

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه شهید باهنر کرمان

دانشکده فنی و مهندسی

بخش مهندسی برق

پایان نامه تحصیلی برای دریافت درجه کارشناسی ارشد رشته مهندسی برق  
مخابرات گرایش سیستم

---

رفع اعوجاج هندسی، پرسپکتیو و کجی تصاویر متن دوربینی فارسی

---

مؤلف :

احسان کیومرثی

استاد راهنما :

دکتر حسین نظام آبادی پور

استاد مشاور :

دکتر سعید سریزدی

بهمن ماه ۱۳۹۱



این پایان نامه به عنوان یکی از شرایط درجه کارشناسی ارشد به

**بخش مهندسی برق**

**دانشکده فنی و مهندسی**

**دانشگاه شهید باهنر کرمان**

تسلیم شده است و هیچگونه مدرکی به عنوان فراغت از تحصیل دوره مزبور شناخته نمی شود.

دانشجو: احسان کیومرثی

استاد راهنما: دکتر حسین نظام آبادی پور

استاد مشاور: دکتر سعید سر یزدی

داور ۱: دکتر مریم پور محی آبادی

داور ۲: دکتر مهدی افتخاری

نماینده تحصیلات تکمیلی دانشکده در جلسه دفاع: دکتر سیامک طالبی

معاون آموزشی و پژوهشی دانشکده: دکتر مریم احتشام زاده

**حق چاپ محفوظ و مخصوص به دانشگاه شهید باهنر کرمان است**

## تقدیم به :

همه کسانی که دوستشان دارم  
که به من آموختند دوست داشتن را...

و به

همه کسانی که دوستم دارند  
که بی وجودشان زندگی عبث خواهد بود.

کس نیست که این گوهر تحقیق بسفت

این بحر وجود آمده بیرون ز نهفت

زان روی که هست کس نمی داند گفت

هر کس سخنی از سر سودا گفته است

"خیام"

## تشکر و قدر دانی :

از زحمات بی دریغ استاد گرانقدر

جناب آقای دکتر حسین نظام آبادی پور

که راهنمایی هایشان مشعل فروزان راه دانش بوده است.

## چکیده :

تصاویر متن دوربینی غالباً با اعوجاج هندسی، پرسپکتیو و کجی همراه هستند. این اعوجاج‌ها، عملکرد سیستم‌های نویسه خوانی نوری را با مشکل روبرو می‌سازند. برای رفع این اعوجاج‌ها روش‌های مختلفی ارائه شده است. اما اغلب این روش‌ها بر متون لاتین تمرکز دارند و قابل استفاده مستقیم برای متون فارسی نمی‌باشند. در این پایان نامه روشی نوین جهت رفع اعوجاج هندسی، پرسپکتیو و کجی تصاویر متن دوربینی فارسی ارائه شده است. این روش شامل چهار مرحله کلی پیش پردازش، رفع اعوجاج کلی، تصحیح اندازه و رفع اعوجاج جزئی می‌باشد. اعوجاج‌های تصویر متن در دو مرحله رفع می‌شوند و یک مرحله تصحیح اندازه نیز برای جبران سازی فشردگی افقی تصویر صورت می‌گیرد. در مرحله رفع اعوجاج کلی، نمای دور تصویر متن مورد نظر است و مدل دوبعدی سند متنی با استفاده از خطوط حاشیه متن و مسیر خطوط متن به دست می‌آید. در مرحله رفع اعوجاج جزئی، کلمات به صورت تک تک بررسی شده و زاویه کجی آن‌ها با استفاده از تبدیل هاف تعیین می‌شود. نتایج آزمایش‌ها بر روی تصاویر متن فارسی نشان از ۷۳/۵٪ بهبود در عملکرد سیستم نویسه خوانی نوری دارد.

## کلید واژه‌ها :

پردازش تصاویر متن، نویسه خوانی نوری تصاویر متن دوربینی، رفع اعوجاج هندسی، تصاویر متن فارسی.

## فهرست مطالب :

۱	فصل ۱ : مقدمه .....
۲	۱-۱ انگیزش .....
۹	۲-۱ اهداف پایان نامه .....
۹	۳-۱ ساختار پایان نامه .....
۱۰	فصل ۲ : مروری بر کارهای گذشته .....
۱۱	۱-۲ مقدمه .....
۱۲	۲-۲ روش‌های خاص .....
۱۵	۳-۲ روش‌های عمومی .....
۲۷	فصل ۳ : بررسی یک الگوریتم رفع اعوجاج نمونه .....
۲۸	۱-۳ مقدمه .....
۲۸	۲-۳ ساختار کلی یک سیستم نمونه .....
۳۳	۳-۳ بررسی روش نمونه .....
۳۷	فصل ۴ : روش پیشنهادی .....
۳۸	۱-۴ مقدمه .....
۴۰	۲-۴ پیش پردازش .....
۴۲	۳-۴ رفع اعوجاج کلی .....
۵۷	۴-۴ تصحیح اندازه .....

۵۸ ..... ۵-۴ رفع اعوجاج جزئی (رفع کجی کلمات)

**۶۷ ..... فصل ۵: نتایج و تحلیل**

۶۸ ..... ۱-۵ مقدمه

۶۸ ..... ۲-۵ معرفی پایگاه داده

۶۸ ..... ۳-۵ نتایج آزمایش‌ها

۷۵ ..... ۴-۵ تحلیل نتایج

۷۷ ..... ۵-۵ جمع بندی و پیشنهادها

۷۹ ..... فهرست منابع

۸۶ ..... پیوست (واژه نامه)



## فهرست شکل ها :

- شکل ۱-۱ مراحل نویسه خوانی نوری ..... ۲
- شکل ۲-۱ دستگاه‌های ویژه‌ی تصویربرداری از اسناد ..... ۵
- شکل ۳-۱ تفاوت تصویر دوربینی با تصویر روبش شده؛ (الف) تصویر روبش شده، (ب-ج)
- تصاویر دوربینی ..... ۷
- شکل ۱-۲ استفاده از دستگاه‌های ویژه در عکس‌برداری از متن؛ (الف) مولد لیزر، دو دوربین و پایه دوربین [۱۶]، (ب) دو پروژکتور و دو دوربین [۱۵] ..... ۱۲
- شکل ۲-۲ مدل سازی سه بعدی سطح سند [۱۷]؛ (الف) تصویر اولیه سند، (ب) مدل مش سه بعدی سند، (ج) تصویر بدون اعوجاج ..... ۱۳
- شکل ۳-۲ مدل سازی سه بعدی به کمک دو تصویر از سند [۱۸]؛ (الف) دو تصویر از یک سند، (ب) مدل سه بعدی سند ..... ۱۳
- شکل ۴-۲ مدل سازی سه بعدی سند بر اساس غلظت سایه در تصویر [۲۹]؛ (الف) تصویر اولیه، (ب) ماسک بخش‌های چاپی، (ج) inpainting با استفاده از ماسک و توابع RBF، (د) سایه تصویر ..... ۱۶
- شکل ۵-۲ رفع اعوجاج به کمک چارچوب کلی متن [۳۵]؛ (الف) چارچوب متن در تصویر اولیه، (ب) تصویر بدون اعوجاج ..... ۱۷
- شکل ۶-۲ رفع اعوجاج با استفاده از دو ویژگی مسیر خط متن و جهت گیری VSB ها [۴۱]؛ (الف) مسیر خطوط متن، (ب) جهت گیری VSB ها، (ج) چهار جهت ممکن برای یک VSB، (د) تصویر بدون اعوجاج ..... ۱۹
- شکل ۷-۲ تفاوت نمودار افکنش از دو جهت و مکان مختلف [۴۹] ..... ۲۳
- شکل ۸-۲ به دست آوردن زاویه کجی با دوروش نوارهای عمودی و فضای خالی [۱۳]؛ (الف) تصویر اولیه، (ب) تشخیص بخش‌های متنی بر اساس طول بخش‌های عمودی، (ج) حذف خطوط

- عمودی بلند، (د) مسیر خطوط متن با استفاده از بخش‌های عمودی باقی‌مانده، (ه) فضای خالی بین کلمات و تکه‌های پیوسته..... ۲۶
- شکل ۱-۳ بخش‌های یک سیستم رفع اعوجاج ..... ۲۹
- شکل ۲-۳ اثرات اعوجاج پرسپکتیو ..... ۳۲
- شکل ۳-۳ اثر تبدیل‌های هندسی مختلف بر روی تصویر ..... ۳۲
- شکل ۴-۳ تاثیر گسترش به روش RLSA [۲]؛ (الف) تصویر خاکستری اولیه، (ب) تصویر دو سطحی، (ج) تصویر گسترش یافته ..... ۳۴
- شکل ۵-۳ مسیر خط متن و چارچوب تکه‌های پیوسته [۲]..... ۳۵
- شکل ۶-۳ نقاط تبدیل پرسپکتیو [۲] ..... ۳۶
- شکل ۷-۳ تصویر بدون اعوجاج [۲]؛ (الف) نوارهای عمودی، (ب) تصویر نهایی ..... ۳۶
- شکل ۱-۴ مراحل روش پیشنهادی ..... ۳۹
- شکل ۲-۴ دوسطحی سازی؛ (الف) تصویر اولیه، (ب) تصویر دوسطحی ..... ۴۱
- شکل ۳-۴ حذف بخش‌های غیر متنی ..... ۴۲
- شکل ۵-۴ مراحل رفع اعوجاج کلی ..... ۴۳
- شکل ۴-۴ یک کلمه نمونه ..... ۴۴
- شکل ۶-۴ تصویر بدون بخش‌های غیر متنی ..... ۴۶
- شکل ۷-۴ تصویر گسترش یافته با SE لوزی شکل ..... ۴۷
- شکل ۸-۴ تصویر هموار شده با فیلتر median ..... ۴۸
- شکل ۹-۴ خطوط گذرنده از نقاط بیشینه ..... ۴۹
- شکل ۱۰-۴ ماسک مورد استفاده برای گسترش تصویر در مرحله جداسازی خطوط متن ..... ۵۱
- شکل ۱۱-۴ تصویر گسترش یافته به وسیله ماسک شش ضلعی ..... ۵۲

- شکل ۴-۱۲ تصویر پس از مرحله سایش ..... ۵۲
- شکل ۴-۱۳ مدل تصویر متن دچار اعوجاج پرسپکتیو شدید ..... ۵۳
- شکل ۴-۱۴ چارچوب کلی متن (خط چین) ..... ۵۴
- شکل ۴-۱۵ یافتن نقاط متناظر تبدیل ..... ۵۵
- شکل ۴-۱۶ خطوط و کلمات پس از رفع اعوجاج کلی ..... ۵۶
- شکل ۴-۱۷ تصویر حاصل از تصحیح اندازه ..... ۵۸
- شکل ۴-۱۸ هموار سازی تصویر ..... ۵۹
- شکل ۴-۱۹ مراحل رفع اعوجاج جزئی ..... ۶۱
- شکل ۴-۲۰ جداسازی خطوط متن ..... ۶۲
- شکل ۴-۲۱ یک تکه کردن اجزای هر کلمه ..... ۶۲
- شکل ۴-۲۲ جداسازی کلمات و خطوط متن ..... ۶۳
- شکل ۴-۲۳ تعدادی از کلمات پیش (سمت چپ) و پس (سمت راست) از رفع اعوجاج جزئی به همراه زاویه‌ی تشخیص داده شده ..... ۶۴
- شکل ۴-۲۴ به دست آوردن مکان کرسی کلمات به وسیله افکنش افقی؛ (الف) یک کلمه نمونه، (ب) نمودار افکنش افقی ..... ۶۵
- شکل ۴-۲۵ تصویر نهایی ..... ۶۶
- شکل ۵-۱ نتایج تصویر ۱؛ (الف) تصویر اولیه، (ب) نتیجه روش مرجع [۲]، (ج) نتیجه روش پیشنهادی ..... ۶۹
- شکل ۵-۲ نتایج تصویر ۲؛ (الف) تصویر اولیه، (ب) نتیجه روش مرجع [۲]، (ج) نتیجه روش پیشنهادی ..... ۷۰
- شکل ۵-۳ نتیجه تصویر ۳؛ (الف) تصویر اولیه، (ب) نتیجه روش پیشنهادی ..... ۷۱
- شکل ۵-۴ نتیجه تصویر ۴؛ (الف) تصویر اولیه، (ب) نتیجه روش پیشنهادی ..... ۷۲

- شکل ۵-۵ نتیجه تصویر ۵؛ (الف) تصویر اولیه، (ب) نتیجه روش پیشنهادی ..... ۷۳
- شکل ۶-۵ نتیجه تصویر ۶؛ (الف) تصویر اولیه، (ب) نتیجه روش پیشنهادی ..... ۷۴
- شکل ۷-۵ نتیجه تصویر ۷؛ (الف) تصویر اولیه، (ب) نتیجه روش پیشنهادی ..... ۷۵
- شکل ۸-۵ خطا در رفع کجی کلمات؛ (الف) کلمات پایین خط پایه، (ب) کلمات بالای خط پایه،  
(ج) کلمات لاتین ..... ۷۷

# فصل اول

## مقدمه

این پایان نامه در رابطه با پردازش تصویر متون چاپی<sup>۱</sup> تهیه شده است. در واقع موضوع این پایان نامه بخشی از پروسه نویسه خوانی نوری<sup>۲</sup> است. در شکل ۱-۱ مراحل مختلف نویسه خوانی نوری نشان داده شده است.

نویسه خوانی توری که با سرواژه‌ی OCR شناخته می‌شود، عبارت است از بازشناسی خودکار متون موجود در تصاویر اسناد و تبدیل آنها به متون قابل جستجو و ویرایش توسط رایانه. تصویر سند غالباً توسط روبش‌گرها یا دوربین‌های دیجیتال تولید می‌شوند و شامل تعدادی پیکسل با رنگ‌های مختلف و سطوح روشنایی گوناگون است. از دید انسان، یک سند ممکن است ارزش اطلاعاتی زیادی داشته باشد، لیکن از دید رایانه تصویر یک سند با تصویر یک منظره تفاوتی ندارد، چرا که هر دوی آنها مجموعه‌ای از پیکسل‌ها هستند. برای اینکه بتوان از اطلاعات نوشتاری تصویر سند استفاده کرد، باید به نحوی نوشته‌های موجود در سند را بازشناسی کرد. چنین کاری توسط نرم‌افزارهای نویسه‌خوان نوری انجام می‌شود.



شکل ۱-۱ مراحل نویسه خوانی نوری

## ۱-۱ انگیزش

گسترش روز افزون تلفن‌های همراه مجهز به دوربین، با دوربین‌هایی با درجه تفکیک بالا، باعث شده تا تهیه تصاویر از مناظر مختلف از جمله متن کتاب‌ها و مجلات، جهت جمع آوری اطلاعات امری عادی و همگانی شود. از طرفی حجم بالای این تصاویر و عدم امکان ویرایش و جستجوی متن موجود در آنها، کمی از ارزش این شیوه جمع آوری اطلاعات می‌کاهد. در اینجا بحث نویسه خوانی نوری این گونه تصاویر به میان می‌آید. سیستم‌های نویسه خوان نوری قادر به تشخیص و استخراج متن از تصاویر متنی هستند که دچار اعوجاج هندسی، پرسپکتیو و کجی نباشند [۵]؛ مشکلاتی که در اکثر تصاویر متن دوربینی وجود دارد.

<sup>1</sup> printed text

<sup>2</sup> Optical Character Recognition (OCR)

همان گونه که در شکل ۱-۱ مشخص است، تصاویر متن را می‌توان به وسیله روبش‌گر و دوربین دیجیتال تهیه نمود. رشد سریع دسترسی به دستگاه‌های عکس برداری دیجیتال ارزان با کیفیت و قابل حمل یک فرصت برای تغییر روش سنتی تصویر برداری از متن به وسیله روبش‌گر ایجاد کرده است. چه دوربین‌های عکاسی دیجیتال و چه دوربین‌های دیجیتالی که بر روی تلفن‌های همراه و رایانه‌های قابل حمل<sup>۳</sup> تعبیه شده‌اند، این قابلیت را دارند که تصاویر کتاب‌ها، نسخ خطی و تاریخی آسیب‌پذیر<sup>۴</sup> و به طور کلی هر جایی که متن نوشته شده باشد را ثبت کنند و این امر، کارایی بیشتر این دستگاه‌ها را نسبت به روبش‌گرهای رومیزی<sup>۵</sup> نشان می‌دهد.

بنابراین اگر مشکل پردازش و تجزیه و تحلیل متون بدست آمده از این دستگاه‌ها رفع شود، تقاضای زیادی برای استفاده از آن‌ها در این زمینه وجود خواهد داشت. پژوهش‌های انجام گرفته بر روی تصاویر متن روبش شده، مرجع و نقطه شروع خوبی برای رسیدن به این هدف است. اما این روش -ها مستقیماً نمی‌توانند برای تصاویر دوربینی مورد استفاده قرار گیرند. در تصاویر متن دوربینی کاستی‌هایی وجود دارد، از جمله اعوجاج هندسی<sup>۶</sup>، اعوجاج پرسپکتیو<sup>۷</sup>، درجه تفکیک<sup>۸</sup> کم، محوشدگی<sup>۹</sup> تصاویر، پس زمینه پیچیده<sup>۱۰</sup> و ...

پردازش و تجزیه و تحلیل متن نسبت به دیگر مسائل بینایی ماشین<sup>۱۱</sup> و پردازش تصویر<sup>۱۲</sup> از جایگاه ویژه‌ای برخوردار شده است. این تصاویر معمولاً به صورت سیاه و سفید مورد پردازش قرار می‌گیرند. در ابتدای دهه ۶۰ میلادی نویسه‌خوانی نوری (OCR) به عنوان یکی از ابتدایی‌ترین کارکردهای بازشناسی الگو<sup>۱۳</sup> مورد توجه قرار گرفت؛ مسأله‌ای که امروزه حل شده در نظر گرفته می‌شود. اما متأسفانه هنوز مسائلی در این زمینه به صورت حل نشده باقی مانده‌اند که حل آن‌ها می‌تواند بسیاری از نیازهای امروز کاربران را مرتفع سازد. مسائلی مثل پیچیدگی محتوای تصاویر متن<sup>۱۴</sup>، وجود نویز در این گونه تصاویر یا فونت‌های مختلف نوشتاری، راه توسعه را در این زمینه باز گذاشته‌اند. پردازش و تجزیه و تحلیل تصاویر متن در طی پنجاه سال گذشته مورد مطالعه بسیار قرار گرفته است [۵].

---

<sup>3</sup> Personal Digital Assistant (PDA)

<sup>4</sup> historical pages too fragile to touch

<sup>5</sup> desktop scanner

<sup>6</sup> geometric distortion

<sup>7</sup> perspective distortion

<sup>8</sup> resolution

<sup>9</sup> blurring

<sup>10</sup> complex background

<sup>11</sup> machine vision

<sup>12</sup> image processing

<sup>13</sup> pattern recognition

<sup>14</sup> complex content

بعضی از زیر مجموعه‌های این رشته عبارتند از پیش پردازش‌های تصاویر، تجزیه و تحلیل طرح بندی فیزیکی و منطقی، نویسه خوانی نوری حروف و نویسه خوانی هوشمندانه نوری حروف<sup>۱۵</sup>، تحلیل تصاویر گرافیکی<sup>۱۶</sup>، پردازش فرم<sup>۱۷</sup>، تأیید امضاء<sup>۱۸</sup>، تشخیص هویت نویسنده<sup>۱۹</sup> و ...

اخیراً استفاده از دوربین‌های عکاسی دیجیتال برای تهیه تصویر متن مورد توجه قرار گرفته است. دستگاه‌های تصویر برداری و عکس برداری دیجیتال، رایانه‌های قابل حمل و حتی تلفن‌های همراه این قابلیت را از خود نشان داده‌اند که به عنوان جایگزین روبش‌گرها مورد توجه قرار گیرند. اگر چه به طور کامل نمی‌توان روبش‌گرها را با این دستگاه‌ها جایگزین کرد، اما قابلیت‌هایی مثل وزن کم، ابعاد کوچک، قابلیت اتصال آسان به شبکه‌های رایانه‌ای و قابلیت تصویر برداری از هر محیط، مزایایی هستند که نمی‌توان آن‌ها را نادیده گرفت. این قابلیت‌ها منجر به استفاده زیاد از آن‌ها جهت تصویر برداری از متون در هر جایی و با هر شرایطی می‌شود.

بخش صنعت هم این نکته را درک کرده و OCR تصاویر دوربینی را در کنار OCR تصاویر روبش شده مورد توجه قرار داده است. برای مثال شرکت XEROX نرم افزاری را در این زمینه ارائه داده است. از آنجایی که تصاویر ویدئویی دارای درجه تفکیک کمتری هستند، به فناوری پردازش‌گر پیشرفته‌تری نیاز دارند. در حال حاضر بیشتر پژوهش‌ها بر روی تصاویر گرفته شده با دوربین عکاسی دیجیتال تمرکز یافته‌اند و کار بر روی تصاویر ویدئویی را می‌توان در آینده بیشتر مورد توجه قرار داد [۵].

#### - دستگاه‌های عکس برداری

در دهه‌های گذشته روبش‌گرهای دیجیتال، در زمینه ثبت تصاویر متن، دستگاه‌های تصویربرداری غالب بوده‌اند. انواع مختلفی از این دستگاه‌ها ساخته شده است. از روبش‌گرهای در اندازه‌های بزرگ برای تصویر برداری از طرح‌های مهندسی گرفته تا روبش‌گرهای کوچک رومیزی جهت استفاده‌های اداری و خانگی. دستگاه‌هایی با سرعت‌های مختلف: از چندین صفحه در ثانیه تا یک خط در ثانیه.

حتی بعضی این قابلیت را دارند که صفحات کتاب را ورق زده و یکی یکی تصویربرداری کنند. درجه تفکیک این دستگاه‌ها اخیراً از مرز 4800 dpi (dot per inch) نیز گذشته است. در حالی که قیمت آن‌ها کمتر از ۱۰۰ دلار است. همین مزیت باعث شده که در کنار هر رایانه شخصی

---

<sup>15</sup> Intelligent Character Recognition (ICR)

<sup>16</sup> graphics analysis

<sup>17</sup> form processing

<sup>18</sup> signature verification

<sup>19</sup> writer identification



معمولاً یک دستگاه روبش گر باشد. معمولاً برای پردازش تصاویر اسناد درجه تفکیک 300 dpi کفایت می کند.

از آنجایی که تصاویر روبش شده متن از کیفیت کافی برخوردارند، OCR این گونه تصاویر تقریباً یک مساله حل شده قلمداد می شود. دوربین های دیجیتال ویژه در زمینه تصویر برداری از متون بسیار مورد استفاده قرار گرفته اند، در حالی که مطالعه بر روی این تصاویر بسیار کم است. این گونه دوربین ها در زمره دوربین های صنعتی قرار می گیرند و از کیفیت و قیمت بالایی برخوردار هستند. در این دستگاه ها شرایطی مهیا می شود که اسناد به طور صاف و در شرایط نوری مناسب باشند.

هدف از ساختن چنین دوربین هایی ثبت تصاویر از متون و اسنادی است که به راحتی نمی توان با روبش گر آن ها را روبش کرد مانند: کتاب های قطور و اسناد تاریخی آسیب پذیر. همراه با بعضی از این دستگاه ها نرم افزارهایی ارائه می شود که قادر به بازیابی متون دارای اعوجاج می باشند. اخیراً بسیاری از شرکت ها برای تجاری کردن این دوربین ها سرمایه گذاری کرده اند. در حالی که برخی



شکل ۱-۲ دستگاه های ویژه تصویربرداری از اسناد

دیگر برای ثبت تصاویر با قابلیت اطمینان بیشتر، به جای دوربین از روبش گر استفاده کرده اند. شکل ۱-۲ چند نمونه از این دستگاه ها را نشان می دهد.

#### - دوربین های عکاسی دیجیتال

پیشرفت سریع دوربین های دیجیتال در ۱۰ سال گذشته موقعیت مناسبی را در اختیار افراد عادی قرار داده است تا تصاویر مختلف را ثبت کنند. امروزه تقریباً تمام مدل های تلفن همراه، دارای دوربین دیجیتال هستند. قابلیت مهم این دوربین ها سازگاری<sup>۲۰</sup> آن ها است. آن ها به کوچکی یک کارت اعتباری، مقاوم در آب و هوای مختلف، قابل حمل بوده و استفاده از آن ها آسان است. همچنین معمولاً قابلیت عکس برداری و فیلم برداری را توأم دارند. اگر چه تصاویر متن به دست

<sup>20</sup> flexibility

آمده از آن‌ها ممکن است دارای مشکلاتی از جمله انحنای تصویر<sup>۲۱</sup>، محو شدگی ناشی از لرزش دوربین و توزیع نور غیریکنواخت<sup>۲۲</sup> باشد. تصاویر متن ممکن است به صورت تصویر ویدئویی باشد نه یک تک تصویر. در نتیجه یک سری پردازش‌های اضافی بر روی تصاویر حاصل از این دستگاه‌ها مورد نیاز است. این حقیقت که محیطی که این دستگاه‌ها در آن‌ها مورد استفاده قرار می‌گیرند با محدودیت محیط‌های مورد استفاده در روبش‌گرها یا دوربین‌های صنعتی مواجه نیستند چالش جدیدی را نمایان می‌سازد.

در حال حاضر درجه تفکیک دوربین‌های دیجیتال تجاری تا ۱۶ M pixel می‌رسد. در شرایط ایده‌آل این درجه تفکیک برای تجزیه و تحلیل و پردازش تصاویر متن کافی است. دوربین‌های دیجیتال فیلم‌برداری در حال حاضر از درجه تفکیک (۱۲۸۰×۷۲۰) بهره‌مندند. چون اصولاً تصاویر این گونه دوربین‌ها برای پخش و ارسال ماهواره‌ای<sup>۲۳</sup> با پهنای باند کم مورد استفاده قرار می‌گیرند و معمولاً فشرده سازی شدیدی روی آن‌ها صورت می‌گیرد. این حقیقت که آن‌ها اصولاً برای تصویر برداری از متن ساخته نشده‌اند یک چالش مهم را نشان می‌دهد. در نهایت هدف این است که بتوان پردازش‌های مربوط به متن را مستقیماً از طریق همان دستگاه‌های تصویر برداری انجام داد.

- مزایای استفاده از دوربین

تجزیه و تحلیل تصاویر دوربینی مزایای زیادی نسبت به تصاویر روبش شده دارد. دوربین‌ها کوچک، قابل حمل و قابل استفاده آسان هستند. آن‌ها قابل استفاده در هر شرایطی هستند و می‌توانند تصاویر متنی تهیه کنند که روبش آن‌ها مشکل یا حتی غیرممکن است، مثل صفحات بزرگ روزنامه، کتاب یا نوشته روی یک ساختمان، ماشین و ... به طور کلی سیستم‌های دوربینی سازش پذیری بالایی دارند. مطالعه‌ای که اخیراً انجام شده، نشان می‌دهد که سیستم‌های OCR دوربینی می‌توانند کاربرد بیشتری نسبت به سیستم‌های OCR که بر مبنای تصاویر روبش شده کار می‌کنند داشته باشند [۶].

در مطالعه دیگری امکان جایگزینی روبش‌گرهای مورد استفاده در میدان‌های جنگ، با دوربین‌های دیجیتال بررسی شده است [۷]. بر اساس این تحقیق، روبش‌گرهای sheet-fed<sup>۲۴</sup> امکان تصویر برداری از کتاب‌های قطور را ندارد. هم چنین نگه‌داری و جابجایی آن‌ها برای سربازان در میدان جنگ مشکل است.

---

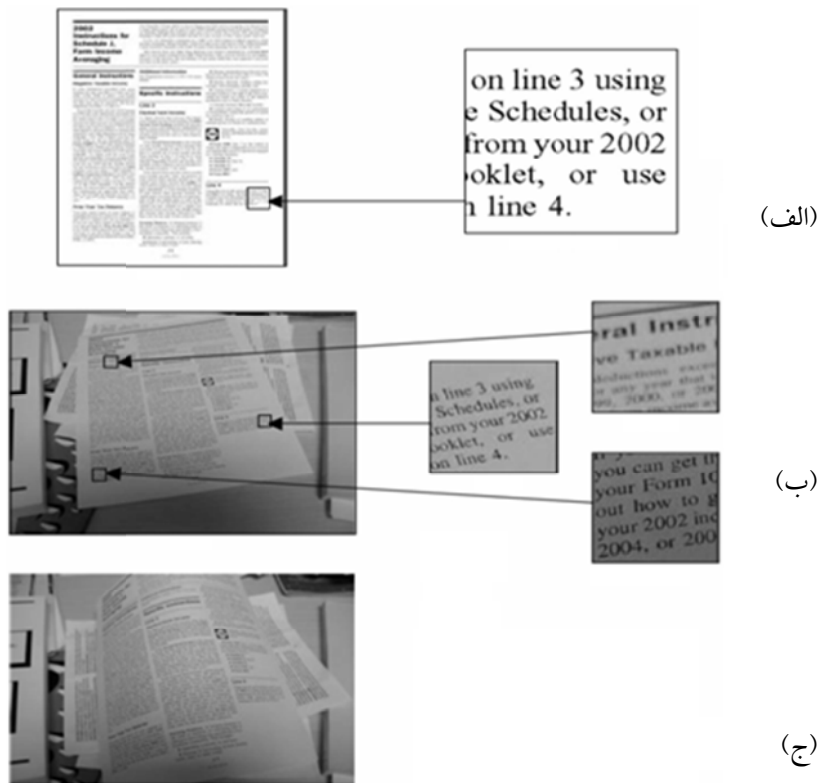
<sup>21</sup> curved surface, non-flat documents

<sup>22</sup> uneven lighting

<sup>23</sup> digital video broadcasting

<sup>24</sup> sheet-fed scanner

با توجه به تفاوت‌هایی که این دو گونه تصویر دارند، نحوه پردازش آنها نیز متفاوت خواهد بود. در تصاویر متن دوربینی امکان وجود اعوجاج پرسپکتیو، هندسی و هم چنین کجی<sup>۲۵</sup>، بسیار بیشتر از تصاویر روبش شده است (شکل ۱-۳).



شکل ۱-۳ تفاوت تصویر دوربینی با تصویر روبش شده؛ (الف) تصویر روبش شده، (ب-ج) تصاویر دوربینی

همچنین شرایط نوری محیط نیز بر تصاویر دوربینی ممکن است تاثیر منفی بگذارد. از طرفی، نویسه خوان‌های نوری موجود، تنها تصاویری را پردازش می‌کنند که اعوجاجی در آنها وجود نداشته باشد. از این رو پیش از بازشناسی این تصاویر نیاز است که رفع اعوجاج روی آنها صورت‌گیرد. رفع اعوجاج‌های پرسپکتیو، هندسی و هم چنین کجی تصاویر متن از جمله این پیش‌پردازش‌ها است. در سال‌های اخیر، پژوهش‌های زیادی در این زمینه صورت گرفته است. اما بیشتر این پژوهش‌ها در زمینه تصاویر متن لاتین است و در مورد تصاویر متن فارسی تنها می‌توان به مراجع [۱۴-۹ و ۳-۱] اشاره کرد. با توجه به تفاوت‌هایی که بین دستور خط فارسی و لاتین وجود دارد، بسیاری از روش‌های ارائه شده برای تصاویر متن لاتین، در مورد تصاویر متن فارسی کارایی ندارد [۱۴]. از این رو ضروری است در این زمینه روش‌هایی متناسب با ویژگی‌های متن فارسی

<sup>25</sup> skew

ارائه گردد. در ادامه به برخی ویژگی‌های دستور خط فارسی و لاتین اشاره شده است. موارد الف تاج عیناً از صفحه تارنمای فرهنگستان زبان و ادب فارسی نقل شده است. الف) در خط فارسی غالباً یک حرف به دو یا چند صورت نوشته می‌شود و این بستگی به جایگاه آن حرف در کلمه دارد:

- ء / ذ / و / ا
- ب / با / ب / پ / پ / پ / پ ؛ ت / تا / ت / ت
- ج / جا / ج / چ / چ / چ / چ ؛ ح / ح / ح / ح ؛ خ / خ / خ / خ
- س / سا / س / س ؛ ش / ش / ش / ش / ش
- ص / ص / ص / ص ؛ ض / ض / ض / ض / ض
- ع / ع / ع / ع ؛ غ / غ / غ / غ / غ
- ف / ف / ف / ف ؛ ق / ق / ق / ق / ق
- ک / ک / ک / ک ؛ گ / گ / گ / گ / گ
- ل / ل / ل / ل / ل
- م / م / م / م / م
- ن / ن / ن / ن / ن
- ه / ه / ه / ه / ه
- ی / ی / ی / ی / ی

ب) حروف فارسی دو دسته است:

- منفصل (پیوندناپذیر) که به حرف بعد از خود نمی‌چسبد (ا، د، ذ، ر، ز، ژ، و)؛
- متصل (پیوندپذیر) که به حرف بعد از خود می‌چسبد.

ج) علاوه بر حروف، در خط فارسی نشانه‌های دیگری به شرح زیر وجود دارد:

- حرکات یا مصوّت‌های کوتاه (ـَ، ـِ، ـُ) مانند پَر، پِر، پُر؛
- مدّ ( ~ ) روی الف، مانند آرد، مآخذ؛
- تشدید ( ّ ) مانند عدّه، پلّه، بقالی، ارّه؛
- سکون ( ° ) مانند لب‌تشنه؛
- یای کوتاه ( ء ) مانند نامه من؛
- تنوین ( ً، ٍ، ٌ )، مانند ظاهراً، عبارۀ اُخری، مضاف‌الیه.