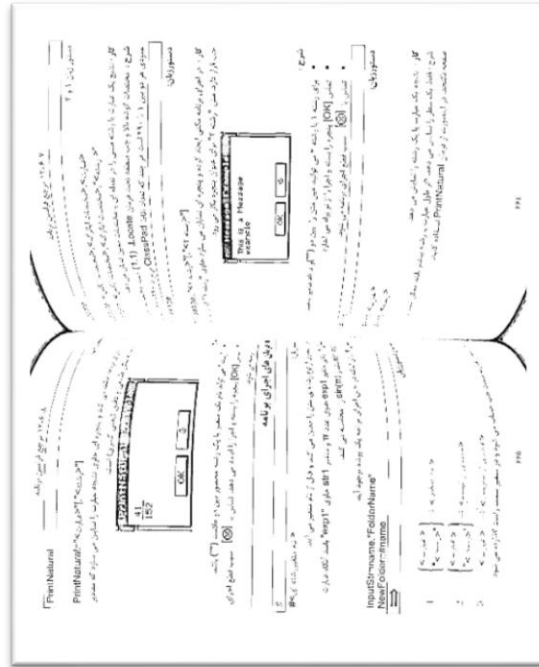
با اینکار نشانه Modify محو شده و به پنجره معمولی گراف باز می گردد

د

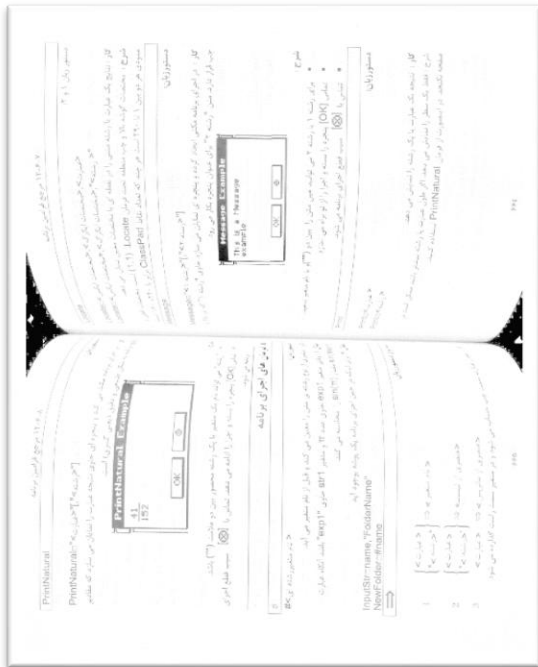
شکل ۵-۲: نتایج روش های مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ الف، الف) روش اتسو (ب) روش نیبلک (ج) روش ژانگ (د) روش پیشنهادی



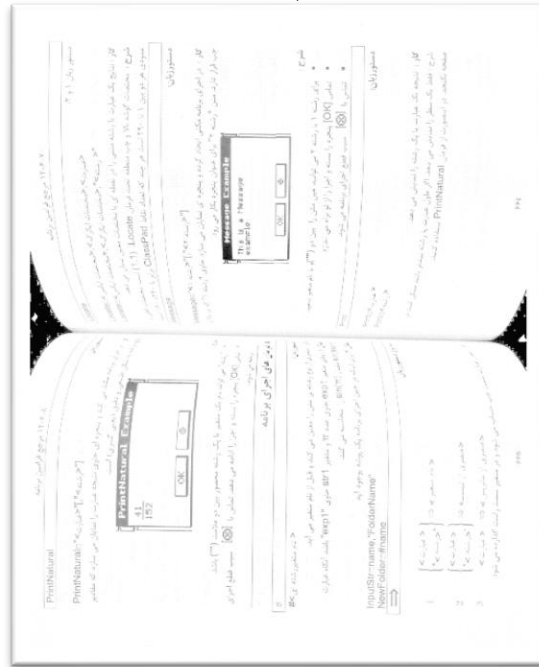
الف



ب

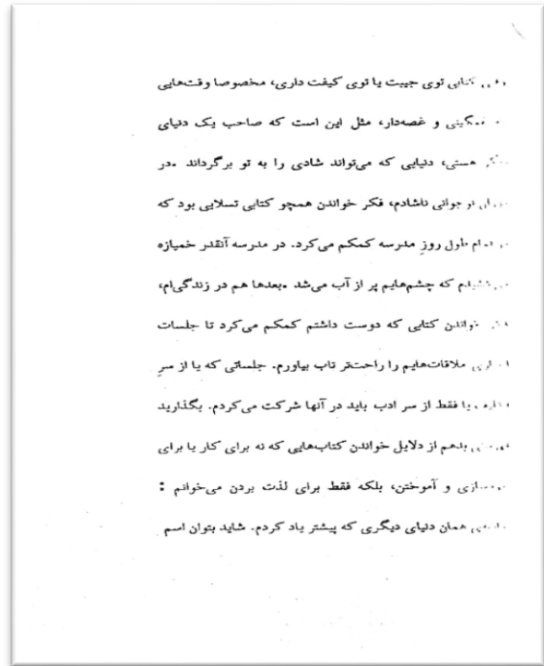
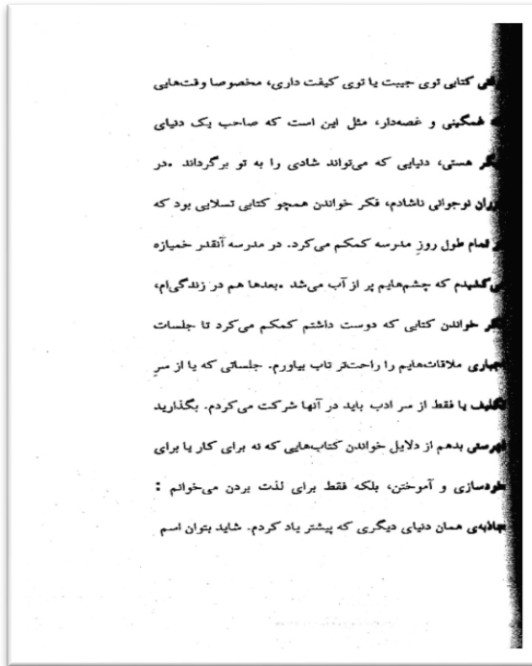


ج



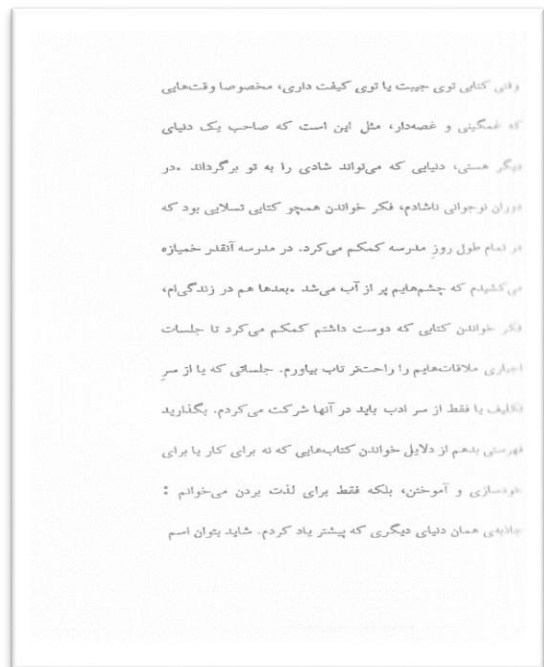
د

شکل ۳-۵: نتایج روش‌های مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ (ب، الف) روش اتسو (ب) روش نیبلک (ج) روش ژانگ (د) روش پیشنهادی



الف

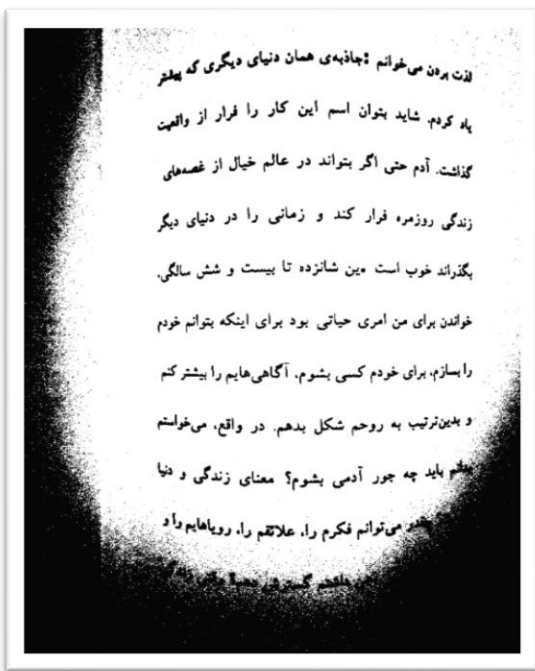
ب



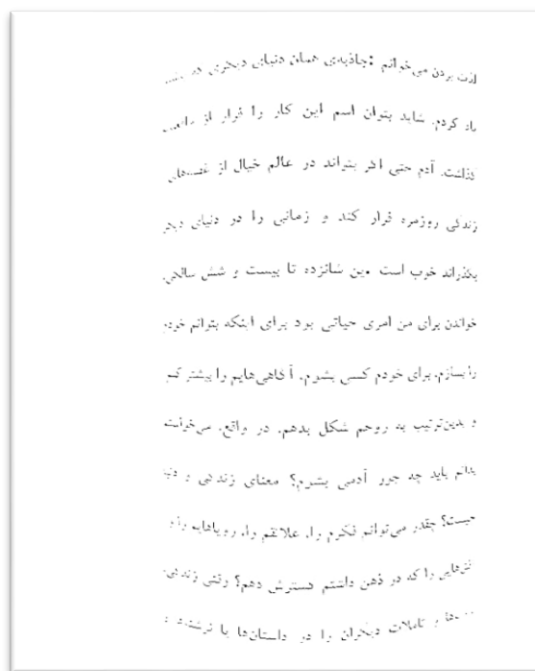
ج

د

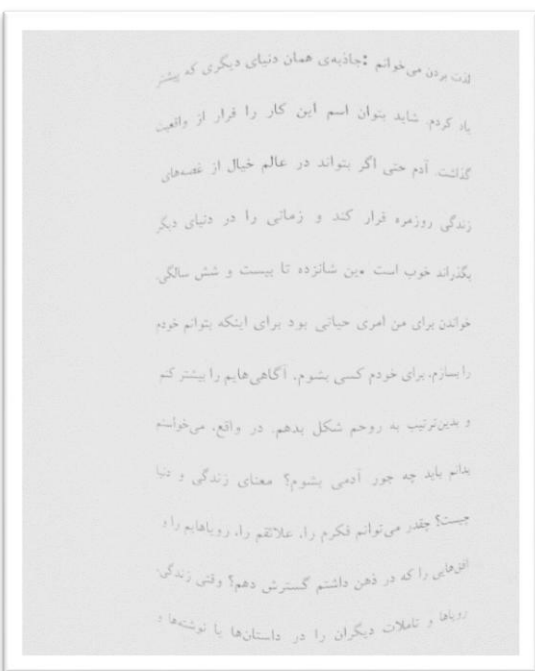
شکل ۵-۴: نتایج روش‌های مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ (ج، الف) روش اتسو (ب) روش نیبلک (ج) روش ژانگ (د) روش پیشنهادی



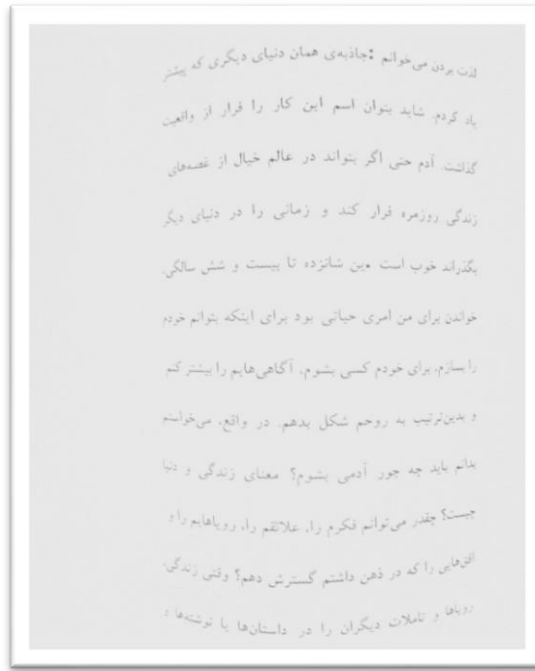
الف



ب



ج

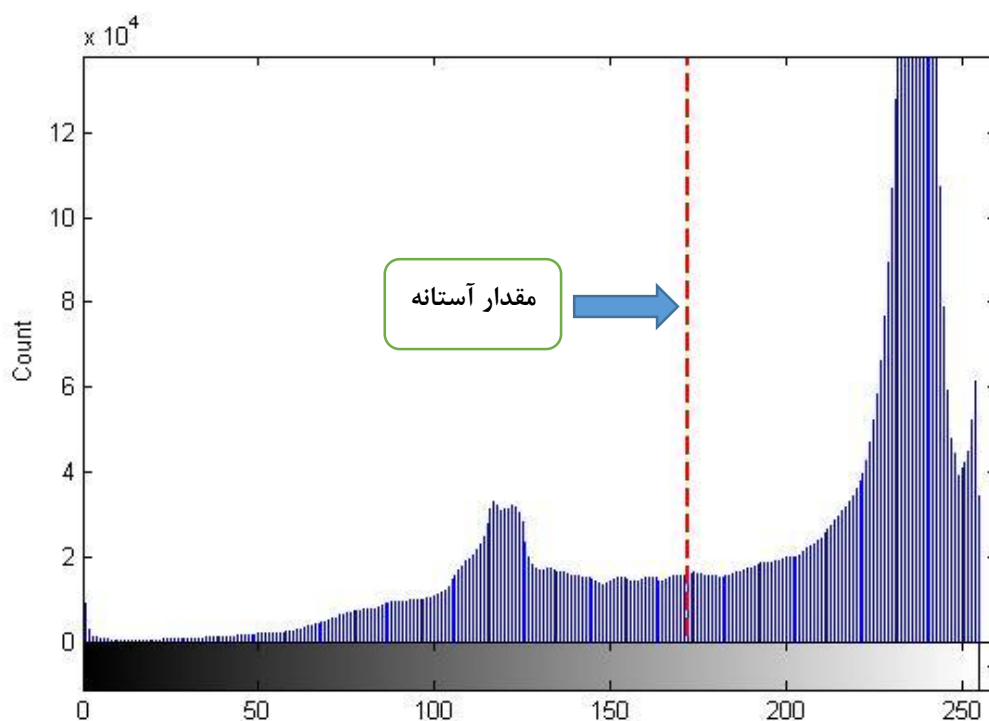


د

شکل ۵-۵: نتایج روش‌های مختلف برای رفع اعوجاج فتومتریک بر روی تصویر ۵-۱ (الف) روش اتسو (ب) روش نیبلک (ج) روش ژانگ ($k=0.9$) (د) روش پیشنهادی ($k=0.9$)

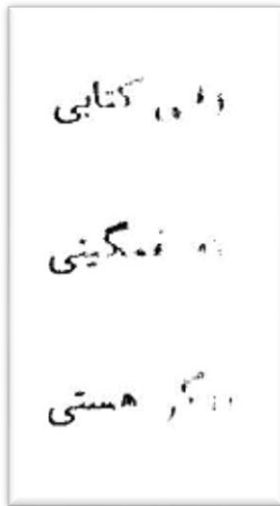
۵-۲-۱- بررسی نتایج روش‌های نیبلک و اتسو

همان‌طور که در تصاویر ۵-۲ تا ۵-۵ مشاهده می‌کنید در نتایج روش اتسو حاشیه‌های سیاه و نوفه فلفل نمکی زیادی وجود دارد در نتیجه، این روش نتایج قابل‌قبولی را ارائه نمی‌دهد. علت ناکارآمد بودن این روش، وجود روشنایی غیریکنواخت در تصاویر بوده و با توجه به این نکته که روش اتسو با یک مقدار آستانه تصویر را دوسطحی می‌کند، لذا در این مواقع این روش مناسب نیست. به‌عنوان نمونه هیستوگرام تصویر ۵-۱ ب و مقدار آستانه‌ی محاسبه‌شده توسط روش اتسو در شکل ۵-۶ نشان داده‌شده است. با توجه به تعداد قله‌ها در تصویر هیستوگرام مشخص است که با یک مقدار آستانه این تصویر به‌خوبی دوسطحی نخواهد شد.

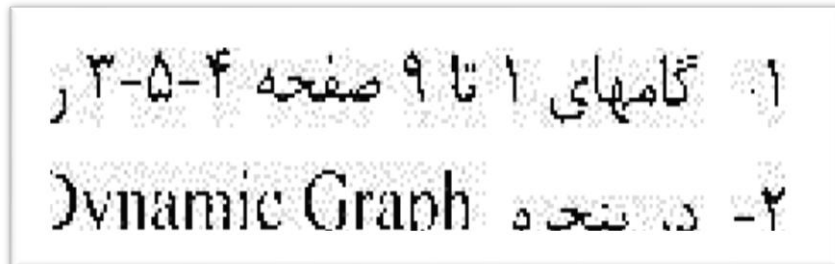


شکل ۵-۶: هیستوگرام تصویر ۵-۱ ب به همراه مقدار آستانه محاسبه شده به روش اتسو

با مشاهده‌ی نتایج روش نیبلک به این موضوع دست می‌یابیم که این روش نیز از دقت کافی به‌منظور حذف اعوجاج فتومتریک برخوردار نیست. هرچند که این روش نسبت به روش اتسو برای تصاویر با روشنایی غیریکنواخت نتیجه‌ی بهتری ارائه می‌دهد؛ ولی نوفه فلفل نمکی زیاد و حروف شکسته و حساسیت بیش‌ازحد به پارامتر k ، از مهم‌ترین مشکلات این روش بوده که باعث بروز اشتباه در تشخیص حروف می‌شود. در شکل ۵-۷ دو نمونه از این مشکلات نشان داده‌شده است.



الف



ب

شکل ۵-۷: دو نمونه از مشکلات روش نیبلک الف) حروف شکسته قسمتی از تصویر ۴-۵ ب) نوفه فلفلی قسمتی از تصویر ۲-۵ ب

اما در روش ژانگ و روش پیشنهادی مشکلات دو روش نیبلک و اتسو وجود ندارد و تصویر خروجی تا حد امکان بدون نوفه فلفلی و حروف شکسته شده می‌باشد.

۵-۲-۲- بررسی نتایج روش ژانگ و روش پیشنهادی

قبل از مقایسه نتایج روش ژانگ با روش پیشنهادی ابتدا لازم است مقدار پارامترهای این دو روش معرفی گردند. اولین پارامتر مهم مربوط به تعداد مراحل تکرار الگوریتم درون‌نگاری می‌باشد. هرچه این مقدار بزرگ‌تر باشد تصویر درون‌نگاری شده نتیجه‌ی بهتری خواهد داشت؛ ولی در عوض مدت‌زمان اجرای الگوریتم نیز افزایش می‌یابد. پس از آزمایش‌های مختلف به این نتیجه رسیدیم که تعداد مراحل تکرار بین ۱۰ تا ۴۰ مرحله برای تصاویر مختلف مناسب می‌باشد و از این بین ما مقدار ۲۰ مرحله را برای آزمایش انتخاب کردیم. همچنین پارامترهای Δt ، λ و h که مربوط به رابطه‌ی درون‌نگاری بوده نیز توسط آزمایش و به صورت سعی و خطا انتخاب می‌شوند. با توجه به رابطه‌ی درون‌نگاری (رابطه‌ی ۳-۷) پارامترهای λ و h به صورت $\frac{\lambda}{h^2}$ استفاده شده است، پس بهتر است به جای مقداردهی به هر کدام از آن‌ها، تنها مقدار این نسبت را انتخاب کنیم. اگر مقدار این نسبت بزرگ (بزرگ‌تر از ۱) باشد، باعث می‌شود تا تصویر درون‌نگاری شده به سمت خاکستری و مشکی تمایل پیدا کند و نتیجه‌ی خروجی را خراب کند. درحالی‌که مقدار خیلی کوچک نیز

باعث کم‌ارزش کردن مقدار $(I_{i+1,j}^n + I_{i-1,j}^n + I_{i,j+1}^n + I_{i,j-1}^n - 4I_{i,j}^n)$ در رابطه‌ی ۴-۱ می‌شود، بنابراین نسبت $\frac{\lambda}{h^2}$ مقداری بین ۰,۲ تا ۰,۶, پیشنهاد می‌شود که البته ما از مقدار ۰,۵ برای آزمایش استفاده کرده‌ایم. مقدار Δt نیز تعیین‌کننده‌ی میزان تأثیرپذیری هر مرحله‌ی درون‌نگاری بر روی نتایج مراحل قبلی می‌باشد که بر اساس آزمایش مقداری ۰,۱ تا ۰,۴ مناسب بوده که البته ما مقدار ۰,۲ را برای آن در نظر گرفته‌ایم. همان‌طور که در فصل چهارم بیان شد، در روش پیشنهادی از ملایم‌کننده‌های گوسی و میانگین‌گیر استفاده می‌کنیم که البته به‌منظور کاهش پیچیدگی پیشنهاد شد که از ملایم‌کننده‌های سطری با ستونی استفاده کنیم. برای دو تصویر ۵-۱ الف و ب از ملایم‌کننده‌ی میانگین‌گیر به طول ۲۰ پیکسل و برای دو تصویر ۵-۱ ج و ۵-۱ د، از ملایم‌کننده‌ی گوسی به طول ۲۰ پیکسل و واریانس ۰,۵ استفاده کرده‌ایم.

برای مقایسه‌ی نتایج روش ژانگ با روش پیشنهادی بررسی دو نکته مهم می‌باشد:

۱: مقایسه‌ی سرعت اجرا

۲: مقایسه‌ی دقت تصاویر رفع اعوجاج شده

از لحاظ سرعت اجرا همان‌طور که قبلاً ذکر شده است، در روش ژانگ، استفاده از ملایم‌کننده‌های RBF به علت وقت‌گیر بودن برای تصاویر با درجه‌ی تفکیک‌پذیری بالا پیشنهاد نمی‌شود. با مشاهده‌ی جدول ۵-۲ می‌توان به این نتیجه رسید که سرعت اجرای روش پیشنهادی در مقایسه با روش ژانگ با استفاده از ملایم‌کننده‌ی RBF بسیار بیشتر می‌باشد. نتایج این جدول بر اساس روشی که در فصل چهارم به‌منظور کاهش حلقه‌های تودرتو بیان شد، ارائه شده است.

همان‌طور که در فصل چهارم بیان شد؛ با کاهش حلقه‌های تودرتو سرعت اجرای الگوریتم افزایش می‌یابد. در جدول ۵-۳ نتایج مدت‌زمان اجرای روش ژانگ و روش پیشنهادی بدون کاهش حلقه‌های تودرتو نشان داده شده است. با مقایسه نتایج دو جدول ۵-۲ و ۵-۳ می‌توان به این نتیجه رسید که با کاهش حلقه‌های تودرتو سرعت اجرای الگوریتم افزایش می‌یابد.

جدول ۲-۵: مدت زمان اجرای روش پیشنهادی و روش ژانگ

روش پیشنهادی	روش ژانگ		مدت زمان روش (s) ابعاد تصویر (پیکسل)
	بدون استفاده از ملایم کننده RBF	با استفاده از ملایم کننده RBF	
۴	۳,۵	۵۶	شکل ۱-۵ الف (۷۴۹*۶۲۷)
۴۵	۴۳	۸۷۰	شکل ۱-۵ ب (۲۷۵۸*۲۴۵۷)
۳۰	۲۹	۴۷۰	شکل ۱-۵ ج (۲۴۰۸*۱۵۱۶)
۶۸	۶۶	۱۲۵۰	شکل ۱-۵ د (۴۰۰۰*۳۰۰۰)

جدول ۳-۵: مدت زمان اجرای روش پیشنهادی و روش ژانگ بدون کاهش حلقه‌های تودرتو

روش پیشنهادی	روش ژانگ بدون استفاده از ملایم کننده RBF	مدت زمان روش (s) ابعاد تصویر (پیکسل)
۱۲۲	۱۲۰	شکل ۱-۵ ب (۲۷۵۸*۲۴۵۷)
۹۲	۹۰	شکل ۱-۵ ج (۲۴۰۸*۱۵۱۶)
۱۷۹	۱۷۷	شکل ۱-۵ د (۴۰۰۰*۳۰۰۰)

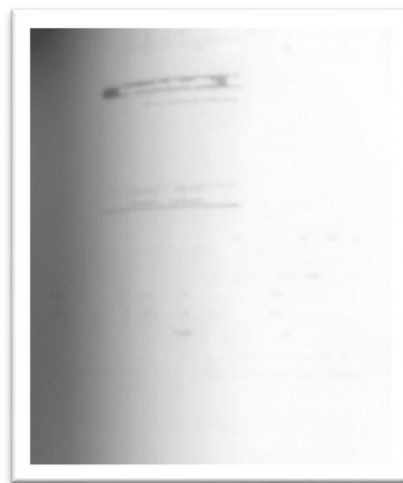
به منظور مقایسه‌ی روش درون نگاری پیشنهادی که به کمک هر ۸ همسایه و ملایم کننده‌ی میانگین گیر یا گوسی معرفی شده، بهتر است نتایج مرحله‌ی درون نگاری را باهم مقایسه کنیم. همان طور که در فصل چهارم ذکر شد هدف از درون نگاری حذف لایه‌ی پیش زمینه از تصویر می باشد. در شکل ۵-۸ دو نمونه از نتایج این مرحله برای دو روش ژانگ و روش پیشنهادی نشان داده شده است. همان طور که در این تصاویر مشاهده می کنید؛ روش پیشنهادی به علت استفاده از ملایم کننده‌ی میانگین گیر بهتر عمل کرده و تصویر را به خوبی ملایم کرده است.

با مشاهده‌ی تصاویر ۲-۵ تا ۵-۵ نیز می توان به این نتیجه رسید که روش پیشنهادی با توجه به اینکه سرعتی به مراتب بالاتر نسبت به روش ژانگ دارد، در حذف اعوجاج فتومتریک برای تصاویر با درجه‌ی

تفکیک پذیری بالا نیز مناسب می باشد. در واقع روش پیشنهادی با حذف لایه ی پیش زمینه از تصاویر به خوبی توانسته است، اعوجاج فتومتریک را حذف نماید.



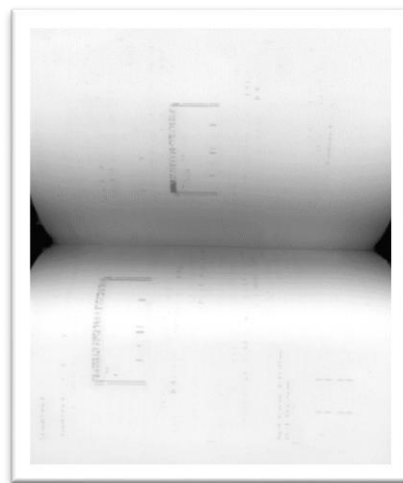
الف



ب



ج



د

شکل ۵-۸: نتایج مرحله ی درون نگاری دو تصویر ۱-۵ الف و ۱-۵ ب، الف و ج (روش پیشنهادی ب و د) روش ژانگ

۳-۵- نتایج روش پیشنهادی به منظور رفع اعوجاج هندسی

در فصل چهارم بخش‌های مختلف الگوریتم پیشنهادی، به منظور رفع اعوجاج هندسی معرفی شده است. همان‌طور که قبلاً اشاره شده است این الگوریتم با استخراج خطوط سعی در رفع اعوجاج هندسی، هر خط را دارد؛ بنابراین با پیدا کردن هر خط و یافتن نقشه‌ی خمیدگی آن خط اعوجاج هندسی برطرف می‌گردد. در این بخش چند نمونه از نتایج روش پیشنهادی برای اسناد فارسی و انگلیسی نشان داده می‌شود. همین‌طور به منظور مقایسه‌ی بهتر نتایج توسط نرم افزار OmniPage [۶۵] و پرشیانگار [۶۶] مورد ارزیابی قرار می‌گیرند. در شکل ۵-۹ چند نمونه از تصاویر پایگاه‌های داده‌ی خسروی‌راد و DFKI نشان داده شده است.

از آنجایی که روش پیشنهادی با تشخیص خطوط به رفع اعوجاج هندسی می‌پردازد و شکل و نمودار می‌تواند در پیدا کردن خطوط اخلال ایجاد کنند، بنابراین تصاویر پایگاه داده‌ی خسروی‌راد به دلیل اینکه در این پایگاه داده تصاویر کاملاً به صورت نوشتاری بوده و هیچ‌گونه نمودار یا شکلی در تصاویر وجود ندارد، پایگاه داده‌ای مناسب می‌باشد؛ بنابراین از بین تصاویر پایگاه داده‌ی DFKI، تصاویری انتخاب شده و مورد تحلیل قرار می‌گیرند که شامل نمودار و یا شکل نباشند.

از بین روش‌های معرفی شده در فصل دوم به منظور مقایسه‌ی نتایج، دو روش خسروی‌راد و شامقلی برای اسناد فارسی و روش شامقلی، SEG و SKEL برای اسناد انگلیسی مورد ارزیابی قرار می‌گیرند.

۳-۵-۱- نتایج تجربی روش پیشنهادی بر روی تصاویر

قبل از بررسی نتایج ابتدا مقادیر پارامترهای روش پیشنهادی بیان می‌شوند. اولین پارامتر مربوط به طول جزء ساختاری عملگر گسترش می‌باشد. همان‌طور که در فصل چهارم بیان شد تعیین مقدار مناسب برای جزء ساختاری به منظور تشخیص صحیح خطوط، بسیار مهم می‌باشد. به همین منظور این مقدار برای پایگاه داده‌ی خسروی‌راد برابر ۱۵۰ و برای پایگاه داده‌ی DFKI برابر با ۶۰ می‌باشد؛ البته ممکن است این مقادیر برای برخی از تصاویر اندکی تغییر یابد. در تعیین این مقادیر درجه‌ی تفکیک‌پذیری تصویر و فاصله‌ی بین خطوط نقشه مهمی دارند. به گونه‌ای که هرچه درجه‌ی تفکیک‌پذیری تصویر بیشتر باشد، طول جزء ساختاری می‌تواند بزرگ‌تر باشد؛ همچنین هرچه فاصله‌ی بین خطوط بیشتر باشد نیز طول جزء ساختاری می‌تواند بزرگ‌تر شود البته تا زمانی که ماسک دو خطوط به هم متصل نشوند.

پارامتر بعدی مربوط به عرض نوارهای عمودی می‌باشد این پارامتر نیز تا حدودی وابسته به درجه‌ی تفکیک‌پذیری تصویر می‌باشد و باید به گونه‌ای انتخاب شود که در هر نوار چند حرف قرار گرفته باشد. به

همین منظور برای پایگاه داده خسروی راد مقداری بین ۱۰۰ تا ۲۵۰ و برای پایگاه داده‌ی DFKI مقداری بین ۸۰ تا ۲۰۰ مناسب می‌باشد.

CHAPTER 5 FULL CONTENT DATA

USING ETHEREAL TO READ STORED FULL CONTENT DATA

Ethereal can read in a capture file via a command-line invocation, as shown here. Figure 5.2 shows the result in the Ethereal graphical user interface (GUI).

```
ethereal -n -r ee0.tpc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI fans. We're going to look at the `ee0.tpc` trace and search for evidence of the DDoS attacks launched during the reference intrusion model in Chapter 4. Once the trace is loaded, we must use a different filter syntax to focus our analysis efforts. To see traffic to or from port 7793 UDP, for example, use `udp.port == 7793` in the Filter field at the bottom of the main window, as shown in Figure 5.3.

The output displayed in Figure 5.3 shows an odd UDP packet. The content of the application data reads `msrcan/172.27.20.2/10.2/10`, followed by `0x0a`. This is the command sent from the Metstream DDoS master daemon running on 172.27.20.5 to the Metstream

A web publishing framework attempts to address these complicated issues. Just as a web server is responsible for responding to a URL request for a file, a web publishing framework is responsible for responding to a similar request; however, instead of responding with a file, it often will respond with a published version of a file. In this case, a published file refers to a file that may have been transformed with XSLT, managed at an application level, or converted into another format such as a PDF. The requestor does not see the raw data that may underlie the published result, but also does not have to explicitly request that publication occur. Often, a URL base (such as `http://www.Hot.com/publish`) signifies that a publishing engine that sits on top of the web server should handle requests. As you may suspect, the concept is much simpler than the actual implementation of a framework like this, and finding the correct framework for your needs is not a trivial task.

Selecting a Framework

You might expect to find a list of hundreds of possible solutions. As you've seen, the Java language offers an easy interface into XML through several APIs. Additionally, Java servlets offer a simple means of handling web requests and responses. However, the list of frameworks is small, and the list of good, stable ones is even smaller. One of the best resources for seeing what products are currently available is XML Software's list at <http://xmlsoft.com/publishing/>. This list changes so frequently that it is not worth repeating here. Still, some important criteria for determining what framework is right for you are worth mentioning.

Stability

Don't be surprised if you (still) have a hard time finding a product whose versioning is greater than 2.x. In fact, you may have to search diligently to even find a second-generation framework. While a higher version number is not a guarantee of stability, it often reflects the amount of time, effort, and review that a framework has undergone. The XML publishing system is such a new beast that the market has been flooded with 1.0 and 1.1 products that simply are not stable enough for practical use.

You can often ascertain the stability of a product by investigating other products from the same vendor. Often a vendor releases an entire suite of tools; if their other tools do not offer SAX 2.0 and DOM Level 2 support, or are all also 1.0 and 1.1 products, you might be wise to pass on the framework until it has matured and conformed to newer XML standards. Try to steer away from platform-specific technologies. If the framework is tied to a platform (such as Windows, or even a specific flavor of Unix), you aren't dealing with a pure Java solution. Remember that a publishing framework must serve clients on any platform; why use a product that can't run on any platform?

For HTML, DOM provides a set of interfaces that model the various HTML elements. For example, you can use the `HTMLDocument` class, the `HTMLAnchorElement`, and the `HTMLSelectElement` (all in the `org.w3c.dom.html` package) to represent their analogs in HTML (`<HTML>`, `<a>`, and `<select>` in this case). All of these provide convenience methods like `setTitle()` (on `HTMLDocument`), `setHref()` (on `HTMLAnchorElement`), and `getOptions()` (on `HTMLSelectElement`). All of these extend core DOM structures like `Document` and `Element`, and so can be used as any other DOM node could.

However, it turns out that the HTML bindings are rarely used (at least directly). It's not because they aren't useful; instead, many tools have already been written to provide this sort of access through even more user-friendly tools. XSLC, a project within the Enhydra application server framework, is one such example (located online at <http://xml.enhydra.org>), and Cocoon, covered in Chapter 10, is another. These allow developers to work with HTML and web pages in a way that does not necessarily require even basic DOM knowledge, making it more accessible to web designers and newer Java developers. The end result of using these tools is that the HTML DOM bindings are rarely needed. But if you know about them, you can use them if you need to. Additionally, you can use standard DOM functionality on well-formed HTML documents (XHTML), treating elements as `Element` nodes and attributes as `Attr` nodes. Even without the HTML bindings, you can use DOM to work with HTML. Piece of cake.

Odds and Ends

What's left in DOM Level 2 besides these modules and `namespaceAwareness`? Very little, and you've probably already used most of it. The `createElement()` and `createElementNS()` methods are new to the `DOMImplementation` class, and you've used both of them. Additionally, the `getSystemId()` and `getPublicId()` methods used in the `DOMSerializer` class on the `DocumentType` interface are also DOM Level 2 additions. Other than that, there isn't much; a few new `DOMException` error code constants, and that's about it. You can see the complete list of changes online at <http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/changes.html>. The rest of the changes are the additional modules, one of which I'll cover next.

DOM Level 3

Before closing the book on DOM and looking at common gotchas, I will spend a little time letting you know what's coming in DOM Level 3, which is underway right now. In fact, I expect this specification to be finalized early in 2002, not long

الف

هنگام عروسی محمدرضا شاه پهلوی و زن مصری اش فوزه در سال ۱۳۱۷ خورشیدی چون مقرر بود، میهمانان مصری و همراهان عروس به وسیله راه آهن جنوب به تهران وارد شوند از طرف دربار و شهربانی دستور اکید صادر شده بود که دیوارهای تمام دهات طول راه و خانه های دهقانی مجاور خط آهن را سفید کنند.

در یکی از دهات چون گنج در دسترس نبود بخشدار دستور می‌دهد که با کتشک و ماست که در آن ده فراوان بوده دیوارها را موقتاً سفید نمایند، و به این منظور با پولی که از کدخدای ده می‌گیرند و با خرید مقدار زیادی ماست کلیه دیوارها را ماست مالی کردند.

نیمت تاریخی این اصطلاح ماست مالی از شصت سال نمی‌گذرد؛ و تاجرای این ماست مالی مدت ها موضوع اصلی شوخی‌های محافل مجلس بود. حکیم ارد بزرگ می‌گوید: فرمانروایان تنها پاسخگوی زمان حال خویش نیستند آنها به گذشتگان و آیندگان نیز پاسخگو هستند.

ت

شکل ۵-۹: نمونه تصاویر دارای اعوجاج هندسی الف، ب، پ) پایگاه داده‌ی DFKI

ت، ج، د) پایگاه داده‌ی خسروی راد

ب

بذت بردن می‌خوانم مجاذبه‌ی همان دنیای دیگری که بیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. این شانزده تا بیست و شش سالگی، خوانند برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم. برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و

بدین ترتیب به روح شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ جعفر می‌توانم فکرم را، علاقه‌م را، رویاهایم را و افتخاری را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، روایا و

ج

پ

خواننده کتاب‌هایی که نه برای کار یا برای هوسهای و آموختن، بلکه فقط برای لذت بردن می‌خوانم؛ مجاذبه‌ی همان دنیای دیگری که بیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. این شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم. برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین ترتیب به روح شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ جعفر می‌توانم فکرم را، علاقه‌م را، رویاهایم را و افتخاری را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، روایا و

د

پارامتر بعدی مربوط به حداکثر زاویه چرخش هر نوار می‌باشد؛ که این مقدار وابسته به میزان چرخش نسبی کلمات یا خطوط می‌باشد. مقدار بزرگ برای این پارامتر به طولانی شدن مدت اجرای الگوریتم می‌انجامد و کوچک بودن آن می‌تواند در تشخیص اشتباه انحنا خطوط بیانجامد. با توجه به میزان انحنا نسبی تصاویر پایگاه داده مقداری بین ۲۰ تا ۴۰ درجه مناسب می‌باشد.

در ادامه به بررسی نتایج بر روی اسناد انگلیسی می‌پردازیم؛ پس‌از آن نتایج بر روی اسناد فارسی ارزیابی می‌شوند.

۵-۳-۲- مقایسه‌ی نتایج بر روی اسناد انگلیسی

به‌منظور مقایسه‌ی بهتر، نتایج روش پیشنهادی همراه با نتایج سه روش SEG [۴۳]، SKEL [۴۰] و شامقلی [۴۷] در تصاویر ۵-۱۰ تا ۵-۱۲ نشان داده‌شده است.

در روش SEG، خمیدگی هر کلمه به‌تنهایی توسط شیب آن کلمه برطرف می‌گردد. پس‌از آن با توجه به چپ‌ترین کلمه‌ی موجود در آن خط هم‌ترازی عمودی صورت می‌گیرد. در این روش به خاطر چرخش و جابجایی صورت گرفته برای قرار گرفتن در مکان نهایی شکل حروف تا حدود زیادی تخریب خواهد شد. این روش تنها بر روی اسناد بدون اجزای گرافیکی کارایی دارد ولی قادر به حذف اجزای گرافیکی با اعمال فیلتری که اجزای گرافیکی را بر اساس سائزشان حذف می‌کند، می‌باشد.

ایده‌ی اصلی روش SKEL بر اساس استخراج اسکلت بیرونی تصاویر متنی می‌باشد که در آن شاخه‌های به‌هم‌پیوسته توسط منحنی‌هایی تقریب زده می‌شوند و فضای درون خطوط متن را تعیین می‌کنند. این روش برخلاف روش SEG قادر به بازیابی اجزای گرافیکی می‌باشد؛ اما این روش بسیار حساس به یافتن پیکره‌بندی بوده و در صورت ایجاد خطا در محاسبه‌ی آن خروجی اعوجاج خواهد داشت.

در تصویر ۵-۱۰ به‌خوبی می‌توان با چشم نتایج را بررسی نمود. همان‌طور که در این تصاویر مشاهده می‌کنید. روش SEG نتوانسته به خوبی در رفع اعوجاج هندسی به خصوص در نواحی که نشان داده شده است موفق باشد. این روش در رفع اعوجاج مکان‌هایی که عدد در متن قرار دارد؛ به علت عدم تشخیص صحیح شیب آن ناتوان است. با توجه به نواحی مشخص شده در شکل ۵-۱۰ می‌توان به این نتیجه رسید که روش پیشنهادی به خصوص در نقاط ابتدایی و انتهایی که میزان اعوجاج زیاد است نیز نسبت به روش SKEL و شامقلی از نتیجه‌ی بهتری در رفع اعوجاج برخوردار می‌باشد. علت برتری روش پیشنهادی اسن است که اعوجاج هر خط براساس انحنا همان خط برطرف می‌شود.

USING ETHEREAL TO READ STORED FULL CONTENT DATA

Ethereal can read in a capture file via a command-line invocation, as shown here. Figure 5.2 shows the result in the Ethereal graphical user interface (GUI).

```
ethereal -n -r en0.1pc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI fans. We're going to look at the en0.1pc trace and search for evidence of the DDoS attack launched during the reference intrusion model in Chapter 4. Once the trace is loaded, we must use a different filter syntax to focus our analysis efforts. To see traffic to or from port 7793 UDP, for example, use `udp.port == 7783` in the Filter field at the bottom of the main window, as shown in Figure 5.3.

The output displayed in Figure 5.3 shows an odd UDP packet. The content of the application data reads `mstream/172.27.20.2/10`, followed by `0x0a`. This is the command sent from the Mstream DDoS master daemon running on 172.27.20.5 to the Mstream

الف

USING ETHEREAL TO READ STORED FULL CONTENT DATA

Ethereal can read in a capture file via a command-line invocation, as shown here. Figure 5.2 shows the result in the Ethereal graphical user interface (GUI).

```
ethereal -n -r en0.1pc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI fans. We're going to look at the en0.1pc trace and search for evidence of the DDoS attack launched during the reference intrusion model in Chapter 4. Once the trace is loaded, we must use a different filter syntax to focus our analysis efforts. To see traffic to or from port 7793 UDP, for example, use `udp.port == 7783` in the Filter field at the bottom of the main window, as shown in Figure 5.3.

The output displayed in Figure 5.3 shows an odd UDP packet. The content of the application data reads `mstream/172.27.20.2/10`, followed by `0x0a`. This is the command sent from the Mstream DDoS master daemon running on 172.27.20.5 to the Mstream

ج

USING ETHEREAL TO READ STORED FULL CONTENT DATA

Ethereal can read in a capture file via a command-line invocation, as shown here. Figure 5.2 shows the result in the Ethereal graphical user interface (GUI).

```
ethereal -n -r en0.1pc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI fans. We're going to look at the en0.1pc trace and search for evidence of the DDoS attack launched during the reference intrusion model in Chapter 4. Once the trace is loaded, we must use a different filter syntax to focus our analysis efforts. To see traffic to or from port 7793 UDP, for example, use `udp.port == 7783` in the Filter field at the bottom of the main window, as shown in Figure 5.3.

The output displayed in Figure 5.3 shows an odd UDP packet. The content of the application data reads `mstream/172.27.20.2/10`, followed by `0x0a`. This is the command sent from the Mstream DDoS master daemon running on 172.27.20.5 to the Mstream

ب

USING ETHEREAL TO READ STORED FULL CONTENT DATA

Ethereal can read in a capture file via a command-line invocation, as shown here. Figure 5.2 shows the result in the Ethereal graphical user interface (GUI).

```
ethereal -n -r en0.1pc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI fans. We're going to look at the en0.1pc trace and search for evidence of the DDoS attack launched during the reference intrusion model in Chapter 4. Once the trace is loaded, we must use a different filter syntax to focus our analysis efforts. To see traffic to or from port 7793 UDP, for example, use `udp.port == 7783` in the Filter field at the bottom of the main window, as shown in Figure 5.3.

The output displayed in Figure 5.3 shows an odd UDP packet. The content of the application data reads `mstream/172.27.20.2/10`, followed by `0x0a`. This is the command sent from the Mstream DDoS master daemon running on 172.27.20.5 to the Mstream

د

شکل ۵-۱۰: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ (پ الف) روش SEG (ب) روش SKEL (ج) روش شاملی (د) روش پیشنهادی

همچنین در تصاویر ۵-۱۱ و ۵-۱۲ به خصوص در پاراگراف اول که شدت اعوجاج هندسی زیاد است برتری روش پیشنهادی قابل مشاهده است.

از جمله مشکلات دیگر روش SEG می‌توان دقت پایین در رفع اعوجاج خطوط با طول کم را بیان کرد. این نکته در تصاویر ۵-۱۱ و ۵-۱۲ نشان داده شده است و به گونه ایست که حتی در تصویر ۵-۱۲ کلمه‌ی Stabiliy طبق این روش حذف شده است. اما با توجه به این نکته که در روش پیشنهادی میزان انحنای خطوط با طول کم بر اساس انحنای خط قبلی برطرف می‌شود؛ بنابراین انحنای خطوط کوتاه نیز به خوبی برطرف می‌شود.

از میان سه روش مورد مقایسه با روش پیشنهادی، روش شامقلی که با تعیین خطوط غالب در بالا و پایین صفحه به رفع اعوجاج هندسی تصویر می‌پردازد نتیجه‌ی بهتری نسبت به دو روش SEG و SKEL دارد؛ اما از جمله مشکلات این روش می‌توان به سه مورد اشاره کرد:

۱- زمانی که تصاویر دارای خطوط غالب در بالا یا پایین تصویر نباشد، اعوجاج هندسی به خوبی برطرف نمی‌شود.

۲- با توجه به اینکه رفع اعوجاج هندسی طبق این روش بر اساس میزان انحنای دو خط غالب در تصویر می‌باشد، در مواقعی که در طول تصویر انحنای چندین بار تغییرشدید پیدا کند، اعوجاج هندسی به خوبی برطرف نمی‌شود.

۳- به علت اینکه در این روش تمام نقاط تصویر انتقال پیدا می‌کنند (حتی نواحی سفید که حاوی اطلاعات نیست)، این الگوریتم زمان بر خواهد بود.

در روش پیشنهادی با توجه به اینکه اعوجاج هر خط بر اساس میزان انحنای همان خط برطرف می‌شود، در مواقعی که در طول تصویر انحنای چندین بار تغییر کند، باز هم اعوجاج هندسی به خوبی برطرف می‌شود. دو مشکل ۱ و ۲ با توجه به تصاویر موجود در پایگاه داده کمتر اتفاق می‌افتد به همین دلیل به خوبی نمی‌توان به صورت شهودی نتایج را مشاهده کرد.

اما مشکل سوم روش شامقلی مربوط به سرعت کم این روش در مقایسه با روش پیشنهادی می‌باشد. با توجه به این نکته که در روش پیشنهادی فقط نقاطی از تصویر که دارای اطلاعات بوده انتقال پیدا می‌کنند، مدت‌زمان اجرا نسبت به روش شامقلی کمتر می‌باشد. این مسئله زمانی که در تصویر خطوط کمتری باشد، بیشتر نشان داده می‌شود. سرعت روش پیشنهادی بر روی تصاویر پایگاه داده‌ی مورد استفاده ۱,۵ برابر سرعت روش شامقلی می‌باشد.

A web publishing framework attempts to address these complicated issues. Just as a web server is responsible for responding to a URL request for a file, a web publishing framework is responsible for responding to a similar request; however, instead of responding with a file, it often will respond with a *published* version of a file. In this case, a published file refers to a file that may have been transformed with XSLT, massaged at an application level, or converted into another format such as a PDF. The requestor does not see the raw data that may underlie the published result, but also does not have to explicitly request that publication occur. Often, a URI base (such as <http://yourHost.com/publish>) signifies that a publishing engine that sits on top of the web server should handle requests. As you may suspect, the concept is much simpler than the actual implementation of a framework like this, and finding the correct framework for your needs is not a trivial task.

Selecting a Framework

You might expect to find a list of hundreds of possible solutions. As you've seen, the Java language offers an easy interface into XML through several APIs. Additionally, Java servlets offer a simple means of handling web requests and responses. However, the list of frameworks is small, and the list of good, stable ones is even smaller. One of the best resources for seeing what products are currently available is XML Software's list at <http://xmlsoftware.com/publishing/>. This list changes so frequently that it is not worth repeating here. Still, some important criteria for determining what framework is right for you are worth mentioning.

Don't be surprised if you (still!) have a hard time finding a product whose version tag is greater than 2.x. In fact, you may have to search diligently to even find a second-generation framework. While a higher version number is not a guarantee of stability, it often reflects the amount of time, effort, and review that a framework has undergone. The XML publishing system is such a new beast that the market has been flooded with 1.0 and 1.1 products that simply are not stable enough for practical use.

You can often ascertain the stability of a product by investigating other products from the same vendor. Often a vendor releases an entire suite of tools; if their other tools do not offer SAX 2.0 and DOM Level 2 support, or are all also 1.0 and 1.1 products, you might be wise to pass on the framework until it has matured and conformed to newer XML standards. Try to steer away from platform-specific technologies. If the framework is tied to a platform (such as Windows, or even a specific flavor of Unix), you aren't dealing with a pure Java solution. Remember that a publishing framework must serve clients on any platform; why use a product that can't run on any platform?

الف

A web publishing framework attempts to address these complicated issues. Just as a web server is responsible for responding to a URL request for a file, a web publishing framework is responsible for responding to a similar request; however, instead of responding with a file, it often will respond with a *published* version of a file. In this case, a published file refers to a file that may have been transformed with XSLT, massaged at an application level, or converted into another format such as a PDF. The requestor does not see the raw data that may underlie the published result, but also does not have to explicitly request that publication occur. Often, a URI base (such as <http://yourHost.com/publish>) signifies that a publishing engine that sits on top of the web server should handle requests. As you may suspect, the concept is much simpler than the actual implementation of a framework like this, and finding the correct framework for your needs is not a trivial task.

Selecting a Framework

You might expect to find a list of hundreds of possible solutions. As you've seen, the Java language offers an easy interface into XML through several APIs. Additionally, Java servlets offer a simple means of handling web requests and responses. However, the list of frameworks is small, and the list of good, stable ones is even smaller. One of the best resources for seeing what products are currently available is XML Software's list at <http://xmlsoftware.com/publishing/>. This list changes so frequently that it is not worth repeating here. Still, some important criteria for determining what framework is right for you are worth mentioning.

Stability

Don't be surprised if you (still!) have a hard time finding a product whose version tag is greater than 2.x. In fact, you may have to search diligently to even find a second-generation framework. While a higher version number is not a guarantee of stability, it often reflects the amount of time, effort, and review that a framework has undergone. The XML publishing system is such a new beast that the market has been flooded with 1.0 and 1.1 products that simply are not stable enough for practical use.

You can often ascertain the stability of a product by investigating other products from the same vendor. Often a vendor releases an entire suite of tools; if their other tools do not offer SAX 2.0 and DOM Level 2 support, or are all also 1.0 and 1.1 products, you might be wise to pass on the framework until it has matured and conformed to newer XML standards. Try to steer away from platform-specific technologies. If the framework is tied to a platform (such as Windows, or even a specific flavor of Unix), you aren't dealing with a pure Java solution. Remember that a publishing framework must serve clients on any platform; why use a product that can't run on any platform?

ج

A web publishing framework attempts to address these complicated issues. Just as a web server is responsible for responding to a URL request for a file, a web publishing framework is responsible for responding to a similar request; however, instead of responding with a file, it often will respond with a *published* version of a file. In this case, a published file refers to a file that may have been transformed with XSLT, massaged at an application level, or converted into another format such as a PDF. The requestor does not see the raw data that may underlie the published result, but also does not have to explicitly request that publication occur. Often, a URI base (such as <http://yourHost.com/publish>) signifies that a publishing engine that sits on top of the web server should handle requests. As you may suspect, the concept is much simpler than the actual implementation of a framework like this, and finding the correct framework for your needs is not a trivial task.

Selecting a Framework

You might expect to find a list of hundreds of possible solutions. As you've seen, the Java language offers an easy interface into XML through several APIs. Additionally, Java servlets offer a simple means of handling web requests and responses. However, the list of frameworks is small, and the list of good, stable ones is even smaller. One of the best resources for seeing what products are currently available is XML Software's list at <http://xmlsoftware.com/publishing/>. This list changes so frequently that it is not worth repeating here. Still, some important criteria for determining what framework is right for you are worth mentioning.

Stability

Don't be surprised if you (still!) have a hard time finding a product whose version tag is greater than 2.x. In fact, you may have to search diligently to even find a second-generation framework. While a higher version number is not a guarantee of stability, it often reflects the amount of time, effort, and review that a framework has undergone. The XML publishing system is such a new beast that the market has been flooded with 1.0 and 1.1 products that simply are not stable enough for practical use.

You can often ascertain the stability of a product by investigating other products from the same vendor. Often a vendor releases an entire suite of tools; if their other tools do not offer SAX 2.0 and DOM Level 2 support, or are all also 1.0 and 1.1 products, you might be wise to pass on the framework until it has matured and conformed to newer XML standards. Try to steer away from platform-specific technologies. If the framework is tied to a platform (such as Windows, or even a specific flavor of Unix), you aren't dealing with a pure Java solution. Remember that a publishing framework must serve clients on any platform; why use a product that can't run on any platform?

ب

A web publishing framework attempts to address these complicated issues. Just as a web server is responsible for responding to a URL request for a file, a web publishing framework is responsible for responding to a similar request; however, instead of responding with a file, it often will respond with a *published* version of a file. In this case, a published file refers to a file that may have been transformed with XSLT, massaged at an application level, or converted into another format such as a PDF. The requestor does not see the raw data that may underlie the published result, but also does not have to explicitly request that publication occur. Often, a URI base (such as <http://yourHost.com/publish>) signifies that a publishing engine that sits on top of the web server should handle requests. As you may suspect, the concept is much simpler than the actual implementation of a framework like this, and finding the correct framework for your needs is not a trivial task.

Selecting a Framework

You might expect to find a list of hundreds of possible solutions. As you've seen, the Java language offers an easy interface into XML through several APIs. Additionally, Java servlets offer a simple means of handling web requests and responses. However, the list of frameworks is small, and the list of good, stable ones is even smaller. One of the best resources for seeing what products are currently available is XML Software's list at <http://xmlsoftware.com/publishing/>. This list changes so frequently that it is not worth repeating here. Still, some important criteria for determining what framework is right for you are worth mentioning.

Stability

Don't be surprised if you (still!) have a hard time finding a product whose version tag is greater than 2.x. In fact, you may have to search diligently to even find a second-generation framework. While a higher version number is not a guarantee of stability, it often reflects the amount of time, effort, and review that a framework has undergone. The XML publishing system is such a new beast that the market has been flooded with 1.0 and 1.1 products that simply are not stable enough for practical use.

You can often ascertain the stability of a product by investigating other products from the same vendor. Often a vendor releases an entire suite of tools; if their other tools do not offer SAX 2.0 and DOM Level 2 support, or are all also 1.0 and 1.1 products, you might be wise to pass on the framework until it has matured and conformed to newer XML standards. Try to steer away from platform-specific technologies. If the framework is tied to a platform (such as Windows, or even a specific flavor of Unix), you aren't dealing with a pure Java solution. Remember that a publishing framework must serve clients on any platform; why use a product that can't run on any platform?

د

شکل ۵-۱۱: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ (الف) روش SEG (ب) روش SKEL

(ج) روش شامقلی (د) روش پیشنهادی

For HTML, DOM provides a set of interfaces that model the various HTML elements. For example, you can use the `HTMLDocument` class, the `HTMLAnchorElement`, and the `HTMLSelectElement` (all in the `org.w3c.dom.html` package) to represent their analogs in HTML (`<HTML>`, `<A>`, and `<SELECT>` in this case). All of these provide convenience methods like `setTitle()` (on `HTMLDocument`), `setHref()` (on `HTMLAnchorElement`), and `getOptions()` (on `HTMLSelectElement`). All of these extend core DOM structures like `Document` and `Element`, and so can be used as any other `DOMNode` could.

However, it turns out that the HTML bindings are rarely used (at least directly). It's not because they aren't useful; instead, many tools have already been written to provide this sort of access through even more user-friendly tools. XMLC, a project within the Enhydra application server framework, is one such example (located online at <http://xmlc.enhydra.org>), and Cocoon, covered in Chapter 10, is another. These allow developers to work with HTML and web pages in a way that does not necessarily require even basic DOM knowledge, making it more accessible to web designers and newer Java developers. The end result of using these tools is that the HTML DOM bindings are rarely needed. But if you know about them, you can use them if you need to. Additionally, you can use standard DOM functionality on well-formed HTML documents (XHTML), treating elements as `Element` nodes and attributes as `Attr` nodes. Even without the HTML bindings, you can use DOM to work with HTML. Piece of cake.

Odds and Ends

What's left in DOM Level 2 besides these modules and namespace awareness? Very little, and you've probably already used most of it. The `createDocument()` and `createDocumentType()` methods are new to the `DOMImplementation` class, and you've used both of them. Additionally, the `getSystemId()` and `getPublicId()` methods used in the `DOMSerializer` class on the `DocumentType` interface are also DOM Level 2 additions. Other than that, there isn't much; a few new `DOMException` error code constants, and that's about it. You can see the complete list of changes online at <http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/changes.html>. The rest of the changes are the additional modules, one of which I'll cover next.

DOM Level 3

Before closing the book on DOM and looking at common gotchas, I will spend a little time letting you know what's coming in DOM Level 3, which is underway right now. In fact, I expect this specification to be finalized early in 2002, not long

For HTML, DOM provides a set of interfaces that model the various HTML elements. For example, you can use the `HTMLDocument` class, the `HTMLAnchorElement`, and the `HTMLSelectElement` (all in the `org.w3c.dom.html` package) to represent their analogs in HTML (`<HTML>`, `<A>`, and `<SELECT>` in this case). All of these provide convenience methods like `setTitle()` (on `HTMLDocument`), `setHref()` (on `HTMLAnchorElement`), and `getOptions()` (on `HTMLSelectElement`). All of these extend core DOM structures like `Document` and `Element`, and so can be used as any other `DOMNode` could.

However, it turns out that the HTML bindings are rarely used (at least directly). It's not because they aren't useful; instead, many tools have already been written to provide this sort of access through even more user-friendly tools. XMLC, a project within the Enhydra application server framework, is one such example (located online at <http://xmlc.enhydra.org>), and Cocoon, covered in Chapter 10, is another. These allow developers to work with HTML and web pages in a way that does not necessarily require even basic DOM knowledge, making it more accessible to web designers and newer Java developers. The end result of using these tools is that the HTML DOM bindings are rarely needed. But if you know about them, you can use them if you need to. Additionally, you can use standard DOM functionality on well-formed HTML documents (XHTML), treating elements as `Element` nodes and attributes as `Attr` nodes. Even without the HTML bindings, you can use DOM to work with HTML. Piece of cake.

Odds and Ends

What's left in DOM Level 2 besides these modules and namespace awareness? Very little, and you've probably already used most of it. The `createDocument()` and `createDocumentType()` methods are new to the `DOMImplementation` class, and you've used both of them. Additionally, the `getSystemId()` and `getPublicId()` methods used in the `DOMSerializer` class on the `DocumentType` interface are also DOM Level 2 additions. Other than that, there isn't much; a few new `DOMException` error code constants, and that's about it. You can see the complete list of changes online at <http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/changes.html>. The rest of the changes are the additional modules, one of which I'll cover next.

DOM Level 3

Before closing the book on DOM and looking at common gotchas, I will spend a little time letting you know what's coming in DOM Level 3, which is underway right now. In fact, I expect this specification to be finalized early in 2002, not long

الف

For HTML, DOM provides a set of interfaces that model the various HTML elements. For example, you can use the `HTMLDocument` class, the `HTMLAnchorElement`, and the `HTMLSelectElement` (all in the `org.w3c.dom.html` package) to represent their analogs in HTML (`<HTML>`, `<A>`, and `<SELECT>` in this case). All of these provide convenience methods like `setTitle()` (on `HTMLDocument`), `setHref()` (on `HTMLAnchorElement`), and `getOptions()` (on `HTMLSelectElement`). All of these extend core DOM structures like `Document` and `Element`, and so can be used as any other `DOMNode` could.

However, it turns out that the HTML bindings are rarely used (at least directly). It's not because they aren't useful; instead, many tools have already been written to provide this sort of access through even more user-friendly tools. XMLC, a project within the Enhydra application server framework, is one such example (located online at <http://xmlc.enhydra.org>), and Cocoon, covered in Chapter 10, is another. These allow developers to work with HTML and web pages in a way that does not necessarily require even basic DOM knowledge, making it more accessible to web designers and newer Java developers. The end result of using these tools is that the HTML DOM bindings are rarely needed. But if you know about them, you can use them if you need to. Additionally, you can use standard DOM functionality on well-formed HTML documents (XHTML), treating elements as `Element` nodes and attributes as `Attr` nodes. Even without the HTML bindings, you can use DOM to work with HTML. Piece of cake.

Odds and Ends

What's left in DOM Level 2 besides these modules and namespace awareness? Very little, and you've probably already used most of it. The `createDocument()` and `createDocumentType()` methods are new to the `DOMImplementation` class, and you've used both of them. Additionally, the `getSystemId()` and `getPublicId()` methods used in the `DOMSerializer` class on the `DocumentType` interface are also DOM Level 2 additions. Other than that, there isn't much; a few new `DOMException` error code constants, and that's about it. You can see the complete list of changes online at <http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/changes.html>. The rest of the changes are the additional modules, one of which I'll cover next.

DOM Level 3

Before closing the book on DOM and looking at common gotchas, I will spend a little time letting you know what's coming in DOM Level 3, which is underway right now. In fact, I expect this specification to be finalized early in 2002, not long

ج

For HTML, DOM provides a set of interfaces that model the various HTML elements. For example, you can use the `HTMLDocument` class, the `HTMLAnchorElement`, and the `HTMLSelectElement` (all in the `org.w3c.dom.html` package) to represent their analogs in HTML (`<HTML>`, `<A>`, and `<SELECT>` in this case). All of these provide convenience methods like `setTitle()` (on `HTMLDocument`), `setHref()` (on `HTMLAnchorElement`), and `getOptions()` (on `HTMLSelectElement`). All of these extend core DOM structures like `Document` and `Element`, and so can be used as any other `DOMNode` could.

However, it turns out that the HTML bindings are rarely used (at least directly). It's not because they aren't useful; instead, many tools have already been written to provide this sort of access through even more user-friendly tools. XMLC, a project within the Enhydra application server framework, is one such example (located online at <http://xmlc.enhydra.org>), and Cocoon, covered in Chapter 10, is another. These allow developers to work with HTML and web pages in a way that does not necessarily require even basic DOM knowledge, making it more accessible to web designers and newer Java developers. The end result of using these tools is that the HTML DOM bindings are rarely needed. But if you know about them, you can use them if you need to. Additionally, you can use standard DOM functionality on well-formed HTML documents (XHTML), treating elements as `Element` nodes and attributes as `Attr` nodes. Even without the HTML bindings, you can use DOM to work with HTML. Piece of cake.

Odds and Ends

What's left in DOM Level 2 besides these modules and namespace awareness? Very little, and you've probably already used most of it. The `createDocument()` and `createDocumentType()` methods are new to the `DOMImplementation` class, and you've used both of them. Additionally, the `getSystemId()` and `getPublicId()` methods used in the `DOMSerializer` class on the `DocumentType` interface are also DOM Level 2 additions. Other than that, there isn't much; a few new `DOMException` error code constants, and that's about it. You can see the complete list of changes online at <http://www.w3.org/TR/2000/REC-DOM-Level-2-Core-20001113/changes.html>. The rest of the changes are the additional modules, one of which I'll cover next.

DOM Level 3

Before closing the book on DOM and looking at common gotchas, I will spend a little time letting you know what's coming in DOM Level 3, which is underway right now. In fact, I expect this specification to be finalized early in 2002, not long

د

شکل ۵-۲: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ پ الف) روش SEG ب) روش SKEL

ج) روش شامق‌لی د) روش پیشنهادی

اما از جمله مشکلات روش پیشنهادی می‌توان به دو مورد اشاره کرد که در آینده سعی می‌شود این مشکلات برطرف گردد.

- ۱- عدم توانایی رفع اعوجاج هندسی اسناد دارای اجزای گرافیکی
- ۲- تشخیص اشتباه نقاط جهت تشخیص انحنا در برخی از نوارها که منجر به تشخیص اشتباه انحنای خط می‌شود.

به منظور حل مشکل اول پیشنهاد می‌شود که بر اساس روش‌های موجود ابتدا اجزای گرافیکی از تصویر حذف شده و پس از رفع اعوجاج جزء حذف شده به مکان خود بازگردد. در ضمن به منظور رفع اعوجاج جزء گرافیکی نیز می‌توان این جزء را به صورت جداگانه رفع اعوجاج نمود. همچنین به منظور حل مشکل دوم نیز پیشنهاد می‌شود، طول نوار در مواقعی که تعداد حروف کمی در آن قرار گرفته به صورت وقتی افزایش پیدا کند.

به منظور مقایسه بهتر از نرم‌افزار OmniPage کمک گرفته و دقت OCR برای روش‌های ذکر شده بررسی می‌گردد. این نرم افزار توسط شرکت ارتباطات نانس^۱ ارائه گردیده است. در حال حاضر این نرم‌افزار بیش از ۱۴۰ زبان مختلف دنیا را پشتیبانی می‌کند. برای آنالیز متون انگلیسی از این نرم‌افزار استفاده خواهیم کرد. به منظور محاسبه دقت OCR روش پیشنهادی، از مجموعه‌ای از تصاویری که در پایگاه داده‌ی DFKI بوده، استفاده شده است. در انتخاب این تصاویر سعی شده است انواع حالات اعوجاج هندسی در نظر گرفته شود. به منظور محاسبه کمی از روش که در [۶۷] برای تعیین صحت کلمات^۲ معرفی شده استفاده خواهیم کرد. این روش به صورت رابطه ۵-۱ بیان می‌شود.

$$Word Accuracy = \frac{\#Words - \#Misrecognized_Words}{\#Words} \quad (۱-۵)$$

که در این رابطه $\#Words$ تعداد کل کلمات موجود در متن و $\#Misrecognized_Words$ تعداد کلماتی است که نرم‌افزار OCR در تشخیص آن ناتوان بوده است. در جدول ۴-۵ نتایج روش پیشنهادی به همراه نتایج سایر روش‌ها نشان داده شده‌اند. از بین ۱۱۱۳۷ کلمه، تعداد ۸۷ کلمه در روش پیشنهادی اشتباه تشخیص داده شده است. همانطور که از روی نتایج پیش‌بینی می‌شد، دقت روش پیشنهادی از دقت دو روش SEG و SKEL بیشتر می‌باشد اما این دقت از دقت روش شامقلی کمتر می‌باشد. البته با رفع مشکلات روش پیشنهادی که پیش‌تر ذکر شد، می‌توان این دقت را افزایش داد.

¹ Nuance Communication

² Word Accuracy

جدول ۵-۴: بررسی درصد دقت صحت کلمات

درصد دقت صحت کلمات	روش
۹۰,۹۳۸۰	بدون بازیابی
۹۴,۹۸۱۱	روش SEG
۹۸,۶۳۲۴	روش SKEL
۹۹,۳۶۲۲	روش شامقلی
۹۹,۲۱۸۸	روش پیشنهادی

۵-۳-۳- مقایسه‌ی نتایج بر روی اسناد فارسی

به‌منظور مقایسه‌ی بهتر، نتایج روش پیشنهادی همراه با نتایج دو روش شامقلی [۴۷] و خسروی‌راد [۴۶] در تصاویر ۵-۱۳ تا ۵-۱۵ نشان داده‌شده است.

همان‌طور که قبلاً اشاره‌شده است در روش خسروی‌راد هر کلمه بر اساس میزان چرخش آن به‌صورت جدا رفع اعوجاج می‌شود، البته تصویر خروجی این الگوریتم می‌تواند خاکستری با دوسطحی باشد. ازجمله مشکلات این روش همان‌طور که در تصاویر ۵-۱۳ تا ۵-۱۵ مشاهده می‌کنید؛ دقت پایین در رفع اعوجاج هندسی و جابجایی بیش‌ازحد حروف و نقاط از کلمه می‌باشد.

مشکلات روش شامقلی نیز که در بخش قبلی هم موردبررسی قرار گرفت، عبارت بودند از:

۱- زمانی که تصاویر دارای خطوط غالب در بالا یا پایین تصویر نباشد، اعوجاج هندسی به‌خوبی برطرف نمی‌شود.

۲- با توجه به اینکه رفع اعوجاج هندسی طبق این روش بر اساس میزان انحنای دو خط غالب در تصویر می‌باشد، در مواقعی که در طول تصویر انحنای چندین بار تغییر کند، اعوجاج هندسی به‌خوبی برطرف نمی‌شود.

۳- به علت اینکه در این روش تمام نقاط تصویر انتقال پیدا می‌کنند (حتی نواحی سفید که حاوی اطلاعات نیست)، این الگوریتم زمان‌بر خواهد بود.

مهم‌ترین مشکل روش شامقلی در تصاویر پایگاه داده خسروی‌راد زمان‌گیر بودن این روش می‌باشد. در تصاویر پایگاه داده‌ی خسروی‌راد به علت درجه‌ی تفکیک‌پذیری بالای تصاویر و همچنین تعداد پیکسل‌های

زیاد سفیدرنگ سرعت روش شامقلی نسبت به روش پیشنهادی کم تر می باشد، این نتیجه در جدول ۵-۵ نشان داده شده است.

در روش شامقلی جداسازی خطوط توسط دو روش^۱ DLCM و استفاده از عملگر گسترش انجام می شود، البته مدت زمان اجرای روش شامقلی با الگوریتم DLCM همان طور که در جدول ۵-۵ مقایسه می کنید؛ بسیار بیشتر از مدت زمان اجرای روش شامقلی با عملگر گسترش می باشد. باین حال روش پیشنهادی سریع تر از هر دو روش شامقلی می باشد.

جدول ۵-۵: مدت زمان اجرای روش شامقلی و روش پیشنهادی

روش	مدت زمان اجرا (بر حسب ثانیه)
روش شامقلی به کمک الگوریتم DLCM	۴۳۰
روش شامقلی به کمک عملگر گسترش	۲۷۰
روش پیشنهادی	۱۴۰

به منظور محاسبه کمی درصد صحت کلمات بر روی مجموعه تصاویر پایگاه داده ی خسروی راد محاسبه شده است. از بین ۲۹۸۸۰ کلمه تعداد ۴۴۸۶ کلمه اشتباه تشخیص داده شده است. با توجه به نتایج، مشخص است که روش پیشنهادی از دقت قابل قبولی برخوردار می باشد. در ضمن می توان با برطرف کردن مشکلات روش پیشنهادی که قبلاً ذکر شده است دقت این روش را افزایش داد.

جدول ۵-۶: بررسی درصد دقت صحت کلمات

روش	درصد دقت صحت کلمات
روش خسروی راد	۶۷,۸۱۷۹۳
روش شامقلی	۸۶,۷۳۰۲۵
روش پیشنهادی	۸۴,۹۸۶۶

¹ Dynamic Local Connectivity Map

هنگام عروسی محمدرضا شاه پهلوی و زن مصری اش فوزیه در سال ۱۳۱۷ خورشیدی چون مقرر بود، میهمانان مصری و همراهان عروس به وسیله راه آهن جنوب به تهران وارد شوند از طرف دربار و شهربانی دستور اکید صادر شده بود که دیوارهای تمام دهات طول راه و خانه های دهقانی مجاور خط آهن را سفید کنند.

در یکی از دهات چون گچ در دسترس نبود بخشدار دستور می‌دهد که با کشک و ماست که در آن ده فراوان بوده دیوارها را موقتاً سفید نمایند. و به این منظور با پولی که از کدخدای ده می‌گیرند و با خرید مقدار زیادی ماست کلیه دیوارها را ماست مالی کردند.

قدمت تاریخی این اصطلاح ماست مالی از شصت سال نمی‌گذرد، و ماجرای این ماست مالی مدت ها موضوع اصلی شوخی‌های محافل و مجالس بود. حکیم ارد بزرگ می گوید: فرمانروایان تنها پاسخگوی زمان حال خویش نیستند آنها به گذشتگان و آیندگان نیز پاسخگو هستند.

هنگام عروسی محمدرضا شاه پهلوی و زن مصری اش فوزیه در سال ۱۳۱۷ خورشیدی چون مقرر بود، میهمانان مصری و همراهان عروس به وسیله راه آهن جنوب به تهران وارد شوند از طرف دربار و شهربانی دستور اکید صادر شده بود که دیوارهای تمام دهات طول راه و خانه های دهقانی مجاور خط آهن را سفید کنند.

در یکی از دهات چون گچ در دسترس نبود بخشدار دستور می‌دهد که با کشک و ماست که در آن ده فراوان بوده دیوارها را موقتاً سفید نمایند، و به این منظور با پولی که از کدخدای ده می‌گیرند و با خرید مقدار زیادی ماست کلیه دیوارها را ماست مالی کردند.

قدمت تاریخی این اصطلاح ماست مالی از شصت سال نمی‌گذرد، و ماجرای این ماست مالی مدت ها موضوع اصلی شوخی‌های محافل و مجالس بود. حکیم ارد بزرگ می گوید: فرمانروایان تنها پاسخگوی زمان حال خویش نیستند آنها به گذشتگان و آیندگان نیز پاسخگو هستند.

الف

هنگام عروسی محمدرضا شاه پهلوی و زن مصری اش فوزیه در سال ۱۳۱۷ خورشیدی چون مقرر بود، میهمانان مصری و همراهان عروس به وسیله راه آهن جنوب به تهران وارد شوند از طرف دربار و شهربانی دستور اکید صادر شده بود که دیوارهای تمام دهات طول راه و خانه های دهقانی مجاور خط آهن را سفید کنند.

در یکی از دهات چون گچ در دسترس نبود بخشدار دستور می‌دهد که با کشک و ماست که در آن ده فراوان بوده دیوارها را موقتاً سفید نمایند، و به این منظور با پولی که از کدخدای ده می‌گیرند و با خرید مقدار زیادی ماست کلیه دیوارها را ماست مالی کردند.

قدمت تاریخی این اصطلاح ماست مالی از شصت سال نمی‌گذرد، و ماجرای این ماست مالی مدت ها موضوع اصلی شوخی‌های محافل و مجالس بود. حکیم ارد بزرگ می گوید: فرمانروایان تنها پاسخگوی زمان حال خویش نیستند آنها به گذشتگان و آیندگان نیز پاسخگو هستند.

ج

شکل ۵-۱۳: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ (الف) روش خسروی‌راد (ب) روش شامقلی (ج) روش پیشنهادی

لذت بردن می‌خوانم :جاذبه‌ی همان دنیای دیگری که بیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. من شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم، برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین‌ترتیب به روح شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکرم را، علائقم را، رویاهایم را و افق‌هایی را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، رویاها و

لذت بردن می‌خوانم :جاذبه‌ی همان دنیای دیگری که بیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. من شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم، برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین‌ترتیب به روح شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکرم را، علائقم را، رویاهایم را و افق‌هایی را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، رویاها و

الف

لذت بردن می‌خوانم :جاذبه‌ی همان دنیای دیگری که بیشتر یاد کردم. شاید بتوان اسم این کار را فرار از واقعیت گذاشت. آدم حتی اگر بتواند در عالم خیال از غصه‌های زندگی روزمره فرار کند و زمانی را در دنیای دیگر بگذراند خوب است. من شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بسازم، برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین‌ترتیب به روح شکل بدهم. در واقع، می‌خواستم بدانم باید چه جور آدمی بشوم؟ معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکرم را، علائقم را، رویاهایم را و افق‌هایی را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، رویاها و

ب

ج

شکل ۵-۱۴: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ ج الف) روش خسروی‌راد ب) روش شامقلی ج) روش پیشنهادی

خواننده کتاب‌هایی که نه برای کار یا برای خودسازی
و آموختن، بلکه فقط برای لذت بردن می‌فوانم :
مادیهی همان دنیای دیگری که پیشتر یاد کردم.
شاید بتوان اسم این کار را فرار از واقعیت گذاشت.
آدم می‌تواند در عالم خیال از غصه‌های زندگی
روزمره فرار کند و زمانی را در دنیای دیگر بگذراند
خوب است. این شانزده تا بیست و شش سالگی.
فواندن برای من امری می‌تواند بود برای اینکه بتوانم
خودم را بسازم. برای خودم کسی بشوم. آکامی‌مایم
را بیشتر کنم و بدین ترتیب به رومم شکل بدم. در
واقع، می‌فواستم بدانم باید چه جور آدمی بشوم؟
معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکره

الف

خواننده کتاب‌هایی که نه برای کار یا برای خودسازی
و آموختن، بلکه فقط برای لذت بردن می‌فوانم :
مادیهی همان دنیای دیگری که پیشتر یاد کردم.
شاید بتوان اسم این کار را فرار از واقعیت گذاشت.
آدم می‌تواند در عالم خیال از غصه‌های زندگی
روزمره فرار کند و زمانی را در دنیای دیگر بگذراند
خوب است. این شانزده تا بیست و شش سالگی.
فواندن برای من امری می‌تواند بود برای اینکه بتوانم
خودم را بسازم. برای خودم کسی بشوم. آکامی‌مایم
را بیشتر کنم و بدین ترتیب به رومم شکل بدم. در
واقع، می‌فواستم بدانم باید چه جور آدمی بشوم؟
معنای زندگی و دنیا چیست؟ چقدر می‌توانم فکره

ج

شکل ۵-۱۵: نتیجه‌ی اعمال روش‌های مختلف بر روی تصویر ۵-۹ (د الف) روش خسروی راد (ب) روش شامقلی
(ج) روش پیشنهادی

فصل ششم

نتیجه‌گیری و پیشنهادات

۱-۶- مقدمه

در این فصل یک جمع‌بندی کلی از پایان‌نامه ارائه‌شده و به‌صورت مختصر به بیان کارهایی که انجام شده است، پرداخته می‌شود و در پایان به بیان کارهایی که در آینده انجام خواهد پذیرفت، می‌پردازیم.

۲-۶- نتیجه‌گیری

در این پایان‌نامه هدف رفع اعوجاج فتومتریک و اعوجاج هندسی از تصاویر است. در این پایان‌نامه به‌منظور تصحیح اعوجاج فتومتریک از روش درون‌نگاری هارمونیک استفاده شده است. به‌منظور افزایش دقت این الگوریتم پیشنهاد نموده‌ایم تا از همسایه‌های قطری هر پیکسل نیز استفاده شود؛ این عمل موجب بهبود تخمین روشنایی پس‌زمینه می‌گردد. همچنین برای کاهش پیچیدگی محاسباتی در نتیجه بکارگیری همسایه‌های قطری پیشنهاد نمودیم تا در اجرای الگوریتم درون‌نگاری یک‌بار از همسایه‌های چهارتایی اصلی و بار دیگر از همسایه‌های قطری استفاده نماییم. با این روش پیچیدگی محاسباتی روش پیشنهادی نیز افزایش نمی‌یابد. همچنین به‌منظور بهبود کیفیت در تخمین پس‌زمینه از فیلترهای ملایم‌کننده‌ی گوسی و میانگین‌گیر نیز استفاده شده است.

همچنین به‌منظور تصحیح اعوجاج هندسی پس از استخراج خطوط از تصویر، هر خط به چندین نوار عمودی تقسیم می‌شود. در هر نوار مختصات نقطه‌ای که مقدار افکنش افقی آن بیشینه است؛ ذخیره می‌شود. انحنای هر خط به کمک این مجموعه نقاط توسط تابع درجه‌ی سوم مشخص می‌گردد. در نهایت با استفاده از تخمین تبدیل پرسپکتیو، اعوجاج هندسی هر خط برطرف می‌گردد. همچنین به‌منظور تشخیص انحنای خطوط با طول کم از انحنای خطوط با طول بزرگ‌تر قبلی یا بعدی استفاده می‌شود.

بر اساس نتایج هر دو روش پیشنهادی از کارایی خوبی برخوردار بوده و دقت نتایج قابل قبول است.

کارهای انجام‌شده در این پایان‌نامه به‌منظور رفع اعوجاج فتومتریک را می‌توان به‌صورت زیر خلاصه کرد:

استفاده از درون‌نگاری هارمونیک به کمک هر هشت همسایه که به‌منظور کاهش پیچیدگی محاسباتی در نتیجه بکارگیری همسایه‌های قطری پیشنهاد نمودیم تا در اجرای الگوریتم درون‌نگاری یک‌بار از همسایه‌های چهارتایی اصلی و بار دیگر از همسایه‌های قطری استفاده نماییم.

همچنین با استفاده از ملایم کننده‌های میانگین گیر و گوسی به جای ملایم کننده‌ی RBF که به خصوص برای تصاویر با قدرت تفکیک پذیری بالا زمان بر است؛ سرعت روش پیشنهادی افزایش یافته است.

روش پیشنهادی به منظور رفع اعوجاج فتومتریک بر روی پایگاه‌های داده‌ی شامقلی و خسروی راد اعمال شده است. این نتایج با سه روش اتسو، نیبلک و ژانگ مقایسه شده است؛ که این نتایج بیانگر قدرت و دقت روش پیشنهادی است.

کارهای انجام شده در این پایان نامه به منظور رفع اعوجاج هندسی را نیز می توان به صورت زیر خلاصه کرد: پیدا کردن خطوط در تصویر که به کمک عملگر گسترش این کار صورت پذیرفته است.

تقسیم هر خط به نوارهای عمودی و پیدا کردن زاویه چرخش بر اساس بیشینه مقدار افکنش افقی تابع شدت روشنایی و محاسبه‌ی مختصات بیشینه مقدار افکنش افقی تابع گرادیان تصویر چرخانده شده در تصویری با ابعاد تصویر اولیه.

به منظور به دست آوردن انحنای هر خط با استفاده از مجموعه نقاط به دست آمده از ماتریس وندرموند به منظور تخمین انحنای هر خط به کمک منحنی درجه سه استفاده شده است. در این مرحله به منظور تخمین انحنای خطوط با طول کم که انحنای آن‌ها به درستی تخمین زده نشده است از انحنای خطوط بالایی یا پایینی آن خط استفاده شده است.

هر خط بین دو منحنی با انحنای محاسبه شده قرار گرفته و فاصله‌ی بین این دو منحنی به تعداد چهارضلعی تقسیم شده است سپس هر چهارضلعی به کمک روش تخمین تبدیل پرسپکتیو به یک مستطیل تبدیل می گردد و در نهایت اعوجاج هندسی برطرف می گردد.

همواره پس از انتقال نقاطی وجود دارند که مقداردهی نشده اند؛ برای همپوشانی این نقاط انتقال از آنالیز عمودی و افقی پیکسل‌های تصویر استفاده شده است.

در این پایان نامه اعوجاج هندسی اسناد انگلیسی و فارسی دو پایگاه داده‌ی خسروی راد و DFKI به کمک روش پیشنهادی برطرف شده است و نتایج با سه روش SEG، SKEL و شامقلی به صورت کمی و کیفی مقایسه شده است؛ که این نتایج نشان دهنده‌ی دقت و سرعت قابل قبول روش پیشنهادی در بین این روش‌ها است.

۳-۶- کارهای آینده

با توجه به الگوریتم پیشنهادی در پایان نامه، کارهای زیر را در آینده انجام خواهیم داد:

استفاده از سایر روش‌های درون‌نگاری به جای درون‌نگاری هارمونیک و مقایسه نتایج این روش‌ها با روش پیشنهادی از لحاظ دقت و سرعت تخمین پس‌زمینه.

ارزیابی روش پیشنهادی بر روی خطوط دست‌نویس فارسی و انگلیسی و مقایسه‌ی نتایج با روش‌های رایج؛ که سعی می‌شود دو نوع اعوجاج فتومتریک و هندسی این اسناد برطرف گردد.

ارتقای روش پیشنهادی برای رفع اعوجاج هندسی اسناد دارای اجزای گرافیکی با استفاده از روش‌های رایج استخراج اجزای گرافیکی از تصویر و برطرف کردن اعوجاج هندسی این اسناد.

ارتقای روش پیشنهادی برای رفع اعوجاج هندسی اسناد دو ستونه که برای این کار ابتدا باید خطوط به‌درستی تشخیص داده شوند و پس از آن اعوجاج هندسی هر خط برطرف گردد.

افزایش وفقی طول هر نوار در مواقعی که تعداد حروف کمی در آن نوار قرار گرفته باشد. برای این کار زمانی که تعداد پیکسل‌های مشکلی در یک نوار از یک مقدار آستانه‌ای کمتر باشد طول نوار به‌صورت خودکار افزایش می‌یابد. این کار می‌تواند به تشخیص دقیق‌تر انحناهای هر خط کمک کند.

مراجع

- [1] A. Rosenfeld and P. Dela Torre, "Histogram concavity analysis as an aid in threshold selection," IEEE Trans, 1983.
- [2] M. I. Sezan, "A peak detection algorithm and its application to histogram-based image data," Image Process, vol. 29, p. 47-59, 1985.
- [3] N. Ramesh, J. H. Yoo and I. K. Sethi, "Thresholding based on histogram approximation," Vision Image Signal Process, vol. 142.5, pp. 271-279, 1995.
- [4] T. Kampke and R. Kober, "Nonparametric optimal binarization," Pattern Recognition, pp. 27-29, 1998.
- [5] T. w. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," IEEE Trans, Vols. SMC-8, pp. 630-632, 1978.
- [6] N. Otsu, "A threshold selection method from gray level histograms," IEEE Trans, Vols. SMC-9, pp. 62-66, 1979.
- [7] D. E. Lloyd, "Automatic target classification using moment invariant of image shapes," in RAE IDN AW126, Farnborough, UK, 1985.
- [8] J. Kittler and J. Illingworth, "Minimum error thresholding," Pattern Recognition, vol. 19, pp. 41-47, 1986.
- [9] S. Cho, R. Haralick and S. Yi, "Improvement of Kittler and Illingworths's minimum error thresholding," Pattern Recogn, p. 609-617, 1989.
- [10] J. N. Kapur, P. K. Sahoo and A. K. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," Image Process, vol. 29, pp. 273-285, 1985.
- [11] C. H. Li and C. Lee, "Minimum cross-entropy thresholding," Pattern Recogn, vol. 26, pp. 617-625, 1993.
- [12] C. H. Li and P. S. Tam, "An iterative algorithm for minimum cross-entropy thresholding," Pattern Recognition, vol. 19, pp. 771-776, 1998.
- [13] L. K. Huang and M. J. Wang, "Image thresholding by minimizing the measures of fuzziness," Pattern recognition, Elsevier, vol. 28, pp. 41-51, 1995.

- [14] W. H. Tsai, "Moment-preserving thresholding: A new approach," *Graph. Models Image Process*, vol. 19, p. 377–393, 1985.
- [15] L. Hertz and R. W. Schafer, "Multilevel thresholding using edge matching," *Comput. Vis. Graph. Image Process*, vol. 44, pp. 279-295, 1988.
- [16] R. L. Kirby and A. Rosenfeld, "A note on the use of gray level, local average gray level space as an aid in threshold selection," *IEEE Trans, Vols. SMC-9*, pp. 860-864, 1979.
- [17] C. Chang, K. Chen, J. Wang and M. G. Althouse, "A relative entropy based approach in image thresholding," *Pattern Recogn*, vol. 27, p. 1275–1289, 1994.
- [18] W. Niblack, "An introduction to digital image processing," in Prentice-Hall, Englewood Cliffs, 1986.
- [19] B. Gatos, I. Pratikakis and S. J. Perantonis, "Adaptive degraded document image binarization," *Pattern recognition*, vol. 15, pp. 1544-1554, 2006.
- [20] S. Bukhari, F. Shafait and T. Breuel, "Foreground-Background Regions Guided Binarization of Camera-Captured Document Images," in *Proceedings of the Third International Workshop on Camera Based Document Analysis and Recognition, Barcelona, 2009*.
- [21] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, vol. 32, pp. 112-117, 2000.
- [22] J. M. White and G. D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Application Requiring Character Image Extraction," *IBM Journal on Research and Development*, vol. 27, pp. 400-411, 1983.
- [23] A. Jain, "Fundamentals of Digital Image Processing," in Prentice-Hall, Englewood Cliffs, NJ,, 1989.
- [24] A. Masalovitch and L. Mestetskiy, "Usage of continuous skeletal image representation for document images dewarping," in *2nd Int. Workshop on Camera-Based Document Analysis and Recognition, Curitiba, Brazil, 2007*.
- [25] M. Bertalmio, G. Sapiro, C. Ballester and V. Caselle, "Image inpainting," *SIGGRAPH*, vol. 1, p. 417–424, 2000.
- [26] A. Criminisi, P. Perez and K. Toyama, "Region Filling and Object Removal by Exemplar-Based Image Inpainting," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 13, 2004.
- [27] K. Timothy and R. Chang, "Digital Inpainting – Survey and Multilayer Image Inpainting Algorithms," in *Proceedings of the Third International Conference on Information Technology and Applications, 2005*.

- [28] L. Zhang, A. M. Yip and C. Tan, "A restoration framework for correcting photometric and geometric distortions in camera-based document images," in International Conference on Computer Vision, 2007.
- [29] L. Zhang, A. M. Yip, M. S. Brown and C. Tan, "A unified framework for document restoration using Inpainting and shape-from-shading," Pattern Recognition, p. 2961 – 2978, 2009.
- [30] V. Janarthanan and G. Jananii, "A Detailed Survey on Various Image Inpainting Techniques," Bonfring International Journal of Advances in Image Processing, vol. 2, 2012.
- [31] S. Ravi, P. Pasupathi, S. Muthukumar and N. Krishnan, "Image In-Painting Techniques - A Survey and Analysis," in Innovations in Information Technology (IIT), 9th International Conference, 2013.
- [32] M. Pilu, "Undoing Page Curl Distortion Using Applicable Surfaces," Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 67-72, 2001.
- [33] A. UlgeS, C. Lampert and C. M. Breuel, "Document capture using stereo vision," in In Proceedings of the ACM Symposium on Document Engineering, Wisconsin, USA, 2004.
- [34] H. Cao, X. Ding and C. Liu, "A cylindrical surface model to rectify the bound document image," Computer Vision, pp. 228-233, 2003.
- [35] J. Liang, D. DeMenthon and D. Doermann, "Geometric Rectification of Camera-captured Document Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 7, 2006.
- [36] T. Wada, H. Ukida and T. Matsuyama, "Shape from Shading with Interreflections under Proximal Light Source: 3D Shape Reconstruction of Unfolded Book Surface from a ScannerImage," in Computer Vision (ICCV '95), 1995.
- [37] M. S. Brown and W. B. Seales, "Document Restoration using 3D Shape: a General Deskewing Algorithm for Arbitrarily Warped Documents," Computer Vision, vol. 2, pp. 367-374, 2001.
- [38] T. Wada, H. Ukida and T. Matsuyama, "Shape from Shading with Interreflections Under a Proximal Light Source Distortion-Free Copying of an unfolded Book," Int'l Journal of Computer Vision, vol. 2, pp. 125-135, 1997.
- [39] O. Lavaille, X. Molines and X. Molines, "Active contours network to straighten distorted text lines," Image Processing, p. 1074–1077, 2001.
- [40] A. Masalovitch and L. Mestetskiy, "Usage of continuous skeletal image representation for document images dewarping," Pattern Recognition and Information Processing, pp. 279-284, 2007.

- [41] A. Ulges, C. H. Lampert and T. M. Breuel, "Document Image Dewarping using Robust Estimation of Curled Text Lines," in Proceedings of the eight international conference on document analysis and recognition, USA, 2005.
- [42] H. Cao, X. Ding and C. Liu, "Rectifying the bound document image captured by the camera: A model based approach," in 7th Int. Conf. Document Anal. Recognition, USA, 2003.
- [43] B. Gatos, N. Pratikakis and K. Ntirogiannis, "Segmentation Based Recovery of Arbitrarily Warped Document Images," in Ninth International Conference on Document Analysis and Recognition (ICDAR), 2007.
- [۴۴] ه. ده بوید، ف. رزازی و ش. علیرضایی، "ارائه روشی نوین برای کاهش اعوجاج تصویربرداری در تصاویر متنی فارسی تصویربرداری شده توسط دوربین"، در ششمین کنفرانس ماشین بینایی و پردازش تصاویر ایران، اصفهان، 1389.
- [۴۵] م. گل پرداز، ح. نظام آبادی پور و س. صیدنژاد، "رفع اعوجاج هندسی در تصاویر متن فارسی گرفته شده با دوربین دیجیتال"، در شانزدهمین کنفرانس ملی سالانه انجمن کامپیوتر ایران، تهران، تهران، 1389.
- [۴۶] س. خسروی راد، "رفع اعوجاجات غیرخطی در تصاویر اسناد فارسی، پایان‌نامه ی ارشد"، دانشگاه صنعتی شاهرود، شاهرود، 1391.
- [۴۷] م. شامقلی، "رفع اعوجاج و بهبود کیفیت تصاویر اسکن شده از کتب فارسی، پایان‌نامه ی ارشد"، دانشگاه صنعتی شاهرود، شاهرود، 1392.
- [48] R. Gonzalez and R. Woods, Digital Image Processing, 3rd Edition. Published: Upper Saddle River, NJ : Pearson/Prentice Hall, 2008.
- [49] T. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," SIAM, Journal on Applied Mathematics, vol. 3, p. 1019–1043, 2002.
- [50] L. Likforman, A. Zahour and B. Taconet, "Text Line Segmentation of Historical Documents: a Survey," International Journal on Document Analysis and Recognition, Springer, 2006.
- [51] A. Zahour, B. Taconet, P. Mercy and S. Ramdane, "Ramdane," Proceedings of the 6th ICDAR, pp. 281-285, 2001.
- [52] A. Antonacopoulos and D. Karatzas, "Document Image analysis for World War II personal records," First Int. Workshop on Document Image Analysis for Libraries, vol. 4, pp. 336-341, 2004.
- [53] S. Sangsawad, R. Chamchong and F. Fung, "Using Local Maxima Profile and Piece-Wise Technique for Line Segmentation on Thai Handwritten Historical Documents," in International Conference on Machine Learning and Cybernetics, Guilin, 2011.
- [54] D. C. Schneider, M. Block and R. Rojas, "Robust document warping with interpolated vector fields," in 9th Int. Conf. Document Anal. Recognit, Curitiba, Brazil, 2007.

- [55] J. Mundy and A. Zisserman, "Geometric Invariance in Computer Vision," MIT, 1992.
- [56] A. Criminisi, I. Reid and A. Zisserman, "A Plane Measuring Device," University of Oxford, 1993.
- [57] F. Shafait and M. Breuel. "Document Image Dewarping Contest", 2nd Int. Workshop on Camera-Based Document Analysis and Recognition, CBDAR'07, Brazil. Sep. 2007.
- [58] J. Carr, R. Beatson and B. McCallum, "Smooth surface reconstruction from noisy data," in NZ Ltd, Graphite,2003.
- [59] H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics from images," Computer Vision Systems, vol. 3, 1978.
- [60] K. Wong, R. Casey and F. Wahl, "Document analysis systems," IBM Journal of research and development, vol. 26, 1982.
- [61] Z. Shi and V. Govindaraju, "Line Separation for Complex Document Images Using Fuzzy Runlength," Proc. of the Int. Workshop on Document Image Analysis for Libraries, pp. 23-27, 2004.
- [62] L. Likforman and C. Faure, "Extracting lines on handwritten documents by perceptual grouping," in Advances in Handwriting and drawing: a multidisciplinary approach, Europa, Paris, 1994.
- [63] L. Likforman, A. Hanimyan and C. Faure, "A Hough Based Algorithm for Extracting Text Lines in Handwritten Document," Proc. of ICDAR, pp. 774-777, 1995.
- [64] A.T.Vandermonde, "https://en.wikipedia.org/wiki/Vandermonde_matrix. [Online]," [Online].
- [65] OmniPage 18 "<http://www.nuance.com>," [Online].
- [66] H. Khosravi, and E. Kabir, "A blackboard approach towards integrated Farsi OCR system", Document Analysis and Recognition, International Journal of Document Analysis and Recognition (IJ DAR), Vol. 12, pp. 21 -32, 2009.
- [67] J. Kanai, T. A. Nartker, S. Rice, and G. Nagy. "Performance metrics for document understanding systems" 2nd Int. Conf. Document, pp. 424–427, Japan, 1993.

Abstract:

Document images produced by scanner or digital camera, usually suffer from geometric distortions and photometric distortions. Both of them deteriorate the performance of OCR systems.

In this thesis harmonic inpainting is used to remove photometric distortion. To improve accuracy of this algorithm, we propose that diagonal neighbors can be used too. This will lead to an improvement in background intensity estimation. Although, to reduce computational complexity, we propose that in implementation of inpainting algorithm, first we must use main quad neighbors and then the diagonal neighbors. Using this idea, computational complexity of proposed method won't be increased. Also, Gaussian smoothing and Moving Average filters are used to improve quality in background estimation.

Although, in order to correct Geometric distortion, after extracting lines from image, each line is divided into some vertical strips. In each strip, coordinates of the point which its horizontal projection is maximum, is stored. Curvature of each line is determined by third order function using this set of points. At the end, geometric distortion of each line is corrected using Perspective transform estimation. Also, to determine curvature of short lines, curvature of previous or next long lines are used.

The proposed method was implemented on Persian and English databases and has been compared with other methods. The obtained results indicate that proposed method is powerful and precise in eliminating both types of distortions.

Keywords: Inpainting, Photometric distortion, Geometric distortion, Two-dimensional processing documents, Optical character recognition (OCR).



University of Shahrood

Faculty of Electronic and Robotics Engineering

**Removing Photometric and Geometric Distortion from text
Images using Background Intensity and Geometrical Text Lines
Cues**

Mohhammad Amin Tolou Beydokhti

Supervisor:

Dr. Alireza Ahmadifard

Advisor:

Dr. Hossein Khosravi

Winter 2015