

بررسی چند روش آستانه‌گیری برای بهبود تصاویر نسخ خطی فارسی

فرشید باباپور مفرد^۱، الناز قربانی ممقانی^۲ و میترا ابراهیمی^۳

دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران، تهران، ایران

farshid.mofrad@yahoo.com^۱

ghorbani6987@gmail.com^۲

mitraEbrahimi@gmail.com^۳

چکیده

در این مقاله به بررسی و مقایسه سه روش Otsu، Shafait و Sauvola برای بهبود تصاویر نسخ خطی فارسی می‌پردازیم. نسخه‌های خطی فارسی با قدمت چند صد سال با وجود حفظ و نگهداری در کتابخانه‌ها، واضح و خوانا نیستند. یافتن روش‌هایی برای حفظ این متون ارزشمند که نشان‌دهنده میراث فرهنگی هر کشوری می‌باشد موضوعی است که نظر محققین را در این زمینه به خود جلب کرده است. نتایج شبیه‌سازی، کارایی بالای روش Shafait رانسبت به دو روش دیگر در کاهش زمان پردازش و بهبود کنتراست نشان می‌دهد.

کلمات کلیدی

پردازش تصویر، بهبود تصویر، پس زمینه‌های آسیب دیده، انتگرال تصویر، باینری (دودویی) کردن، آستانه‌گیری محلی تطبیقی.

۱- مقدمه

آستانه‌گیری مرمت می‌کنند. هدف دودویی کردن متن تبدیل تصویر ورودی داده به نمایش دوسطحی سفید و سیاه است. تکنیک‌های دودویی کردن برای اسناد می‌توانند در دو گروه بزرگ طبقه بندی شوند: ۱- دودویی کردن سراسری ۲- دودویی کردن محلی. روش‌های دودویی کردن سراسری مانند Otsu [2] سعی در یافتن تنها یک مقدار آستانه برای تمام تصویر دارند. سپس هر پیکسل در پیش‌زمینه و پس‌زمینه صفحه بر پایه مقدار خاکستری اش علامت گذاری می‌شود. روش‌های دودویی کردن سراسری خیلی سریع هستند و نتایج خوبی را برای معمولی بدست می‌دهند. اگر شدت روشنایی روی تصویر یکنواخت نباشد برای مثال در مورد صفحات اسکن شده کتاب، روش‌های دودویی کردن سراسری، باعث تولید نویز در طول لبه‌های صفحه می‌شود. روش‌های دودویی کردن محلی مانند Sauvola [3] سعی در غلبه بر این مشکل‌ها با محاسبه آستانه‌های منحصر بفرد برای هر پیکسل با استفاده از اطلاعات همسایه‌های محلی اطراف پیکسل دارند. این روش‌ها معمولاً قادر به بدست آوردن نتایج خوبی روی تصاویر که شدیداً تخریب شدند هستند، اما اغلب آنها کند هستند [4].

مقایسه تکنیک‌های مختلف برای دودویی کردن متن در گذشته مورد توجه قرار گرفته بود. Trier [5] روش دودویی کردن تطبیقی محلی را برای تصاویر خاکستری با کنتراست کم، شدت و نویز متغیر در پس زمینه ارائه داده است. ارزیابی آنها نشان می‌دهد که روش Niblack [6] بهتر از سایر روش‌های آستانه‌گیری محلی اجرا می‌شود. Badekas [7] هفت الگوریتم مختلف را

نسخه‌های خطی باستان با قدمت صدها تا هزاران سال اغلب دارای پس زمینه‌ی بد یا آسیب دیده هستند. عوامل متعددی مانند عوامل بیولوژیکی، عوامل فیزیکی، عوامل شیمیایی و ... باعث این آسیب‌ها می‌شوند. علاوه بر مسائل زیست محیطی، غفلت انسان نیز منجر به نابودی نسخ خطی فارسی می‌شود. با توجه به این موضوع بسیار مهم، به منظور افزایش بقای عمر نسخ خطی و امکان استفاده‌ی آسان، تهیه اسکن و میکروفیلم و عکس دیجیتال از نسخ مزبور ضروری است. با این کار، دوام و کیفیت نسخه‌های خطی قدیمی تضمین شده و می‌تواند توسط اشخاص علاقمند و نسل‌های آینده به اشتراک گذاشته شود. کتابخانه‌ی ملی ایران نیز فعالیت‌هایی برای حفظ نسخه‌های خطی قدیمی فارسی انجام داده است. از جمله‌ی این اقدامات می‌توان به مرمت منابع آسیب دیده (مرمت سنتی و مرمت با استفاده از دستگاه ریخته‌گری)، صحافی، خدمات رقمی سازی و تهیه‌ی اسکن از اصل آثار اشاره کرد [۱]. با این حال، همه‌ی نسخه‌های خطی از طریق این فرایندها، واضح و قابل خواندن نمی‌شوند. گاهی اوقات، این اسناد دارای مشکلات کنتراست در پس زمینه هستند با رنگ جوهر آسیب دیده است. از این رو، به منظور بهبود نسخه‌های خطی، اشخاص علاقمند باید روش‌های پردازش تصویر دیجیتال را بکار گیرند. به طور کلی، برخی از محققان در ابتدا تصویر نسخه‌ی خطی را از یک پس زمینه‌ی تمیز و عاری از نویز با استفاده از باینری کردن یا فرایند