مجله دانشکده پزشکی اصفهان **مقاله پژوهشی**
سال بیست و نهم/شماره ۱۷٤/ ویژه نامه (مهندسی پزشکی) ۱۳۹۰
تاریخ دریافت: ۹۰/۸/۱۵ **تاریخ پذیرش:** ۹۰/۱۱/۱۷

استفاده از الگوریتم آنالیز اجزای اصلی جهت تشخیص نرم‌افزاری حروف در متون فارسی

شهروز گشمرد[۱]، دکتر علیرضا مهری دهنوی[۲]، دکتر حسین ربانی[۳]

چکیده

مقدمه: ابداع روش منابع گویای خودکار، همواره یکی از دغدغه‌های فکری پژوهشگران بوده است تا ابزاری کارآمد جهت مطالعه‌ی متون معمولی توسط نابینایان تولید نمایند. در این راه، علاوه بر تلاش در جهت ساخت ابزارها، روش‌های بهینه جهت پردازش، شناخت و تفکیک نوشته‌ها و متون فارسی نیز مورد نیاز می‌باشد.

روش‌ها: در این مقاله با استفاده از روش آنالیز اجزای اصلی، روشی جدید جهت تشخیص اتوماتیک حروف در تصاویر تهیه شده از حروف فارسی تایپ شده ارایه می‌گردد که قادر به بهبود روش‌های موجود از لحاظ دقت و سرعت می‌باشد. در این روش، که برگرفته از روش آنالیز اجزای اصلی می‌باشد، در اصل اجزای وابسته به تغییرات کم از داده‌ها حذف می‌شود و بدون از دست رفتن قسمت عمده‌ی اطلاعات تصویر، می‌توان شباهت‌ها و اختلافات بین تصاویر مختلف را با سرعت و دقت بیشتری مورد بررسی قرار داد.

یافته‌ها: در این مقاله با انجام آزمایش‌های عملی متعدد، کارآیی بهتر روش اجزای اصلی در مقایسه با روش‌های قبلی مانند روش‌های آماری و زرنیک سریع مورد تأیید قرار گرفته است. برای مثال با آزمایش هر سه این روش‌ها بر یک مجموعه‌ی داده‌ی آموزشی یکسان و چهار مجموعه‌ی داده‌ی آزمونی یکسان، مشخص گردید که دقت این روش در مقایسه با روش آماری ۱/۷۷۵ درصد بهتر و سرعت آن ۷/۵ برابر شده است و در مقایسه با روش زرنیک سریع دقت ۲/۲ درصد بهبود داشته و سرعت آن نیز ۵/۱۲ برابر بیشتر شده است.

نتیجه‌گیری: مقایسه‌ی نتایج نهایی اعمال داده‌ها به الگوریتم‌های مطرح شده در مقاله و مقایسه‌ی دقت و زمان آن‌ها نشان داد که الگوریتم آنالیز اجزای اصلی (PCA) ترکیبی بهترین نتایج از لحاظ سرعت و دقت را به همراه دارد.

واژگان کلیدی: الگوریتم ترکیبی و آنالیز اجزای اصلی، پردازش تصویر، شناسایی حروف فارسی

مقدمه

چندی است پژوهشگران به فکر ابداع روش منابع گویای خودکار افتاده‌اند تا ابزاری کارآمد جهت مطالعه ی متون معمولی توسط نابینایان تولید نمایند (۲-۱). از کاربردهای دیگر این ابزار می‌توان به پخش صوتی متن به جای خواندن آن اشاره کرد که با رواج این روش، مفهوم مطالعه به کلی دگرگون خواهد شد. به علاوه با این کار می‌توان در نگهداری و ذخیره‌سازی اطلاعات نیز از این روش استفاده نمود، زیرا با تبدیل متون نوشتاری معمولی به اطلاعات دیجیتال قابل استفاده برای کامپیوترها، بدون نیاز به تایپ دستی آن‌ها، نگهداری این اطلاعات بسیار راحت‌تر، مطمئن‌تر و کم هزینه‌تر خواهد بود.

تحقیقات در زمینه‌ی شناسایی حروف عربی (فارسی) از سال ۱۹۸۰ میلادی شروع شد و هنوز ادامه دارد. در این سال تعدادی از کارهای اولیه توسط امین و همکاران (۴-۳) در دانشگاه نانسی فرانسه آغاز گردید که در آن از یک سیستم بلادرنگ به نام شناسایی متقابل استفاده شد. پس از آن بدیع و Shimura (۵) از ایده‌ی دنبال کردن خط مرزی حروف و شناسایی از روی

[۱] کارشناس ارشد، گروه فیزیک و مهندسی پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران
[۲] دانشیار، گروه فیزیک و مهندسی پزشکی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران
[۳] استادیار، گروه فیزیک و مهندسی پزشکی، دانشکده‌ی پزشکی، مرکز تحقیقات پردازش تصویر و سیگنال پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران
نویسنده‌ی مسؤول: دکتر علیرضا مهری دهنوی
Email: mehri@med.mui.ac.ir

www.mui.ac.ir

منحنی‌های اجزا استفاده نمودند. امینی و مسینی (۶) یک سیستم بادرنگ پیشنهاد نمودند که به طور جداگانه کار جداسازی و تشخیص حروف را انجام می‌داد. بعد Hachour (۷) با تلفیق منطق فازی و سیستم‌های ماهر به دقت بالاتری از سایرین دست یافت. به مرور تکنیک‌های کارآتری در این زمینه ارایه گردید مانند هاشمی و همکاران (۸) و برومندنیا و شنبه‌زاده (۹) که اولی از روش آماری و دومی از روش زرنیک سریع استفاده نمودند؛ هر دو روش در مقایسه با سایر روش‌های قبلی بهترین نتایج را به دنبال داشتند (۹-۸).

در این طرح پژوهشی سعی بر آن بود تا روشی کارآمد به نام آنالیز اجزای اصلی در مرحله ی تشخیص حروف به کار گرفته شود. این روش در کار تشخیص الگوها به خوبی عمل کرده و نتایج مطلوبی ارایه نموده است (۱۰). استفاده از این روش می‌تواند در نرم‌افزار استفاده شده در ابزار منابع گویای خودکار، سرعت پردازش را بالا ببرد و دقت بیشتری اعمال نماید.

کار اصلی ابزار منابع گویای خودکار، همان عمل تبدیل تصاویر تهیه شده از متون به گفتار خواهد بود. مراحل مختلف کار تبدیل تصویر یک متن به گفتار را می‌توان به طور خلاصه در مراحل زیر تشریح نمود:

ابتدا استفاده کننده‌ی ابزار، عکسی از متن تهیه می‌کند. برای مثال، اگر این ابزار به شکل یک دوربین دیجیتال ساخته شده باشد، آن‌گاه استفاده کننده، که یک نابینا است، کافی است بداند که محل تقریبی این متن کجاست. در گام بعد، پردازش اولیه‌ی سخت‌افزاری و نرم‌افزاری بر روی تصویر انجام می‌گیرد تا اگر در تصویر تهیه شده اشکالاتی مانند پراش رنگی، لرزش دست، آلودگی به نویز یا چرخش تصویر و غیره بروز نموده باشد، بتوان آن‌ها را حذف نمود. از آنجا که عکس تهیه شده توسط استفاده کننده لزوما محدود به متون نمی‌باشد، جدا کردن متن از تصویر در گام بعدی صورت می‌گیرد. سپس جداسازی خطوط و کلمات و حروف از یکدیگر با استفاده از یک الگوریتم بهینه (۸) انجام می‌شود و حروف جدا شده‌ی مربوط به هر کلمه جهت تشخیص به مرحله‌ی بعد ارسال خواهند شد.

در مرحله‌ی تشخیص حروف، یک به یک حروف مجزا شده توسط الگوریتمی تشخیص داده خواهند شد؛ به این معنا که مشخص می‌شود این حرف یا کاراکتر معادل کدامین حرف الفبایی (یا به عبارت صحیح‌تر معادل چه کلاس شناسایی) می‌باشد. خروجی این الگوریتم یا نرم‌افزار می‌تواند کد اسکی مربوط به این حروف تشخیص داده شده باشد.

کدهای خروجی برای هر کلمه جداگانه صادر می شود و به حافظه ارسال می‌گردد. پس از آن کدهای تشخیص داده شده به هم متصل می‌شوند و کدهای به هم پیوسته متناسب با اصول گفتاری و آوایی مربوط به زبان مورد نظر، به حالت گفتاری تبدیل می‌شود. خروجی این مرحله یک سیگنال دیجیتال خواهد بود که در نهایت توسط نرم‌افزار و سخت‌افزار صوتی، این سیگنال به یک سیگنال الکتریکی آنالوگ تبدیل می‌شود و یک بلندگو سیگنال الکتریکی نهایی را به صوت قابل شنود توسط استفاده کننده تبدیل می‌کند.

همانگونه که در بالا توضیح داده شد، الگوریتم می‌بایست حروف از هم مجزا شده (که به صورت یک سری داده‌های تصویری پردازش شده هستند) را تشخیص دهد و کدهای آن‌ها را برای مراحل بعدی ارسال نماید (الگوریتم نرم‌افزاری تشخیص حروف). این مرحله به دلایل زیر یکی از بخش‌های اصلی

www.mui.ac.ir

نرم‌افزارهای تبدیل تصویر متن به گفتار در زبان فارسی می‌باشد:

- حروف فارسی بنا به موقعیتشان در کلمه اشکال مختلفی به خود می‌گیرند. مثلا حرف ب اگر جداگانه، در ابتدا، میانه یا انتهای کلمه قرار گیرد به چهار شکل مختلف ظاهر خواهد شد. بدین منظور تعداد کل کلاس‌های شناسایی به جای ۳۲ عدد به ۱۱۴ عدد افزایش خواهد یافت (۹). (به همه‌ی حروفی که کد شناسایی یکسانی را می‌گیرند، کلاس شناسایی یکسان می‌گویند مانند همه‌ی ب اول چسبان‌ها).
- از لحاظ شکل، اختلاف بسیاری از این حروف با هم بسیار جزیی و به طور مثال تنها در تعداد نقطه‌ها و یا محل قرار گرفتن آن ها می‌باشد مانند حروف ب، پ، ت و ث.
- فونت‌های نوشتاری بسیاری برای این زبان وجود دارد. مطابق مقاله‌ای از Kubba (۱۱) زبان عربی در حدود ۴۵۰ فونت مختلف برای کتاب‌ها و ۱۵۰ فونت برای روزنامه‌ها دارد که البته برای زبان فارسی نیز کمابیش وضع به همین شکل است.

بنابراین، الگوریتم‌هایی که برای مرحله ی تشخیص سایر زبان‌ها به کار گرفته می‌شود، در استفاده برای زبان‌های فارسی و عربی دچار مشکل خواهد شد؛ از این رو ارایه ی الگوریتم‌های قوی‌تر در این مرحله الزامی است.

روش‌ها

روش آنالیز اجزای اصلی

روش آنالیزاجزای اصلی را می‌توان برای استخراج اطلاعات اصلی موجود در دسته‌ای از داده‌ها استفاده نمود. در این روش، پایه‌هایی، که بهترین توزیع داده را نمایش می‌دهند، استخراج می‌شوند (۱۲). آنالیز اجزای اصلی برای ناهمبسته کردن داده‌ها مفید است و تبدیلی ایجاد می‌نماید که داده‌ها را در جهت‌های متعامد با بیشترین دامنه‌ی تغییرات تصویر می‌کند. همچنین آنالیز اجزای اصلی با تقلیل ابعاد از n به p ($p \le n$) برای کد کردن فشرده ی داده‌ها نیز قابل استفاده است. یعنی با استفاده از آنالیز اجزای اصلی بدون از دست رفتن قسمت عمده‌ی اطلاعات، اجزای وابسته به تغییرات کم، از داده حذف می‌شود (۱۳). روش آنالیز اجزای اصلی می‌تواند برای کاربردهای متعدد دیگری همچون طبقه‌بندی الگوها نیز استفاده شود (۱۵-۱۴).

برای اجرای این روش بایستی دو مرحله‌ی آموزش و آزمون را پیاده‌سازی نمود (۱۰). پروسه‌ی استفاده از آنالیز اجزای اصلی در آموزش یک الگوریتم تشخیص دهنده مطابق مراحل زیر خواهد بود:

۱) هر الگوی آموزشی P_i ($1 \le i \le m$) به تعدادی از کلاس طبقه‌بندی شده (کلاس شناسایی) اشاره دارد. (فرض می‌کنیم P_i یک بردار $n \times 1$ است.)

۲) بردار میانگین سری آموزشی را محاسبه و این میانگین را Ψ مینامیم:

$$\Psi = \frac{1}{m}\sum_{i=1}^{m} P_i \qquad (۱)$$

Ψ یک بردار با اندازه‌ی $n \times 1$ می‌باشد.

۳) از روی سری اولیه‌ی الگوهای آموزشی، ماتریس A تعریف می‌شود (فضای الگو). در هر یک از ستون‌ها یکی از m الگوی آموزشی قرار می‌گیرد که از آن بردار میانگین Ψ کسر شده است.

$$A = [P_1 - \Psi \quad P_2 - \Psi \quad \dots \quad P_m - \Psi] \qquad (۲)$$

ماتریس A به دست آمده اندازه‌ای برابر $n \times m$ خواهد داشت.

۴) بردارهای ویژه‌ی ماتریس کوواریانس

www.mui.ac.ir

الگوهای آموزشی را محاسبه می‌نماییم. این ماتریس با نام C به صورت زیر تعریف می‌شود:

$$C = AA^T \qquad (۳)$$

ماتریس کوواریانس با اندازه‌ی $n \times n$ می‌باشد که n اندازه‌ی الگوی آموزشی است. چون اندازه‌ی C بزرگ است، اغلب ماتریس L به فرم زیر در نظر گرفته می‌شود (۱۱):

$$L = A^TA \qquad (۴)$$

ماتریس L اندازه‌ای برابر $m \times m$ داردکه m تعداد الگوهای آموزشی است. قابل اثبات است که می‌توان بردارهای ویژه‌ی C را با استفاده از ترکیب خطی A و بردارهای ویژه ی L به دست آورد (۱۴). در صورتی که λ_i, v_iها به ترتیب بردارهای ویژه و مقادیر ویژه‌ی L باشند، خواهیم داشت:

$$\begin{aligned} &Lv_i = \lambda_i v_i \\ &(A^TA)v_i = \lambda_i v_i \\ &A(A^TA)v_i = A\lambda_i v_i \\ &(AA^T)Av_i = \lambda_i Av_i \\ &CAv_i = \lambda_i Av_i \end{aligned} \qquad (۵)$$

این معادله نشان می‌دهد که Av_i یک بردار ویژه از ماتریس $C = AA^T$ است. بنابراین V را ماتریسی قرار می‌دهیم که در ستون‌های خود شامل m بردار ویژه‌ی L است. ماتریس U –که بردارهای ویژه‌ی C را در ستون‌های خود دارد– را می‌توان این گونه نوشت:

$$U = AV \qquad (۶)$$

که اندازه‌ی ماتریس V برابر $m \times m$ و ماتریس A برابر $n \times m$ و در نتیجه ماتریس U برابر $n \times m$ خواهد بود.

۵) برای ایجاد ماتریس اجزای اصلی تعداد p عدد از بردارهای ویژه ی C را انتخاب می‌کنیم که در آن $p \leq m$. بردارهای ویژه ی انتخاب شده، ستون‌های ماتریس U هستند که به بزرگ‌ترین مقادیر ویژه‌ی پیدا شده در ماتریس در بردارهای ویژه‌ی محاسبه شده از ماتریس کوواریانس مربوط می‌باشند. این انتخاب بر این اساس انجام شده است که ماتریس U تنها شامل p بردار ویژه باشد.

$$U = [U_1 \quad U_2 \quad \dots \quad U_p] \qquad (۷)$$

که ماتریس U اندازه‌ای برابر $p \times n$ خواهد داشت.

۶) هر کدام از الگوهای آموزشی Pi به دست آمده را در فضای بردار ویژه ی انتخاب شده تصویر می‌کنیم تا به یک Ω_i کد شده‌ی جدید از الگو برسیم.

$$\Omega_i = U^T(P_i - \Psi) \qquad (۸)$$

این بردار کد شده Ω_i به دست آمده اندازه‌ی $p \times ۱$ خواهد داشت.

از میان همه ی مراحل برای الگوریتم آموزشی PCA، مرحله ی (۵) بسیار مهم است، چون اجازه می‌دهد که اندازه‌ی الگوهای آموزشی از $n \times ۱$ به $p \times ۱$ تقلیل یابد. به این مرحله می‌توان به عنوان مرحله‌ی فشرده‌سازی نگاه کرد.

پروسه‌ی استفاده از روش آنالیز اجزای اصلی در آزمون یک الگوریتم آموزش داده شده نیز مطابق مراحل زیر خواهد بود:

فرض کنید P_{test} یک الگو با اندازه ی $n \times ۱$ در مرحله‌ی شناسایی و طبقه‌بندی باشد. مراحل زیر انجام می‌شود:

۱) بردار میانگین Ψ از P_{test} کم می‌شود. آن‌گاه بردار به دست آمده بر روی فضای بردار ویژه‌ی U تصویر می‌شود تا یک Ω_{test} کد شده ی جدید از الگو به دست آید:

$$\Omega_{test} = U^T(P_{test} - \Psi) \qquad (۹)$$

Ω_{test} اندازه‌ای برابر $p \times ۱$ دارد.

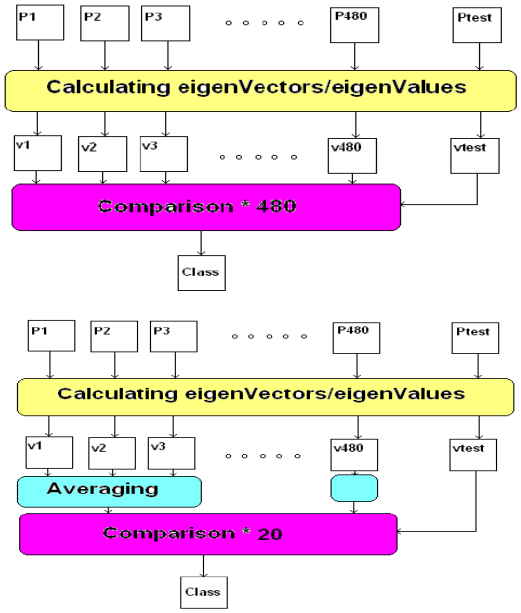
www.mui.ac.ir

۲) الگوهای آزمایشی کد شده Ω_i را آنالیز می کنند تا الگوی Ω به فرم زیر حاصل گردد:

$$\Omega = \min \| \Omega_{test} - \Omega_i \| \quad (۱۰)$$

همانگونه که در معادله‌ی (۱۰) آمده است، الگوی P_{test} به دست آمده به گونه‌ای طبقه‌بندی (شناسایی) می‌شود که با کلاس مشابهی از X_i متناظر باشد.

در ادامه می‌توان برای بالاتر رفتن سرعت الگوریتم به جای آنکه به ازای تک تک داده‌های آموزشی مقادیر ویژه ی آن‌ها را در عمل مقایسه به حساب آورد، در فرایند آموزش با دسته‌بندی داده‌های آموزشی و پس از طی مراحل پردازش از بردارهای ویژه ی این دسته میانگین گرفت ومیانگین‌ها را ملاک مقایسه در پروسه‌ی آزمایش قرار داد (شکل ۱).



شکل ۱. مقایسه‌ی بین ساختار PCA (شکل بالا) و PCA ترکیبی (شکل پایین)

روش الگوریتم آماری

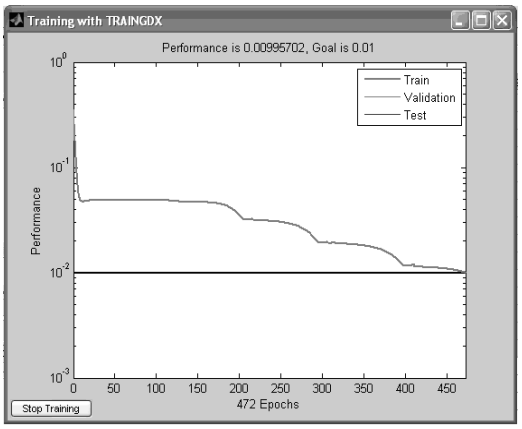
جهت اجرای الگوریتم آماری از مقاله‌ی هاشمی و همکاران (۸) استفاده شده است. بر اساس نتایج این مقاله، روش اعمالی در مقایسه با سایر روش‌های قبلی کارآیی بسیار خوبی داشته است. در این پژوهش در مرحله‌ی شناسایی حروف، از روش آماری استفاده شده است که بنا به اظهار مؤلف نتیجه‌ی این روش با شش روش دیگر از جمله روش‌های نزدیک‌ترین همسایگی نیز مقایسه شده و نتایج قابل توجهی کسب نموده است. در مطالعه‌ی هاشمی و همکاران برای انجام مرحله‌ی تشخیص حروف فارسی ۱۴ روش مختلف آماری ارایه شده است که از این میان بنا به نتایج اخذ شده، جدیدترین و سریع‌ترین روش، روش استفاده از شبکه‌ی عصبی بوده است.

دراین مقاله، یک شبکه‌ی عصبی چند لایه‌ی جلوسو با یک لایه‌ی میانی استفاده شده است. به علت تبدیل تمام نمونه‌های ورودی به اندازه‌ی ۲۰ × ۲۰، تعداد نورون‌های ورودی ۴۰۰ انتخاب شده است. تعداد نورون‌های خروجی ۲۰ عدد می‌باشد که برابر تعداد کلاس‌های شناسایی است. تعداد نورون‌های لایه‌ی پنهان نیز بنا به پیشنهاد مقاله (۸) برابر ۱۸۰ عدد انتخاب می‌گردد. البته پس از اجرای برنامه با این تعداد نورون لایه‌ی میانی، به دلیل تعداد خاص ورودی‌ها و پس از چند بار آزمایش و خطا با تعداد مختلف نورون‌های لایه‌ی میانی، این نتیجه حاصل می‌شود که برای داشتن بالاترین سرعت و دقت در این الگوریتم بهتر است تعداد نورون‌های لایه‌ی پنهان برابر ۴۸۰ انتخاب شود.

در این الگوریتم ابتدا نرم‌افزار به وسیله‌ی داده‌های آموزشی آموزش داده می‌شود و پس از تنظیم خودکار ضرایب، داده‌های آزمونی به نرم‌افزار آموزش داده شده اعمال می‌گردد. پس از اتمام مراحل پردازشی، کمترین

www.mui.ac.ir

مقدار خطا، ملاک تعلق نمونه‌ی آزمونی به یک کلاس شناسایی خاص خواهد بود. شکل ۲ نمونه‌ای از نمودار آموزش یکی از شبکه‌های عصبی مربوط به یک کلاس شناسایی نمونه در الگوریتم آماری را نشان می‌دهد.



شکل ۲. نمودار روند آموزش یکی از شبکه‌های عصبی مربوط به یک کلاس شناسایی نمونه در الگوریتم آماری

روش زرنیک سریع

الگوریتم دیگری که در این مقاله مورد استفاده قرار گرفت، الگوریتم ارایه شده توسط برومندنیا و شنبه‌زاده (۹) موسوم به زرنیک سریع می باشد. در این مقاله با استفاده از پنج روش ویژگی‌های تصاویر حروف فارسی استخراج می‌شود و سپس با استفاده از شبکه‌ی عصبی بر اساس ویژگی‌های استخراج شده، تشخیص حروف صورت می‌پذیرد. این پنج روش شامل روش‌های ZWM (Zernike Wavelet Moment)، ZM (Zernike Moment)، WM (Wavelet Moment)، OFW (Orthogonal Fourrier Mellin Moments) و RCF (Radial Coding Features) می‌باشد که در این مقاله صحت به دست آمده برای داده‌ها با توجه به نتایج عملی در مورد روش ZWM که ترکیبی از ویژگی‌های تبدیل ویولت و تبدیل زرنیک تصاویر می‌باشد، بالاترین میزان گزارش شده است. این الگوریتم با استفاده از ویژگی‌های استخراج شده از تصاویر حروف فارسی و با استفاده از ZWM و همچنین یک شبکه‌ی عصبی جلوسوی چند لایه با یک لایه ی میانی پنهانی شامل ۴۵ نورون اجرا می‌شود. ساختار برنامه ی زرنیک سریع شبیه ساختار الگوریتم آماری است؛ با این تفاوت که ابتدا از هر تصویر مشخصه‌های زرنیک و ویولت استخراج می‌شوند و سپس به عنوان ورودی به شبکه‌های عصبی داده می‌شوند. ادامه‌ی کار مانند روش آماری خواهد بود.

در سیستم تشخیص حروف، نمونه‌های مورد نظر اغلب با مجموعه‌ای از ویژگی‌های عددی با هدف کاهش اطلاعات تکراری و کم کردن حجم اطلاعات نمایش داده می‌شوند. با استفاده از روش زرنیک و ممان ویولت تغییر اندازه و چرخش در تشخیص حروف بی‌تأثیر خواهد بود (۱۸-۱۶). برای به دست آوردن ممان مستقل از چرخش برای تصویر $f(r,\theta)$ (در فضای مختصات قطبی) از فرمول زیر استفاده می‌شود:

$$F_{pq}=\iint f(r.\theta)g_{pq}(r)e^{jq\theta}rdrd\theta, \qquad (۱۱)$$

که در آن F_{pq} ممان درجه‌ی pq‌ام، $g_{pq}(r)$ تابعی از متغیر شعاعی r و p و q پارامترهای صحیحی هستند.

کرنل ممان‌های زرنیک مجموعه‌ای از چند جمله‌ای‌های زرنیک متعامد می‌باشد که در یک بردار مختصات قطبی و در فضای یک دایره ی واحد تعریف می‌شود. ممان زرنیک دو بعدی از درجه ی p با تکرار q برای تصویر $f(r,\theta)$ به صورت زیر تعریف می‌شود:

$$Z_{pq}=\frac{P+1}{\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{1}V^{*}_{pq}(r,\theta)f(r,\theta)rdrd\theta, \qquad (۱۲)$$

$$|r|\le 1$$

که چند جمله‌ای زرنیک از درجه‌ی p با تکرار q، $V_{pq}(r,\theta)$ به صورت زیر تعریف می‌شود:

www.mui.ac.ir

$$V_{pq}(r,\theta) = R_{pq}(r)e^{jq\theta} \quad (۱۳)$$

و $R_{pq}(r)$ چند جمله‌ای دایروی با مقدار حقیقی خواهد بود.

ممان ویولت نیز بدین گونه تعریف می‌شود:

$$W_{mnq} = \int_{\theta=0}^{2\pi} \int_{r=0}^{1} f(r,\theta)\psi_{mn}(r)rdrd\theta \quad |r| \le 1 \quad (۱۴)$$

با انتقال تصویر به فضای دایره‌ی واحد، مشخصه‌های زرنیک و ممان ویولت تصویر به ترتیب از $\|Z_{pq}\|$ و $\|W_{mnq}\|$ به ازای مقادیر مختلف از p، q، m و n محاسبه می‌گردد. پس از محاسبه‌ی این مشخصه‌ها و ترکیب آن‌ها، مشخصه‌های ترکیبی FZWM تصویر استخراج شده و حال، این مشخصه‌ها به جای خود تصویر به ورودی شبکه‌های عصبی اعمال می‌گردد. سایر مراحل همانند روش الگوریتم آماری خواهد بود.

یافته‌ها

جهت مقایسه‌ی کارآیی این روش‌ها با هم از مجموعه‌ی متون تایپی و چاپی فارسی نمونه‌هایی با خصوصیات زیر تهیه شد:

- از روی متون مختلف فارسی تعداد ۳۵۰ تصویر تهیه شد و تمامی مراحل آماده‌سازی تصاویر شرح داده شده در مقدمه، به صورت دستی انجام گرفت. داده‌های نمونه‌ای به دست آمده به صورت تصاویر پردازش شده، از حروف مجزا شده بودند.
- نمونه‌ها از متون مختلف و تحت شرایط مختلف تهیه شده بودند تا وجود نویز و سایر موارد به صورت واقعی‌تر در آن‌ها مد نظر قرار گرفته باشد. در کل، تعداد ۷۲۰ نمونه جهت آموزش الگوریتم‌ها و ۴ دسته ی ۴۸۰ تایی مختلف به ترتیب آسان به دشوار نیز جهت آزمون این الگوریتم‌ها استفاده شد. دشوار به این معنا که نمونه‌ها دارای شرایط نویزی و تصویربرداری بدتری بودند و تشخیص کلاس شناسایی آن‌ها مشکل‌تر بود. این نمونه‌ها شامل ۴ فونت مختلف نازنین، تاهوما، تیتر و زر بودند که از ۴ اندازه‌ی مختلف ۱۲، ۱۴ و ۲۰ انتخاب شده بودند.
- جهت مقایسه‌ی هر سه الگوریتم، ابتدا داده‌های آموزشی به الگوریتم‌ها اعمال شد. پس از آموزش آن‌ها چهار مجموعه ی داده‌های آزمونی جداگانه به هرکدام از آنها اعمال گردید و نتیجه‌ی اعمال این داده‌های آزمونی جداگانه ثبت شد.
- برای محاسبه‌ی دقت الگوریتم‌ها پس از دریافت جواب از برنامه، پاسخ‌های به دست آمده از نظر صحت بررسی شدند و تعداد پاسخ‌های صحیح شمرده می‌شد و بدینوسیله درصد صحت پاسخ‌ها محاسبه می‌گردید.
- جهت محاسبه ی سرعت الگوریتم‌ها نیز زمان طی شده از ابتدای اعمال داده‌های آزمونی (مجموعه‌های ۴۸۰ تایی) به برنامه تا زمان دریافت کامل جواب اندازه‌گیری می‌گردید.
- در انتها، با پنج بار اعمال کردن هر کدام از مجموعه‌های آزمونی به الگوریتم‌ها و گرفتن پاسخ، از تمامی درصدهای صحت و زمان‌های پاسخ میان‌گیری شد.

نتیجه‌گیری

با توجه به جدول نتایج می‌توان استنباط کرد که الگوریتم آنالیز اجزای اصلی ترکیبی توانسته است به میزان ۱/۷۷۵ درصد دقت و ۷/۵ بار سرعت تشخیص حروف چاپی زبان فارسی را نسبت به روش آماری و همچنین به میزان ۲/۲ درصد دقت و ۵/۱۲ بار سرعت را نسبت به روش زرنیک سریع بهبود بخشد. مقایسه‌ی

www.mui.ac.ir

جدول ۱: مقایسه‌ی نتایج نهایی سه الگوریتم

بررسی دقت (درصد)	داده‌های آزمونی ۱	داده‌های آزمونی ۲	داده‌های آزمونی ۳	داده‌های آزمونی ۴
الگوریتم آماری	۹۵/۲	۹۵/۵	۹۶	۸۷/۱
الگوریتم زرنیک سریع	۹۵/۲	۹۴/۵	۹۵/۵	۸۶/۹
الگوریتم PCA	۱۰۰	۹۹/۷۹	۹۶/۲۵	۸۷/۲۹
الگوریتم PCA ترکیبی	۹۹/۵	۹۸/۹	۹۵/۴	۸۷/۱
بررسی سرعت (ثانیه)	**داده‌های آزمونی ۱**	**داده‌های آزمونی ۲**	**داده‌های آزمونی ۳**	**داده‌های آزمونی ۴**
الگوریتم آماری	۳۷	۳۶	۳۶	۳۷
الگوریتم زرنیک سریع	۲۵	۲۴/۵	۲۴/۹	۲۵
الگوریتم PCA	۳۷/۲۴	۳۶/۹۵	۳۶/۰۵	۳۴/۹۴
الگوریتم PCA ترکیبی	۵	۴/۸	۴/۸	۴/۸

نتایج نهایی اعمال داده‌ها به الگوریتم‌های مطرح شده در مقاله و مقایسه‌ی دقت و زمان آن‌ها در جدول ۱ قابل مشاهده است. چنانچه از نتایج مشخص است، الگوریتم آنالیز اجزای اصلی (PCA) ترکیبی بهترین نتایج از لحاظ سرعت و دقت را به همراه دارد.

استفاده از دیگر روش‌های کاهش بعد همچون آنالیز مؤلفه‌های مستقل یا تجزیه‌ی نقاط تکین می‌تواند منجر به بهبود نتایج گردد. همچنین استفاده از پیش پردازش‌های اتوماتیک نیز پیش از شروع عملیات اصلی در تسهیل الگوریتم و بهبود نتایج مؤثر خواهد بود.

تشکر و قدردانی

این مقاله برگرفته از طرح تحقیقاتی شماره‌ی ۳۸۸۲۳۷ مصوب معاونت پژوهشی دانشگاه علوم پزشکی اصفهان می‌باشد. بدین‌وسیله از معاونت پژوهشی دانشگاه علوم پزشکی اصفهان قدردانی به عمل می‌آید.

References

1. Parslow A. An optical real-world text to speech reader. International Congress Series 2005; 1282: 1056-60.
2. Fazilat M. Applied technology for blinds. Electronic Edition of Documents and Information of Iran 2007; 1(7): 1-7.
3. Amin A. Machine recognition of handwritten Arabic words by IRAC II. Proceedings of the 6th International Conference in Pattern Recognition; 1982 Oct; Munich, Germany. p. 729-31.
4. Amin A, Masini G, Haton JP. Recognition of Arabic words and sentences. Proceedings of the 7th International Conference in Pattern Recognition; 1984; Montreal, Canada. p. 1055-7.
5. Badi K, Shimura M. Machine recognition of Arabic cursive scripts. Trans Inst Electron Commun Eng 1982; E65(2): 107-14.
6. Amin A, Masini G. Machine recognition of multifont printed Arabic texts. Proceedings of the 8th International Conference in Pattern Recognition; 1986 Oct 27-31; Paris, France. p. 392-95.
7. Hachour O. The combination of fuzzy logic and expert system for Arabic character recognition. Proceedings of the 3rd International IEEE Conference on Intelligent Systems; 2006; p. 189-91.
8. Hashemi MR, Fatemi O, Safavi R. Persian cursive script recognition. Proceedings of the 3rd International Conference on pattern recognition; 1995 Aug 14-16; Montreal, Canada. p. 869-73.
9. Broumandnia A, Shanbehzadeh J. Fast zernike wavelet moments for Farsi character recognition. Image and Vision Computing 2007; 25(5): 717-26.
10. Curado LAF, Bauchspiess A, CuradoVBF. Using PCA to recognize characters in vehicle license plates. Proceedings of the World Congress in Computer Science; 2006; Las Vegas, USA. p. 264-68.
11. Kubba GA. The impact of computers on Arabic writing, character processing and teaching. Proceedings of I.F.I.P. Congress on Information Processing and Education (Area 9); 1980 Oct; Tokyo, Japan. p. 961-65.

www.mui.ac.ir

12. Shlens J. A Tutorial on principal component analysis: derivation, discussion and singular value decomposition. [Online] 2006 Available from: URL: http://www.snl.salk.edu/~shlens/notes.html.
13. Haykin S. Neural network: a comprehensive foundation. 2nd ed. Englewood Cliffs, NJ: Prentice Hall; 1999.
14. Turk M, Pentland A. Eigenfaces for recognition. Journal of Cognitive Neuroscience 1991; 3(1): 71-86.
15. Silva P, Rosa A. Face recognition based on symmetryzation. Centre for Secure Information Technologies. [Online] 2002. Available from: URL: http//www.wenku.baidu.com/view/8cf86cc558f5f61fb7366610.html.
16. Chong CW, Raveendran P, Mukundan R. Translation invariants of Zernike moments. The Journal of the Pattern Recognition Society 2003; 36: 1765-73.
17. Chim YC KAIY. Character recognition using statistical moments. Image and Vision Computing 1999; 17: 299-307.
18. Mallat S. A Wavelet Tour of Signal Processing. 2nd ed. New York, NY: Academic Press; 1999.

www.mui.ac.ir

Journal of Isfahan Medical School Original Article

Vol 29, No 174, Special Issue (Biomedical Engineering) 2012 Received: 6.11.2011 Accepted: 6.2.2012

Using Principal Component Analysis to Recognize Farsi Alphabetic Characters in Printed Scripts

Shahrouz Gashmard MSc[1], Alireza Mehri Dehnavi PhD[2], Hossein Rabbani PhD[3]

Abstract

Background: From the early stages of script writing and reading, there was a big desire for researchers to make an aid device for the blind to read scripts without help of other. This paper introduced a new approach for recognition of Farsi scripts using principal component analysis (PCA).

Methods: Materials used for this project were selected from ordinary books, magazines, newspapers, and printed documents. Character samples were selected from four fonts in different positions and three sizes resulted in total number of 20 recognition classes. The five methods used in character recognition were statistical method, Fast Zernike Wavelet Moments (FZWM) method, PCA, PCA with sample averaging, and PCA with eigenvectors averaging.

Findings: The accuracy and speed of PCA in recognition of Farsi characters were respectively 1.775% and 7.5 times better than the statistical method. Likewise, it was 2.2% more accurate and 5.12 times faster than FZWM method.

Conclusion: Using PCA with combinational averaging in samples and eigenvectors can be a novel method for recognition of Farsi characters.

Keywords: Farsi character recognition, Image processing, Principal component analysis, Combined algorithm

[1] Department of Medical Physics and Engineering, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
[2] Associate Professor, Department of Medical Physics and Engineering, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
[3] Assistant Professor, Department of Medical Physics and Engineering, School of Medicine AND Medical Signal and Image Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
Corresponding Author: Alireza Mehridehnavi PhD, Email: mehri@med.mui.ac.ir