

رفع اعوجاج هندسی متون به کمک

اطلاعات هندسی خطوط متن

محمد امین طلوع بیدختی و علیرضا احمدی فرد*

دانشکده مهندسی برق، گروه الکترونیک، دانشگاه صنعتی شاهرود، شاهرود، ایران



چکیده

تصاویر سند تهیه شده توسط پویش گر یا دوربین دیجیتال، همواره با اعوجاج‌های فتومتری و هندسی همراه هستند. وجود هر دو نوع اعوجاج، باعث کاهش دقت عملکرد نرم افزارهای شناسایی نویسه‌ها (OCR) می‌شوند. در این مقاله روشی نوین جهت رفع اعوجاج‌های هندسی از تصاویر متنی ارائه شده است. در روش پیشنهادی به منظور تصحیح اعوجاج هندسی، در ابتدا خطوط متن از تصویر استخراج و سپس هر خط متن به ستون‌هایی به عرض مساوی شکسته می‌شوند. برای هر قطعه استخراج شده از یک خط، راستای قطعه به نحوی تصحیح می‌شود که حروف موجود در آن قطعه در راستای افقی قرار گیرد. برای این منظور به ازای چرخش‌های مختلف قطعه متن، افکنش افقی تصویر محاسبه می‌شود و چرخشی از قطعه که بلندترین قله افکنش را ایجاد کند، راستای تصحیح شده آن قطعه در نظر گرفته می‌شود. بر این اساس یک نقطه مرجع که معرف راستای مبنا است، برای هر قطعه خط هم راستا شده با افق استخراج می‌شود. به کمک نقاط مرجع، هر قطعه از خط، انحنای آن خط متن به کمک برازش یک تابع درجه سه به دست می‌آید. در نهایت با استفاده از تخمین تبدیل پرسپکتیو، اعوجاج هندسی هر خط برطرف می‌شود. جهت افزایش پایداری روش پیشنهادی در تخمین انحنای خطوط متن با طول کم، از انحنای خطوط با طول بزرگ‌تر مجاور آن خط استفاده شده است. روش پیشنهادی بر روی پایگاه‌های داده فارسی و انگلیسی پیاده‌سازی و با برخی روش‌های هم تراز آن مقایسه شده است. نتایج بیان گر قدرت و دقت روش پیشنهادی در رفع اعوجاج هندسی است.

واژگان کلیدی: اعوجاج هندسی، پردازش دوبعدی اسناد، تخمین تبدیل پرسپکتیو، نویسه خوان نوری

Document Image Dewarping using geometrical information extracted from document lines

Mohammad Amin Tolou Beydokhti & Alireza Ahmadifard*

Department of Electronic, Faculty of Electrical, Shahrood University of Technology, Shahrood, Iran

Abstract

Document images produced by scanners or digital cameras usually have photometric and geometric distortions. If either of these effects distorts document, recognition of words from such a document image using OCR is subject to errors. In this paper we propose a novel approach to significantly remove geometric distortion from document images. In this method first we extract document lines from document using morphological operators. Then, extracted document lines are divided into a number of equal size column strips.

This allows to assume that each segment of line document is not curved. Each extracted document line segment is aligned horizontally. For this purpose, a segment line of document is rotated at different angles and for each rotation horizontal projection is obtained. The rotation angle with maximum peak at the corresponding projection signal is selected to align the line segment, horizontally. In order to estimate the

* Corresponding author

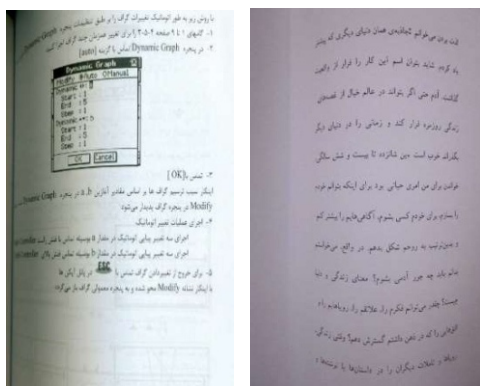
* نویسنده عهده دار مکاتبات

geometrical distortion, for each document line a reference point is extracted from each line segment. These points indicate the position of a document line at starting column of line segments. Using reference points of a document line a polynomial function is fitted to each document line. At the end, geometric distortion for each part of the document is eliminated using a perspective transformation.

This transformation is estimated based on the extracted polynomial function. To increase the stability of the proposed method for short text lines, the curve of adjacent text lines of longer length is used. A post processing stage is required after applying perspective transformation on document patches. Since this transformation is a continuous mapping but it is applied on digital images. To remove this distortion from the result, the consistency of each pixel value with the value of neighboring pixels are considered to correct the value of inconsistency pixels.

The proposed method is implemented on Persian and English databases and has been compared with the existing methods. The results indicate the efficiency and accuracy of the proposed method in elimination of geometric distortions.

Keywords: Geometric distortion, document processing, perspective Transformation, Optical character recognition (OCR)



(شکل ۱-۱): دو نمونه از تصویر اسناد دارای اعوجاج فتومتری و

هندسی

(Figure-1): Two document image samples with photometric and geometric distortion

به خمیدگی ایجاد شده در راستای خطوط سند یا کجی سند که هنگام پویش کردن یا عکس برداری در تصویر ایجاد می شود، «اعوجاج هندسی» می گویند. علت اصلی این اعوجاج، انحنای برگه های سند ناشی از قطر کتاب یا کج قرار گرفتن سند در زمان پویش یا عکس برداری است. این دو نوع اعوجاج علت اصلی تشخیص اشتباه کلمات در نرم افزارهای رایج OCR هستند.

هدف این مقاله رفع اعوجاج هندسی از متون تصویر برداری شده است؛ بنابراین فرض بر این است که تصاویر ورودی بدون اعوجاج فتومتری بوده و یا این اعوجاج به کمک روش های رایج مثل روش نیلک، روش ساولا یا روش حذف سایه به کمک درون نگاری بهبود یافته از تصاویر حذف شده است. همچنین به منظور بهبود کیفیت تصویر

۱- مقدمه

تصاویر متنی که به عنوان داده اولیه برای سامانه OCR مورد استفاده قرار می گیرند، توسط اسکنرها و دوربین های دیجیتال تهیه می شوند. علاوه بر آن، با پیشرفت فناوری دوربین تلفن های همراه، استفاده از آنها به منظور عکس برداری از متون و اسناد افزایش یافته است. با توجه به مشکلات عکس برداری با روش های ذکر شده، نیاز به بازسازی تصاویر بیش از پیش احساس می شود؛ چرا که بدون بهبود کیفیت آنها عملکرد دستگاه های OCR جهت بازشناسی متون بسیار ضعیف است.

در شکل (۱)، دو نمونه از اسناد الکترونیکی که نیاز به بازسازی دارند، نشان داده شده است. اگرچه این مشکلات در دستگاه بینایی انسان قابل چشم پوشی بوده و برای انسان بازشناسی متن امکان پذیر است، اما برای پردازش های رایانه ای از جمله تشخیص^۱، تحلیل^۲ و بازشناسی^۳ مشکل ایجاد می کنند.

به سبب ایجاد شده در تصویر متن که با توجه به نورپردازی غیر یکنواخت در تصویر ایجاد می شود، «اعوجاج فتومتری» گفته می شود. علت رخ دادن اعوجاج فتومتری قطور بودن کتاب در هنگام پویش کردن یا توزیع غیر یکنواخت منابع نوری در صحنه حاوی متن بوده که سبب ایجاد سایه در تصویر شده است.

¹ Understanding

² Analysis

³ Recognition

ماسالوویچ و همکارانش روشی به نام ²(SKEL) را معرفی کرده‌اند که در آن با استفاده از³پیکربندی تصویر یا اسکلت تصویر به رفع اعوجاج هندسی پرداخته‌اند. در این روش ابتدا پیکربندی تصویر ساخته می‌شود؛ سپس شاخه‌ها بر اساس طول و زاویه آن‌ها به دو دسته افقی و عمودی تقسیم‌بندی شده و شاخه‌های افقی، فیلتر می‌شوند. با این کار تنها شاخه‌هایی باقی می‌مانند که بین دو خط نوشتاری قرار گرفته‌اند. درنهایت با تخمین منحنی بی‌زری⁴ تصویر سند مدل‌سازی می‌شود. از معایب این روش این است که اگر در به‌دست آوردن اسکلت تصویر، به‌خصوص تخمین مرز عمودی تصویر، خطایی رخ دهد، کیفیت تصویر خروجی تا حد بسیار زیادی کاهش پیدا می‌کند [14].

در روش گتوس که به روش ⁵(SEG) نیز معروف است؛ خمیدگی هر کلمه به‌تنهایی توسط شیب آن کلمه برطرف می‌شود؛ سپس با توجه به چپ‌ترین کلمه موجود در آن خط، هم‌ترازی عمودی صورت می‌گیرد. با درنظرگرفتن این نکته که در این روش تصحیح اعوجاج هر کلمه به‌طور مستقل انجام می‌شود، درنتیجه بعد از نگاشت هر کلمه، خطوط دچار بی‌نظمی خواهد شد، به‌گونه‌ای که کلمات به‌خوبی در یک راستا در خط قرار نمی‌گیرند؛ همچنین به‌علت اینکه سنجش زاویه هر کلمه با توجه به طول متفاوت آن‌ها به‌مخصوص برای کلمات با طول کوتاه با خطا مواجه می‌شود؛ لذا این امر به کاهش دقت این روش می‌انجامد [8].

ده بوید و همکارانش الگوریتمی جدید به‌منظور کاهش چرخش سه‌بعدی در تصاویر متنی فارسی تصویربرداری‌شده توسط دوربین، معرفی کرده‌اند. 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. از معایب این روش می‌توان به تخریب شکل نهایی کلمات در خط اشاره کرد [4].

خسروی راد الگوریتمی جدید جهت رفع اعوجاج چرخش سه‌بعدی در تصاویر اسناد فارسی پیشنهاد داده است. در این روش اعوجاج هندسی با استفاده از رفع چرخش کلمات بر اساس شیب منحنی خط گذرنده از آن کلمه برطرف می‌شود. از معایب این روش می‌توان به تخریب

ورودی نیز می‌توان از روش بهسازی تصویر با کاهش اثر عوامل خرابی بر مؤلفه روشنایی و بازتابش تصویر استفاده کرد [18]، [16]، [6] و [2].

در سال‌های اخیر روش‌های متعددی در زمینه رفع اعوجاج هندسی پیشنهاد شده‌اند. این روش‌ها به دو دسته کلی تقسیم می‌شوند. دسته نخست تکنیک‌های مبتنی بر بازسازی به کمک مدل سه‌بعدی اعوجاج هندسی تصاویر متن و دسته دوم تکنیک‌هایی هستند که اعوجاج هندسی از روی مشخصه‌هایی از متن تخمین‌زده و سپس اقدام به رفع آن می‌شود.

در برخی از روش‌های دسته نخست دانستن اطلاعات اولیه در مورد نوع پویش‌گر و همچنین سامانه نورپردازی موردنیاز است؛ که البته این اطلاعات همیشه در دسترس نیست. در برخی دیگر از این روش‌ها با مدل‌کردن سطح صفحه متن به سطحی استوانه‌ای شکل به رفع این اعوجاج پرداخته می‌شود که البته این روش‌ها، با توجه به این نکته که نوع پیچش صفحه همیشه استوانه‌ای شکل نیست، جامع نیستند.

اولگس و همکارانش به رفع اعوجاج هندسی از تصاویر گرفته‌شده توسط دوربین‌های دیجیتال پرداخته‌اند. در این روش با توجه به مدل‌سازی سه‌بعدی ثابت از صفحه، به رفع اعوجاج هندسی پرداخته می‌شود [22].

لیانگ و همکارانش روشی برای یک‌سوسازی اعوجاج‌های ناخواسته معرفی کرده‌اند که در آن با به‌وجودآوردن یک مدل استوانه‌ای سعی در تخمین شکل سه‌بعدی صفحه متن شده است. یک محدودیت اساسی این روش فرض مدل هندسه‌ای استوانه برای صفحه متن است [11].

وادا و همکارانش روش دیگری را معرفی کردند که در آن از روش شکل به‌کمک سایه¹ استفاده می‌شود تا شکل سه‌بعدی سطح متن را بازسازی کند. از معایب این روش نیاز به مشخص‌بودن اطلاعاتی در مورد الگوی پخش نور محیط است که این اطلاعات، به‌طورمعمول در دسترس نیست [23].

دسته دوم، روش‌های مبتنی بر شناسایی اعوجاج هندسی از روی متن هستند که در این روش‌ها از روی انحناي خطوط، اعوجاج هندسی صفحه تخمین زده می‌شود و با یک نگاشت معکوس، متن دارای اعوجاج را می‌توان اصلاح کرد.

² Skeletonization

³ Image skeleton

⁴ Cubic Bezier curves

⁵ Segmentation

¹ Shape from Shading

شکل نهایی کلمات در اثر چرخش، تغییر در اندازه حروف و دقت پایین در رفع اعوجاج اشاره کرد [3].

شامقلی روشی برای رفع اعوجاج هندسی پیشنهاد کرده است. در این روش از تخمین خمیدگی خطوط در بالا و پایین صفحه و با انتقال سطح خمیده سند به سطح مستطیلی شکل اعوجاج هندسی تصویر برطرف می‌شود. البته این به علت انتقال تمام نقاط تصویر روشی زمان‌بر است. در جدول (۱) به صورت خلاصه روش‌های مروری ذکر شده بررسی شده‌اند [20]، [5].

از میان روش‌های بازشناسی کلمات فارسی نیز می‌توان به روش بایسته و همکارانش اشاره کرد که در این روش علاوه بر معرفی یک الگوریتم دومرحله‌ای برای بازشناسی کلمات، روشی جدید جهت استخراج ویژگی از هیستوگرام گرادین تصویر معرفی شده است [1].

ساختار مقاله در ادامه بدین صورت است: در بخش دوم روش پیشنهادی به منظور رفع اعوجاج هندسی معرفی می‌شود؛ پس از آن در بخش سوم نتایج روش پیشنهادی با روش‌های رایج مقایسه می‌شود و در بخش چهارم به جمع‌بندی و پیشنهاد برای ادامه کار پرداخته شده است.

(جدول-۱): توصیف مختصر روش‌های مرور شده

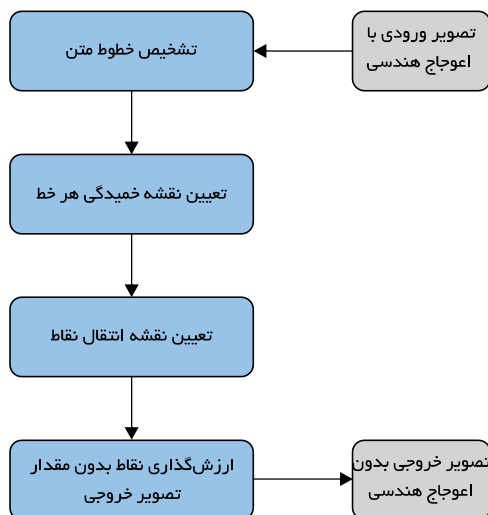
(Table-1): A brief description for the reviewed methods

روش	مشکل اصلی
اولگس [22]	فرض مدل سه بعدی ثابت از صفحه متن
لیانگ [11]	فرض مدل هندسه‌ای استوانه صفحه متن
وادا [23]	نیاز به مشخص بودن اطلاعاتی در مورد الگوی پخش نور محیط
ماسالوویچ [14]	حساسیت در پیکره‌بندی تصویر سند
گتوس [8] روش با استفاده	تخریب در قرارگیری شکل حروف در خط
ده بوید [4]	تخریب شکل نهایی کلمات در خط
خسروی [3] راد	تخریب شکل نهایی کلمات در خط
شامقلی [20]	زمان بالای اجرا

دوربین دیجیتال یا پویش گر و طریقه قرارگرفتن سند نسبت به سامانه تصویربرداری، نیازی نیست. روش پیشنهادی مبتنی بر استخراج خطوط است؛ به گونه‌ای که انحنای هر خط از بالا به پایین مشخص شده و پس از رفع اعوجاج هندسی هر خط، کار برای خطوط بعدی دنبال می‌شود.

در شکل (۲) روندنمای کلی روش پیشنهادی به منظور رفع اعوجاج هندسی نشان داده شده است. طبق این روندنما در مرحله نخست خطوط متن استخراج شده، سپس نقشه خمیدگی هر خط از متن مشخص می‌شود؛ در مرحله بعد، روشی جهت انتقال نقاط انجام شده و در مرحله آخر هر نقطه از تصویر سند که مقداردهی نشده، بر اساس نقاط همسایه آن تخمین زده می‌شوند.

در این مقاله از دو پایگاه داده خسروی راد برای اسناد فارسی و پایگاه داده DFKI برای اسناد لاتین استفاده شده است [19]، [3].



(شکل-۲): روندنمای روش پیشنهادی رفع اعوجاج هندسی
(Figure-2): Flowchart of the proposed method for removing geometric distortion

۲-۱- تشخیص خطوط متن

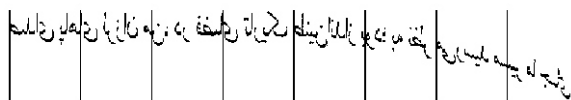
روش‌های زیادی به منظور تشخیص خطوط متن وجود دارد. یکی از روش‌هایی که هم از لحاظ سرعت و هم از لحاظ دقت، رفتار قابل قبولی دارد، روش مبتنی بر عمل گر ریخت‌شناسی گسترش است. در این روش تصویر ورودی می‌بایست بدون نوفه و به صورت دوسطحی باشد [24]، [13] و [12].

با توجه به این نکته که متون فارسی برخلاف متون انگلیسی از سمت راست شروع شده، از آینه افقی تصاویر متن فارسی، به عنوان داده ورودی استفاده می‌شود. با این کار تمام مراحل برای اسناد فارسی مشابه اسناد انگلیسی است. الگوریتم کلی به صورت زیر بیان می‌شود:

۲- روش پیشنهادی رفع اعوجاج هندسی

در روش پیشنهادی برای رفع اعوجاج هندسی، تنها به اطلاعات دوبعدی موجود در سند نیاز است؛ لذا به مشخصات

نوارها بستگی به میزان انحنای خطوط دارد و بر روی هر مجموعه داده متفاوت است. عرض مناسب بر روی پایگاه‌های داده مورد استفاده در این پژوهش، به صورت تجربی مقداری بین ۱۵۰ تا ۲۵۰ پیکسل به دست آمده است. با توجه به تعداد ستون‌های تصویر، عرض آخرین نوار ممکن است از بقیه نوارها کوچک‌تر باشد. در شکل (۴) یک نمونه از این تقسیم‌بندی نشان داده شده است.



(شکل-۴): شکستن یک خط از متن به نوارهای عمودی
(Figure-4): A document line is broken into vertical stripes

۲-۲-۲- افکنش افقی گرادیان تصویر قطعات شکسته شده و یافتن مختصات بیشترین

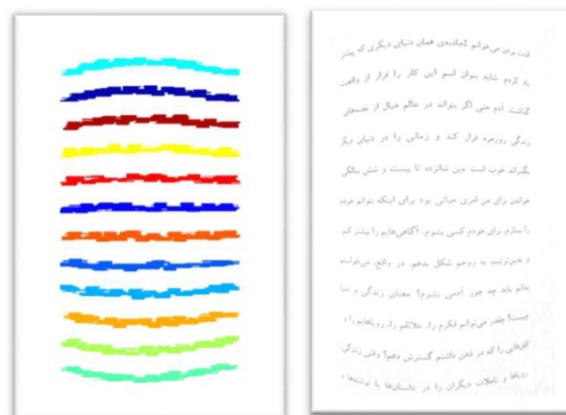
مقدار روی منحنی افکنش

به منظور پیدا کردن تابع انحنای هر خط از متن، خط مبنا برای هر یک از نوارهای عمودی مربوطه تخمین زده می‌شود. برای این منظور بیشینه مقدار تابع افکنش افقی گرادیان متن محاسبه شده است. برای حصول نتیجه بهتر ابتدا هر نوار در بازه $-\alpha$ تا $+\alpha$ چرخانده شده و زاویه‌ای که طبق آن حروف نوار عمودی، در راستای افقی قرار می‌گیرند، در نظر گرفته می‌شود.

در شکل (۵) چند نمونه از چرخش‌ها نشان داده شده است. با توجه به این نکته که نگارش یک خط از متن فارسی بر روی خط مبنا انجام می‌شود، می‌توان نتیجه گرفت که در اثر چرخش متن نوار عمودی 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 نوشته شده داخل نوار در ترازترین حالت ممکن خود است. مختصات نقطه بیشینه در این حالت، نشان‌گر خط مبنا برای نوار عمودی مربوطه در تصویر است؛ اما در اسناد انگلیسی با توجه به این نکته که حروف اغلب بین دو خط نوشته می‌شوند، بر اثر چرخش، زمانی نوار در ترازترین حالت قرار می‌گیرد که مختصات بیشینه نمودار افکنش افقی بر روی یکی از دو خطوط مبنا قرار گیرد.

محدوده مورد انتظار چرخش^۱

- قرینه تصویر ورودی محاسبه می‌شود.
- عمل گر ریخت‌شناسی گسترش با یک جزء ساختاری مناسب به تصویر قرینه شده اعمال می‌شود.
- نقاط به هم چسبیده توسط روش اجزای به هم متصل تشخیص داده می‌شوند.
- با انجام دادن این مراحل، ماسک مورد نیاز به منظور تشخیص خطوط تولید می‌شود. در شکل (۳) یک نمونه ماسک تولید شده توسط این روش نشان داده شده است.



(B)

(الف)

(شکل-۳): ماسک تولیدی جهت استخراج خطوط متن

(الف) تصویر ورودی (ب) ماسک تولید شده
(Figure-3): The produced mask to extract document lines A) input image B) the produced mask

۲-۲- تعیین نقشه خمیدگی هر خط

یکی از اصلی‌ترین مرحله‌های رفع اعوجاج هندسی، مرحله تشخیص نقشه خمیدگی سند است. در این مرحله هدف، مشخص کردن انحنای هر خط است. این مرحله خود از چند گام اساسی تشکیل شده است:

- شکستن هر خط به نوارهای عمودی
 - افکنش افقی گرادیان تصویر قطعات شکسته شده و یافتن مختصات بیشترین مقدار روی منحنی افکنش
 - برازش یک منحنی درجه سه بر نقاط مشخص شده از هر خط
- در ادامه این مراحل شرح شده‌اند.

۲-۲-۱- شکستن هر خط به نوارهای عمودی

در این گام هر خط استخراج شده به نوارهایی با فاصله مساوی تقسیم می‌شود. این عمل به منظور شکستن خط دارای انحنا به قطعاتی است که به دلیل کوتاه بودن، از انحنای آن می‌توان صرف‌نظر کرد. این نوارها با شروع از سمت چپ تصویر آغاز شده و تا انتها ادامه پیدا می‌کند. اندازه عرض

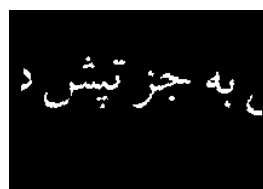
گرادیان قطعه انجام شود، بازهم خط مبنا در نمودار افکنش افقی ایجاد قله می‌نماید پس از به‌دست آوردن بهترین زاویه، مقدار گرادیان عمودی تصویر چرخانده‌شده را محاسبه کرده و این بار عملیات افکنش افقی بر روی گرادیان تصویر انجام می‌شود. در نهایت مختصاتی که نمودار افکنش گرادیان در آن بیشینه مقدار را دارد، ذخیره می‌شود.

در شکل (۷) نتیجه این عملیات مشاهده می‌شود. بیشینه مقدار نمودار افکنش افقی در سطر ۶۶م اتفاق افتاده است؛ بنابراین این نقطه، به‌عنوان سطر مربوط به خط مبنا در نظر گرفته می‌شود؛ اما نکته مهم در این بین، این است که بر اثر چرخش، اندازه تصویر نوارهای عمودی هم در راستای ستون‌ها و هم در راستای سطرها افزایش می‌یابد. در نتیجه نقطه به‌دست‌آمده نسبت به مقدار آفست است؛ لذا مختصات نقاط مرجع به‌دست‌آمده در نوارهای عمودی ترازشده می‌بایست به مختصات واقعی آن در تصویر ورودی نگاشته شود؛ به منظور حذف آفست مراحل زیر پیشنهاد می‌شود:

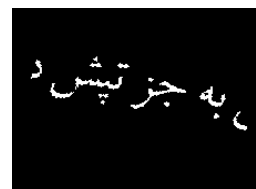
- ۱- جهت چرخش (علامت زاویه چرخش) تعیین شود.
 - ۲- اندازه تعداد سطرها و تعداد ستون‌های نوار اصلی و نوار چرخانده‌شده تعیین شود.
 - ۳- بر اساس زاویه چرخش نقطه یادشده محاسبه شود.
- با توجه به شکل (۸)، زاویه چرخش را با θ و تعداد سطرها و تعداد ستون‌های نوار اصلی را به‌ترتیب با m و n و تعداد سطرها و تعداد ستون‌های نوار چرخانده‌شده را به‌ترتیب با M و N نشان می‌دهیم: ($N \geq n$ و $M \geq m$). هدف پیدا کردن تصویر نقطه مرجع A (که از افکنش افقی در مرحله قبل به‌دست آمد) بر روی نوار چرخانده‌شده (A') و یافتن

$$|A'B'| = |BA'| \text{ فاصله}$$

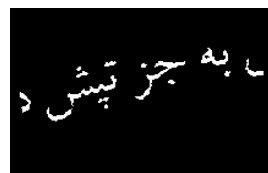
است.



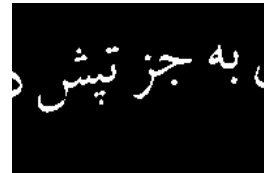
ب (B)



الف (A)



ت (D)



پ (C)

(شکل-۵): چرخش بخشی از خط با زاویه‌های مختلف

الف) چرخش با زاویه ۱۶- درجه ب) چرخش با زاویه ۸-

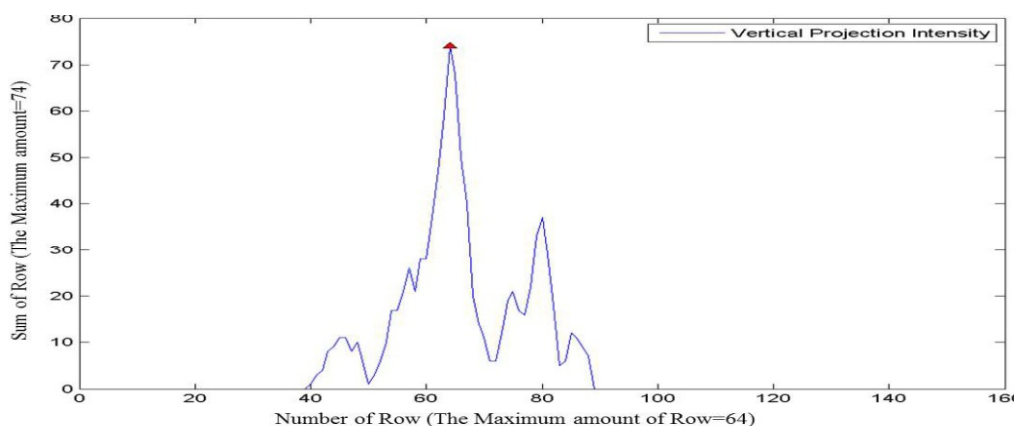
درجه پ) چرخش با زاویه ۸ درجه ت) چرخش با زاویه

۱۶ درجه

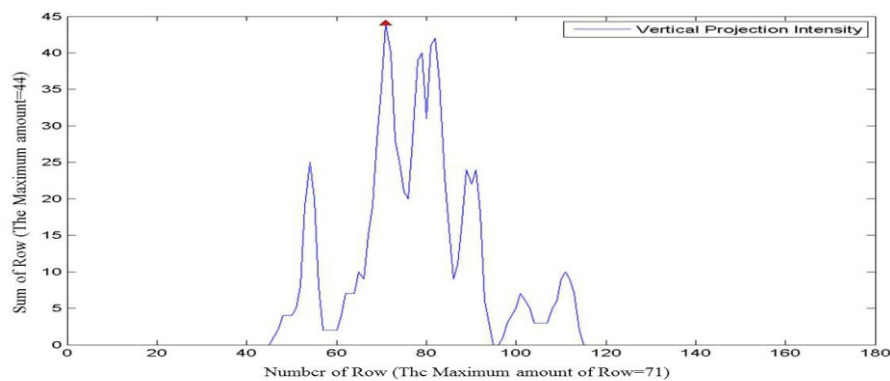
(Figure-5): Rotation of a line segment at different angles. A) -16 degrees B) -8 degrees C) +8 degrees D) +16 degrees

با مشاهده نمودارهای شکل (۶) متوجه می‌شویم که با چرخش نوار به‌اندازه زاویه ۸- درجه، متن در نوار عمودی تراز می‌شود. این زاویه از بیشینه مقدار نمودارهای افکنش به‌دست می‌آید؛ زیرا این مقدار برای زاویه ۸- درجه برابر با ۷۴ بوده درحالی‌که برای سایر زاویه‌ها از مقدار ۷۴ کوچک‌تر است؛ اما مشکل این روش زمانی است که حروف پایین‌رونده‌ای مثل «ر»، «ز» یا «و» پشت سر هم قرار گرفته و یا تعداد حروف کمی در یک قطعه از نوار عمودی قرار گیرند. در این صورت قله‌ی نمودار افکنش افقی تصویر قطعه، خط مبنا را به‌درستی تعیین نمی‌کند.

اگر عملیات افکنش افقی به‌جای تصویر قطعه بر روی

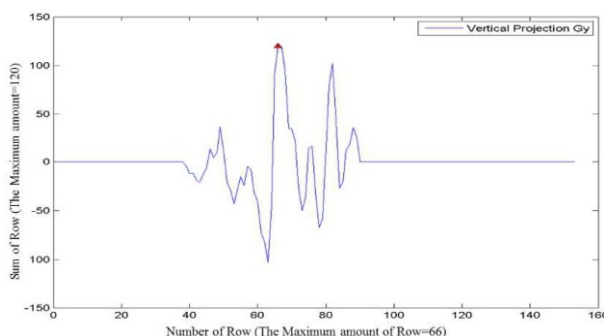


الف (A)

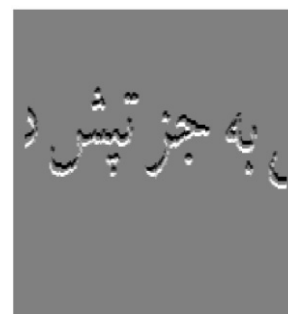


ب(B)

(شکل-۶): نمودار افکنش افقی شدت روشنایی تصویر الف) زاویه چرخش ۸- درجه ب) زاویه چرخش ۱۶- درجه
(Figure-6): Vertical projection of a line segment for two different rotations A) -8 degrees B) -16 degrees.

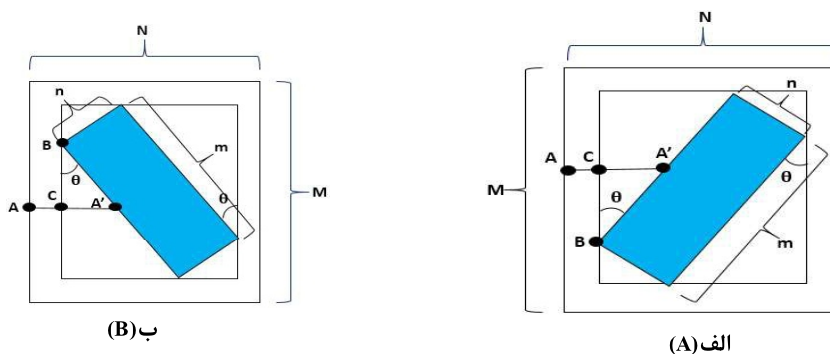


ب(B)



الف(A)

(شکل-۷): نمودار افکنش افقی گرادیان الف) گرادیان تصویر ب) نمودار افکنش افقی گرادیان
(Figure-7): Vertical projection gradient graph A) image gradient B) vertical projection gradient



ب(B)

الف(A)

(شکل-۸): تغییر اندازه تصویر بر اثر چرخش الف) چرخش به سمت راست ($\theta > 0$) ب) چرخش به سمت چپ ($\theta < 0$)

(Figure-8): Image rotation A) rotation to the right $\theta > 0$ B) rotation to the left $\theta < 0$

C با این فاصله از B ایجاد شده است. در این صورت مثلث $A'CB$ مثلثی قائم‌الزاویه بوده و اندازه وتر آن طبق رابطه (۱) به دست می‌آید.

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

در رابطه (۱) مقدار θ برابر با زاویه چرخش است که در مرحله قبلی به دست آمده بود و مقدار $|BC|$ نیز طبق

برای این منظور یک ماتریس با ابعاد نوار عمودی اصلی با مقادیر یک ایجاد کرده و به اندازه θ چرخانده می‌شود. بر اثر چرخش تعدادی سطر و ستون با ارزش پیکسل‌های پس‌زمینه (مشکی) به تصویر اضافه خواهد شد.

از سمت چپ تصویر شروع کرده و به ازای ستون‌های مختلف نخستین سطری که دارای پیکسل سفیدرنگ (روشنایی قلم) است، نقطه مرجع (B) نام‌گذاری می‌شود. فاصله عمودی نقطه A تا نقطه‌ی B را محاسبه کرده و نقطه

روشی که در بند قبلی توضیح داده شد، مشخص است. بنابراین مقدار $|AB|$ به سادگی محاسبه می‌شود.

با توجه به این نکته که نقطه B برای چرخش به سمت راست، در پایین صفحه و برای چرخش به سمت چپ، در بالای صفحه قرار دارد، بنابراین مختصات نقطه جدید برای چرخش به سمت راست $(X(A), m-|A'B|)$ و برای چرخش به سمت چپ برابر با $(X(A), |A'B|)$ است. برای هر نوار، مراحل ذکر شده انجام می‌شود تا نقاط مرجع جهت تخمین انحنای هر خط به درستی به دست آید.

نتایج تجربی اولیه نشان داد که به منظور تخمین دقیق‌تر انحنای خطوط لازم است که محل خط مبنا در ستون انتهایی هر سطر متن در تخمین به کار گرفته شود. همچنین با توجه به تقسیم‌بندی هر خط به نوارهای عمودی که در شکل (۹) الف نیز نشان داده شده است، آخرین نقطه محاسبه شده تا انتهای خط فاصله قابل توجهی دارد که این مسئله باعث تشخیص اشتباه انحنا در انتهای هر خط می‌تواند شود. به همین منظور پیشنهاد می‌شود یک نوار با طول برابر با سایر نوارها از سمت راست تصویر جدا کرده و تصویر آیینه‌شده افقی در آن قرار بگیرد و مراحل طبق روال گذشته انجام شود و نقطه به دست آمده از نوار انتهای آیینه‌شده در آخرین ستون به دست بیاید.

نتیجه این مرحله در شکل (۹) نشان داده شده است. در این تصویر، در آخرین ستون نیز یک نقطه جهت تخمین انحنا به دست آمده است. این امر به تخمین دقیق‌تر انحنا در نوار انتهایی هر خط کمک خواهد کرد.

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-۹): پیدا کردن نقاط جهت تشخیص انحنای هر خط الف) نتیجه روش پیشنهادی برای پیدا کردن نقاط جهت تشخیص انحنا ب) نتیجه روش پیشنهادی جهت یافتن نقطه پایانی

(Figure-9): Finding reference points to determine the curvature of each line A) the result of the proposed method for finding the reference points to estimate curvature B) the result of the proposed method for finding the ending points

همان‌طور که در شکل (۹-ب) مشخص است با به دست آوردن نقطه انتهایی هر خط به تشخیص صحیح انحنای خطوط در مراحل بعدی کمک شده است.

۳-۲-۲-۳- برازش یک منحنی درجه سه بر نقاط مشخص شده از هر خط

برای تخمین خمیدگی هر خط از تقریب چندجمله‌ای درجه سه استفاده می‌کنیم. چندجمله‌ای درجه سه با رابطه (۲) مشخص می‌شود. در این رابطه a_0 تا a_3 ضرایب چندجمله‌ای درجه سوم می‌باشند.

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

با توجه به مجموعه نقاط مرجع به دست آمده به منظور برازش منحنی درجه سوم با این نقاط از روش چند جمله‌ای وندرموند^۱ استفاده می‌شود. طبق این روش برازش تابع درجه سه به صورت رابطه (۳)، فرموله می‌شود [17].

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

در رابطه (۳) مقادیر (x_i, y_i) نقاط مرجع محاسبه شده در مرحله قبلی بوده و n نیز تعداد این نقاط است؛ بنابراین هدف محاسبه ضرایب چندجمله‌ای درجه سوم (a_0 تا a_3) برای هر خط است که طبق رابطه (۳) به راحتی قابل محاسبه است.

در شکل (۱۰) چند نمونه از منحنی‌های تخمین زده شده به وسیله این روش نشان داده شده است. همان‌طور که در این شکل‌ها مشاهده می‌شود، برای متون انگلیسی انحنا بر روی خط مبنا پایینی و برای متون فارسی این انحنا بر روی خط مبنا واقع می‌شود. همان‌طور که در قبل ذکر شد، از آینه افقی متون فارسی استفاده شده است.

اما نکته مهم در این قسمت این است که اگر طول یک سطر متن از حدی کمتر باشد، تعداد نقاط کمی جهت برازش منحنی درجه سه در دسترس است. این مسئله به تشخیص اشتباه انحنای خطوط کوتاه منجر می‌شود.

¹ Vandermonde polynomial

- نداشت قطعات چهارضلعی تشکیل‌دهنده خط متن به مستطیل

۱-۳-۲- قرار دادن هر خط متن بین دو منحنی

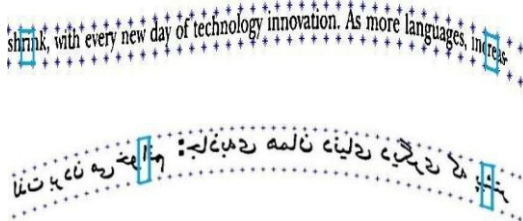
در این گام با توجه به انحنا به‌دست‌آمده برای هر خط متن، هر خط، بین دو منحنی بالایی و پایینی محصور می‌شود. به عبارت دیگر هدف این است که این دو منحنی حدود بالایی و پایینی خط متن را مشخص کنند؛ لذا برای این منظور کافی است طبق رابطه (۲)، دو منحنی درجه سه با ضرایب a_1 تا a_3 یکسان با ضرایب منحنی به‌دست‌آمده ایجاد کرده، ولی ضریب a_0 این دو منحنی با توجه به اینکه این دو منحنی فقط در راستای سطرها جابه‌جا شده‌اند، متفاوت است.

با مطالعه بر روی قلم‌های مختلف و بررسی اسناد گوناگون، فاصله بین دو منحنی بر اساس دو برابر بیشترین ارتفاع حروف در هر خط انتخاب شده و خط به‌گونه‌ای بین دو منحنی قرار می‌گیرد که به‌صورت کامل محصور شده است.

۲-۳-۲- شکستن ناحیه بین دو منحنی محصورکننده

خط متن به تعدادی ستون

در این گام ابتدا باید ناحیه بین دو منحنی بالا و پایین یک خط متن را توسط خطوط عمودی با فاصله یکسان به نوارهای عمودی تقسیم کنیم. ناحیه بین دو ستون متوالی را از نوار متن یک چهارضلعی می‌توان فرض کرد. با محاسبه نقاط تلاقی خطوط عمودی با دو منحنی درواقع مختصات چهار رأس هر چهارضلعی به‌دست خواهد آمد. در شکل (۱۱) دو نمونه را از این چهارضلعی‌ها مشاهده می‌کنید.



(شکل-۱۱): نقاط رئوس چهارضلعی‌ها (الف) متن انگلیسی (ب) متن فارسی

(Figure-11): The points of the quadrilateral vertices A) English text B) Persian text

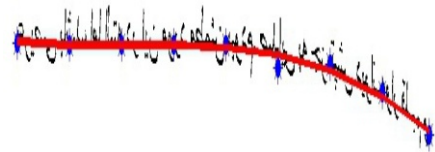
به‌منظور رفع این مشکل پیشنهاد می‌شود؛ در مواردی که طول خطوط از یک‌سوم میانگین طول خطوط تصویر کمتر است، به‌جای انحنای به‌دست‌آمده، از انحنای خطوط مجاور بالا یا پایین که طول بیشتری دارند، جهت تشخیص انحنای آن خط استفاده شود.



(الف) (A)



(ب) (B)



(ج) (C)

(شکل-۱۰): نمونه‌هایی از تخمین انحنای خطوط الف و ب) متون انگلیسی ج) متن فارسی (از آینه خط فارسی استفاده شده‌است).

(Figure-10): Examples for estimating the curvature of document lines A,B) English texts C) Persian text (Persian text has been mirrored)

۳-۲- تعیین نقشه انتقال نقاط

پس از مشخص‌شدن انحنای هر خط متن، هدف، تصحیح تصویر متن جهت رفع اعوجاج هندسی است. به همین منظور و با توجه به محل قرارگیری خطوط انحنا در گام نخست باید هر سطر متن را بین دو منحنی که در بالا و پایین سطر قرار گرفته و از انحنای سطر تبعیت می‌کنند، قرار داد؛ سپس سطر محصورشده بین منحنی‌ها به قطعات کوچکی شکسته می‌شود. در این صورت هر قطعه یک چهارضلعی بوده که با نداشت تخمین تبدیل پرسپکتیو^۱ به یک مستطیل تبدیل می‌شود. [7,15]

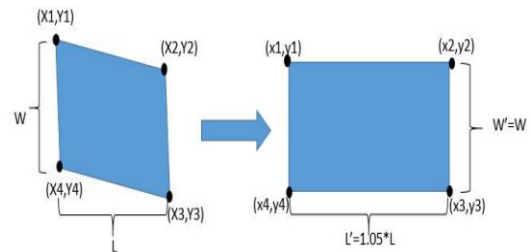
به‌طور خلاصه این گام‌ها به‌صورت زیر بیان می‌شوند:

- قراردادن هر خط متن بین دو منحنی
- شکستن ناحیه بین دو منحنی محصورکننده خط متن به تعدادی ستون

¹ Perspective Transform Estimation

خط متن به مستطیل

با به دست آمدن مختصات چهار رأس چهارضلعی‌ها، می‌توان با استفاده از روش تخمین تبدیل پرسپکتیو^۱ به انتقال نقاط از چهارضلعی‌ها به مستطیل‌هایی با مختصات چهار رأس مشخص شده پرداخت. با مشخص بودن مشخصات چهارضلعی مختصات چهار رأس مستطیل به گونه‌ای انتخاب می‌شود که عرض مستطیل با چهارضلعی برابر بوده و طول مستطیل ۱,۰۵ برابر طول چهارضلعی باشد؛ البته این ضریب به صورت تجربی به دست آمده است. در شکل (۱۲) این موارد نشان داده شده است. در نتیجه این نگاشت اعوجاج هندسی خطوط برطرف می‌شود؛ البته برخی نقاط تصویر خروجی در این نگاشت مقداره‌ی نشده‌اند؛ که در گام بعدی با درون‌یابی مقدار آن‌ها تعیین می‌شود.



(شکل-۱۲): انتقال نقاط از چهار ضلعی به مستطیل

(Figure-12): Transformation of the vertices from the quadrilateral to the rectangular

۳-۲-۴- ارزش‌گذاری نقاط بدون مقدار تصویر خروجی

با توجه به گسسته بودن تصاویر و برابری تعداد پیکسل‌های چهارضلعی و مستطیل، نگاشت پیکسل‌های نواحی چهارضلعی به مستطیل‌ها یک‌به‌یک نیست؛ لذا نقاطی در تصویر نهایی وجود دارد که مقداره‌ی نمی‌شوند. این نقاط با مقدار یک، مقداره‌ی اولیه می‌شوند؛ به همین دلیل تصاویری با نقاط سفید زائد مطابق شکل (۱۳) ایجاد می‌شود. به منظور تصحیح مشکل بالا گام‌های زیر بر روی تصویر نهایی انجام می‌شود:

- ۱- مشخص کردن ستون‌هایی در تصویر که تمام مقادیر آن یک بوده، در این ستون‌ها اگر پیکسلی وجود داشت که دو همسایه سمت چپ و راست آن دارای مقدار صفر بود، مقدار پیکسل مذکور برابر با صفر می‌شود.
- ۲- در هر سطر اگر پیکسلی با مقدار یک وجود داشت که دو همسایه بالایی و پایینی آن مقدار صفر داشتند، مقدار آن

¹ Perspective Transform Estimation

پیکسل نیز صفر می‌شود.

با اعمال گام‌های بالا تصویر خروجی تصحیح می‌شود. نتیجه اعمال این مرحله را در شکل (۱۳ب) نشان داده شده است.

communication

(الف) A

communication

(ب) B

(شکل-۱۳): پس پردازش بر روی نتیجه نگاشت یافته متن (الف)

تصویر متن نگاشت یافته (ب) نتیجه پس پردازش پیشنهادی بر

روی (الف)

(Figure-13): Post processing on transformed text image A) text image after transformation B) the result of proposed post processing on image (A)

۳- نتایج تجربی

به منظور مقایسه بهتر برای متون فارسی، نتایج روش پیشنهادی با نتایج دو روش شامقلی [5] خسروی‌راد [3] و برای متون لاتین نتایج روش پیشنهادی با نتایج سه روش شامقلی [20]، SKEL [14] و SEG [8] مقایسه می‌شود.

در روش خسروی‌راد هر کلمه بر اساس میزان چرخش آن به صورت جدا رفع اعوجاج می‌شود، البته تصویر خروجی این الگوریتم می‌تواند خاکستری یا دوسطحی باشد. از جمله مشکلات این روش دقت پایین در رفع اعوجاج هندسی و جابه‌جایی بیش از حد حروف و نقاط از کلمه است. مشکلات روش شامقلی نیز که در بخش قبلی هم مورد بررسی قرار گرفت، عبارت بودند از:

- 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 است، در مواقعی که در طول تصویر انحنا چندین بار تغییر کند، اعوجاج هندسی به خوبی برطرف نمی‌شود.
 - ۳- با توجه به اینکه در این روش تمام نقاط تصویر انتقال پیدا می‌کنند (حتی نواحی پس‌زمینه که حاوی اطلاعات نیستند)، این الگوریتم زمان‌بر خواهد بود.
- بنابراین مهم‌ترین مشکل روش شامقلی پیچیدگی محاسباتی در نتیجه زمان بالای اجرای این روش است. در تصاویر پایگاه داده‌ی خسروی‌راد به علت درجه تفکیک‌پذیری

نرم افزارهای استفاده شده به خوبی قادر به تشخیص فرمول ها و نمادهای ریاضی نیستند، این عبارات در محاسبه لحاظ نشده اند.

(جدول-۳): دقت شناسایی کلمات متون فارسی
(Table -3): The accuracy of word recognition on Persian texts

درصد دقت شناسایی کلمات	روش
67.81793	روش خسروی راد [3]
86.73025	روش شامقلی [5]
86.9009	روش پیشنهادی

برای متون لاتین نتایج روش پیشنهادی با نتایج سه روش شامقلی [20] و SKEL [14] و SEG [8] مقایسه می شود. در روش SEG، خمیدگی هر کلمه به تنهایی توسط شیب آن کلمه برطرف می شود. پس از آن با توجه به چپ ترین کلمه موجود در آن خط هم ترازای عمودی صورت می گیرد؛

با در نظر گرفتن این نکته که در روش بالا تصحیح اعوجاج هر کلمه به طور مستقل انجام می شود؛ در نتیجه بعد از نگاشت هر کلمه، خطوط دچار بی نظمی هایی خواهد شد؛ به گونه ای که کلمات به خوبی در یک راستا قرار نمی گیرند؛ همچنین به علت اینکه سنجش زاویه هر کلمه با توجه به طول متفاوت آن ها به خصوص برای کلمات با طول کوتاه با خطا مواجه می شود؛ لذا این امر به کاهش دقت این روش می انجامد.

ایده اصلی روش SKEL بر اساس استخراج اسکلت بیرونی تصاویر متنی است. این روش برخلاف روش SEG قادر به بازیابی اجزای گرافیکی است؛ اما این روش بسیار حساس به یافتن پیکربندی بوده و در صورت ایجاد خطا در محاسبه آن، خروجی اعوجاج خواهد داشت.

با توجه به نتایج که در جدول (۴) برای متون لاتین آورده شده است، مشخص است که روش پیشنهادی از دقت قابل قبولی برخوردار است. برای سنجش این دقت، تصاویر رفع اعوجاج شده توسط نرم افزار OmniPage به فایل متن تبدیل شده، سپس درصد دقت (R) طبق رابطه (۶) محاسبه شده است. از بین ۱۵۶۸۱ کلمه تعداد ۹۷ کلمه اشتباه تشخیص داده شده است [25].

همان طور که در جدول (۴) مشاهده می کنید، دقت روش پیشنهادی نسبت به سایر روش های مورد مقایسه بیشتر است؛ همچنین به منظور مقایسه به صورت کیفی دو نمونه از

بالای تصاویر این مشکل نمایان تر است. به منظور مقایسه سرعت اجرا، سرعت روش شامقلی با سرعت روش پیشنهادی در جدول (۲) آورده شده اند که البته این نتایج بر روی تصاویر پایگاه داده خسروی راد بوده که ابعاد تصاویر 3000×4000 پیکسل است.

(جدول-۲): مدت زمان اجرای روشهای رفع اعوجاج هندسی
(Table-2): Computation time for considered de-warping methods

روش	مدت زمان اجرا (بر حسب ثانیه)
روش شامقلی به کمک الگوریتم DLCM [5]	430
روش شامقلی به کمک عملگر گسترش [5]	270
روش پیشنهادی	140

در روش شامقلی جداسازی خطوط توسط دو روش $DLCM^1$ و استفاده از عملگر گسترش انجام می شود، البته مدت زمان اجرای روش شامقلی با الگوریتم DLCM همان طور که در جدول (۲) مشاهده می کنید، بسیار بیشتر از مدت زمان اجرای روش شامقلی با عملگر گسترش است. با این حال روش پیشنهادی سریع تر از روش شامقلی با هر دو روش جداسازی خطوط است.

به منظور محاسبه کمی، درصد صحت کلمات بر روی مجموعه تصاویر پایگاه داده خسروی راد محاسبه شده است [9]. [3].

از بین ۲۹۸۸۰ کلمه تعداد ۳۹۱۴ کلمه اشتباه تشخیص داده شده است. با توجه به نتایج که در جدول (۳) برای متون فارسی آورده شده، مشخص است که روش پیشنهادی از دقت قابل قبولی برخوردار است. برای سنجش این دقت، تصاویر رفع اعوجاج شده توسط نرم افزار پرشیا نگار به فایل متن تبدیل شده، سپس درصد دقت صحت (R) طبق رابطه (۶) محاسبه می شود [10].

$$R = 100 * \frac{\#Words - \#MisWords}{\#Words} \quad (6)$$

که در این رابطه $\#Words$ تعداد کل کلمات موجود در متن و $\#MisWords$ تعداد کل کلماتی بوده که به درستی توسط نرم افزار تشخیص داده نشده است. در محاسبه درصد دقت صحت ذکر این نکته ضروری است که با توجه به این که

¹ Dynamic Local Connectivity Map

(جدول-۴): دقت شناسایی کلمات در متون انگلیسی
(Table -4): The accuracy of word recognition on English texts

درصد دقت شناسایی کلمات	روش
94.9811	روش SEG [8]
98.6324	روش SKEL [14]
99.3622	روش شامقلی [20]
99.3814	روش پیشنهادی

سرعت بالاتری دارد، دقت را نیز افزایش داده است. البته روش پیشنهادی به تنهایی قادر به رفع اعوجاج اسناد دارای اجزای گرافیکی نبوده ولی با استفاده از روش‌های رایج استخراج اجزای گرافیکی از تصویر، می‌تواند اعوجاج این نوع اسناد را نیز برطرف کند.

۴- نتیجه‌گیری و پیشنهادها

در این مقاله به منظور تصحیح اعوجاج هندسی پس از استخراج خطوط از تصویر، هر خط به چندین نوار عمودی تقسیم می‌شود. در هر نوار با چرخش تصویر، مختصات نقطه‌ای که مقدار افکنش افقی آن بیشینه است، ذخیره می‌شود. انحنای هر خط به کمک این مجموعه نقاط توسط یک تابع درجه سه مشخص می‌شود. در نهایت با استفاده از تخمین تبدیل پرسپکتیو، اعوجاج هندسی هر خط برطرف می‌شود. همچنین به منظور تشخیص انحنای خطوط با طول کم از انحنای خطوط با طول بزرگ‌تر قبلی یا بعدی استفاده می‌شود. با مقایسه نتایج مشخص شد که دقت و سرعت روش پیشنهادی مناسب است.

با توجه به الگوریتم پیشنهادی، کارهای زیر برای ادامه و بهبود این کار پیشنهاد می‌شوند:

ارزیابی روش پیشنهادی بر روی خطوط دست‌نویس فارسی و انگلیسی و مقایسه نتایج با روش‌های رایج که سعی می‌شود دو نوع اعوجاج فتومتری و هندسی این اسناد برطرف شود.

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 دوستونه که برای این کار ابتدا باید خطوط به درستی تشخیص داده شوند و پس از آن اعوجاج هندسی هر خط برطرف شود.

افزایش وفقی طول هر نوار در مواقعی که تعداد حروف کمی در آن نوار قرار گرفته باشد. برای این کار زمانی که تعداد پیکسل‌های مشکی در یک نوار از یک مقدار آستانه‌ای کمتر باشد، طول نوار به صورت خودکار افزایش می‌یابد. این کار به تشخیص دقیق‌تر انحنای هر خط می‌تواند بیانجامد.

اسناد الکترونیکی در شکل (۱۴) نشان داده شده‌اند که نتایج روش پیشنهادی و سایر روش‌ها در تصاویر ۱۵ تا ۱۷ نمایش داده شده است.

در تصویر (۱۵ الف، ب و ج) به ترتیب نتایج روش‌های شامقلی، خسروی راد و روش پیشنهادی، بر روی تصویر (۱۴-الف) نشان داده شده است. همان‌طور که مشاهده می‌شود، نتایج روش پیشنهادی و شامقلی از روش خسروی راد بسیار مناسب‌تر بوده و این با نتایج جدول (۳) سازگاری دارد.

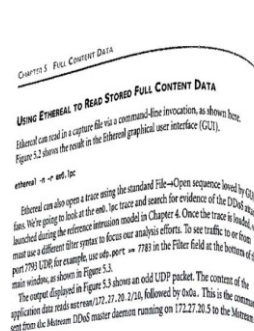
در شکل (۱۶) قسمتی از نتیجه تصویر (۱۵ ب و ج) توسط نرم‌افزار پرشیانگار نشان داده شده است. همان‌طور که مشاهده می‌کنید، نتیجه روش پیشنهادی در رفع اعوجاج هندسی مناسب است.

در شکل (۱۷) نتیجه روش پیشنهادی و سه روش شامقلی، SEG و SKEL برای تصویر (۱۴ ج) را مشاهده می‌کنید. با مشاهده این تصاویر به خصوص نواحی مشخص شده، معلوم می‌شود که دقت روش پیشنهادی در رفع اعوجاج هندسی مناسب است. نتایج روش پیشنهادی برای تصاویر (۱۴ ب و د) در شکل (۱۸) نشان داده شده است.

نکته دیگری که برتری روش پیشنهادی نسبت به روش شامقلی را نشان می‌دهد، این است که روش پیشنهادی نواحی مرزی حروف را به خوبی استخراج می‌کند و نوفه مرز قلم در کلمات نسبت به روش شامقلی کمتر است. این موضوع با مشاهده شکل (۱۹) بهتر درک می‌شود.

در تصویر ۲۰ نتیجه نرم‌افزار OmniPage برای قسمتی از متن تصاویر (۱۷ ج و د) نشان داده شده است. کلمات مشخص شده نمونه‌ای از کلماتی است که در آن روش پیشنهادی نسبت به روش شامقلی برتری داشته است.

در مقایسه با روش شامقلی که بالاترین دقت را در میان سایر روش‌های مورد مقایسه دارد، روش پیشنهادی علاوه بر آنکه



(D)د

(C)ج

(B)ب

(A)الف

(شکل-۱۴): نمونه هایی از اسناد دارای اعوجاج هندسی الف و ب) پایگاه داده‌ی خسروی راد [3] ج و د) پایگاه داده‌ی DFKI [19] (Figure-14): Examples for documents with geometric distortion A,B) Khosravi Rad's database[3] C,D) DFKI database [19]

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.

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.

(C)ج

(B)ب

(A)الف

(شکل-۱۵): نتایج رفع اعوجاج هندسی تصویر (۱۴ الف) الف) روش خسروی راد[3] ب) روش شامقلی[5] ج) روش پیشنهادی (Figure-15): Results for de-warping geometric distortion (Fig. 14-A) A) Khosravi Rad's method[3] B) Shamgholi's method[5] C) the proposed method



(B)ب



(A)الف

(شکل-۱۶): خروجی نرم‌افزار پرشیا نگار تصویر (۱۵ ب، ج) الف) روش شامقلی[5] ب) روش پیشنهادی

(Figure-16): Word recognition using PERSHIA NEGAR OCR software (Fig. 15-B,C) A) de-warping using Shamgholi's method [5] B) de-warping using the proposed method

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. 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 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. 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

(B)

(A) 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 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. 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

(D) د

(C) ج

(شکل-۱۷): نتایج رفع اعوجاج هندسی تصویر (۱۴ ج) الف) روش SEG [8] ب) روش SKEL [14] ج) روش شامقلی [20]

د) روش پیشنهادی

(Figure-17): The result of word recognition using PERSHIA NEGAR software (Fig.14-C) A) de-warpage using SEG method [8] B) de-warpage using SKEL method [14] C) de-warpage using Shamgholi's method [20] D) the proposed method

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 em0.1pc
```

Ethereal can also open a trace using the standard File→Open sequence loved by GUI

fans. We're going to look at the em0.1pc 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

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. این شانزده تا بیست و شش سالگی، خواندن برای من امری حیاتی بود برای اینکه بتوانم خودم را بازم، برای خودم کسی بشوم، آگاهی‌هایم را بیشتر کنم و بدین ترتیب به روحم 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 افق‌هایم را که در ذهن داشتم گسترش دهم؟ وقتی زندگی، رویاها و

(B)ب

(A)الف

(شکل-۱۸): نتایج روش پیشنهادی (الف) تصویر ۱۴ ب (ب) تصویر ۱۴ د
(Figure -18): The result of proposed method A) Fig.14-B B) Fig. 14-D

been written

(A)الف

been written

(B)ب

(شکل-۱۹): قسمتی از شکل‌های ۱۷ ج و د (الف) روش شامقلی [20] ب) روش پیشنهادی
(Figure -19): A part of image in Fig. 17-C, D Using A) Shamgholi's method [20] B) the proposed method

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. 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.

(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 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. 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.

(ب)

(شکل-۲۰): خروجی نرم‌افزار OmniPage بر روی شکل ۱۷ ج، د الف) نتیجه روش شامغلی [20] ب) نتیجه روش پیشنهادی (Figure -20): OmniPage software Output on Fig 17-C,D using A) Shamgholi's method [20] B) the proposed method

- [5] M. Shamgholi, "Distortion correction and Image enhancement in Persian Books" M.S. thesis, Shahrood ut, Shahrood, Iran, 2013.
- [6] م. ا. طلوع بیدختی و ع. احمدی فرد, "رفع اعوجاج فتومتریکی از تصویر اسناد به کمک درون نگاری بهبود یافته", ششمین کنفرانس فناوری اطلاعات و دانش. دانشگاه صنعتی شاهرود شاهرود, ایران, ۱۳۹۳.
- [6] M.A. Tolou Beydokhti and A. Ahmadyfard "Removing Photometric Distortion in Persian documentary images using Improved Inpainting", Sixth Conference on Information and Knowledge Technology. Shahrood ut, Shahrood, Iran, 2014.
- [7] A. Criminisi, I. Reid, and A. Zisserman, "A Plane Measuring Device," University of Oxford, 1993.
- [8] 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.
- [9] J. Kanai, T. A. Nartker, S. Rice, and G. Nagy, "Performance metrics for document understanding systems," in Proceedings of the Second International Conference on Document Analysis and Recognition, 1993, pp. 424-427.
- [10] H. Khosravi and E. Kabir, "A blackboard approach towards integrated Farsi OCR system," International Journal of Document Analysis and Recognition (IJ DAR), vol. 12, pp. 21-32, 2009.
- [11] J. Liang, D. DeMenthon, and D. Doermann, "Geometric Rectification of Camera-captured Document Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, pp. 591-605, 2008.
- [12] L. Likforman-Sulem and F. Claudie, "Extracting text lines in handwritten documents by perceptual grouping," in Advances in handwriting and drawing: a multidisciplinary approach, paris, 1994, pp. 117-135.
- [13] L. Likforman-Sulem, A. Hanimyan, and C. Faure, "A Hough based algorithm for extracting text lines in handwritten documents," in IEEE Proceedings of the Third International Confer-

5-References

۵- مراجع

- [۱] ۱. بایسته تاشک ع. احمدی فرد و ح. خسروی, "یک روش دو مرحله ای برای بازشناسی کلمات دست نوشته فارسی به کمک بلوک بندی تطبیقی گرادیان تصویر", پردازش علائم و داده ها, دوره ۱۲, صفحات ۲۹-۱۵, ۱۳۹۴.
- [1] E. bayesteh Tashak, A. Ahmadyfard and H. Khosravi "A two-step method for recognizing Persian handwritten words using adaptive division of gradient image", JSDP, vol 12, PP 15-29, 2015.
- [۲] ح. حسن پور و ع. رستمی قادی "بهسازی تصویر با کاهش اثر عوامل خرابی بر مولفه ی روشنایی و بازتابش تصویر", پردازش علائم و داده ها, دوره ۹, صفحات ۲۳-۱۳, ۱۳۹۱.
- [2] H. Hasanpour and O. Rostami Ghadi "Image enhancement By Reducing the effect of failure factors on Intensity And reflection of the image" JSDP, vol 9, PP 12-23, 2012.
- [۳] س. خسروی راد, "رفع اعوجاجات غیرخطی در تصاویر اسناد فارسی", پایان نامه ی ارشد, دانشگاه صنعتی شاهرود, شاهرود, ایران, ۱۳۹۱.
- [3] S. Khosravi Rad, "Nonlinear distortion correction in Persian documentary images" M.S. thesis, Shahrood ut, Shahrood, Iran, 2012.
- [۴] ه. ده بوید, ف. رزازی, و ش. علیرضایی, "ارائه روشی نوین برای کاهش اعوجاج تصویربرداری در تصاویر متنی فارسی تصویربرداری شده توسط دوربین", ششمین کنفرانس ماشین بینایی و پردازش تصاویر, اصفهان, ایران, ۱۳۸۹.
- [4] H. Dehboyd, F. Razazi, Sh. Alirezei "Introducing a new method for reducing image distortion in Persian text images captured by the camera" Sixth Conference of the Machine and Image Processing, Esfahan, Iran, 2010.
- [۵] م. شامقلی, "رفع اعوجاج و بهبود کیفیت تصاویر اسکن شده از کتب فارسی", پایان نامه کارشناسی ارشد, دانشگاه صنعتی شاهرود, شاهرود, ایران, ۱۳۹۲.

[24] K. Y. Wong, R. G. Casey, and F. M. Wahl, "Document analysis system," *IBM journal of research and development*, vol. 26, no. 6, pp. 647-656, Nov. 1982.

[25] OmniPage. [Online]. <http://www.nuance.com>



محمد امین طلوع بیدختی مدرک

کارشناسی خود را در رشته مهندسی الکترونیک از دانشگاه صنعتی شاهرود در سال ۱۳۹۱ دریافت کرده است. ایشان مدرک کارشناسی ارشد خود را نیز ارسال سال ۱۳۹۳ در رشته الکترونیک دیجیتال در همان دانشگاه اخذ کرده‌اند.

زمینه‌های پژوهشی وی پردازش تصویر، پردازش سیگنال و شناسایی آماری الگو است.

نشانی رایانامه ایشان عبارت است از:

M.a.Tolou.b@Gmail.com



علیرضا احمدی فرد مدرک

کارشناسی را از دانشگاه صنعتی اصفهان در مهندسی الکترونیک و کارشناسی ارشد را از دانشگاه صنعتی امیرکبیر در مهندسی مخابرات اخذ کرده‌اند. ایشان مدرک دکترای تخصصی را در زمینه پردازش تصویر و ماشین بینایی از مرکز CVSSP دانشگاه Surrey در سال ۲۰۰۲ اخذ کردند.

زمینه‌های پژوهشی مورد علاقه دکتر احمدی فرد تشخیص و آشکارسازی اشیاء در تصاویر، تشخیص عیوب کالا در خط تولید، تشخیص هویت از روی بیومتریک‌ها و انطباق تصاویر است. در حال حاضر ایشان عضو هیئت علمی دانشگاه صنعتی شاهرود می‌باشند.

نشانی رایانامه ایشان عبارت است از:

Ahmadyfard@Shahroodut.ac.ir

ence on Document Analysis and Recognition, 1995, pp. 774-777.

[14] A. Masalovitch and L. Mestetskiy, "Usage of continuous skeletal image representation for document images de-warping," in *Proceedings of International Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, 2007, pp. 45-53.

[15] J. Mundy and A. Zisserman, *Geometric invariance in computer vision*. Cambridge, MA : MIT press, 1992, vol. 92.

[16] W. Niblack, "An introduction to digital image processing," Strandberg Publishing Company, 1985.

[17] A. H. Roger and C. R. Johnson, "Topics in matrix analysis," in *Matrix analysis*. Cambridge university press, 2012.

[18] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, vol. 32, no. 2, pp. 225-236, 2000.

[19] F. Shafait and M. Breuel, "Document image dewarping contest," in *2nd Int. Workshop on Camera-Based Document Analysis and Recognition*, Curitiba, Brazil, 2007, pp. 181-188.

[20] M. Shamgholi, M. H. Khosravi, and S. M. Riazi, "Document Image Dewarping Based on Text Line Detection and Surface Modeling," *International Journal of Engineering-Transactions C: Aspects*, vol. 27, no. 12, p. 1855, 2014.

[21] Z. Shi and V. Govindaraju, "Line separation for complex document images using fuzzy runlength," in *IEEE Proceedings in First International Workshop on Document Image Analysis for Libraries*, 2006, pp. 306-312.

[22] A. Ulges, C. H. Lampert, and T. Breuel, "Document capture using stereo vision," in *ACM Proceedings of the 2004 ACM symposium on Document engineering*, 2004, pp. 198-200.

[23] 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 Scanner Image," in *IEEE Proceedings in Fifth International Conference on Computer*, 1995, pp. 66-71.