کاربرد تطابق شکل در بازشناسی ارقام دستنویس فارسی

علیرضا درویش[۱]، احسان اله کبیر[۲*]، حسین خسروی[۳]

۱- دانش‌آموخته کارشناسی ارشد مهندسی الکترونیک ، دانشکدهٔ فنی و مهندسی، دانشگاه تربیت مدرس

۲- دانشیار بخش مهندسی برق، دانشکدهٔ فنی و مهندسی، دانشگاه تربیت مدرس

۳- دانش‌آموخته کارشناسی ارشد مهندسی الکترونیک، دانشکدهٔ فنی و مهندسی، دانشگاه تربیت مدرس

*** تهران، صندوق پستی ۱۴۳- ۱۴۱۱۵**

kabir@modares.ac.ir

چکیده- در این تحقیق، از نوعی الگوریتم تطابق شکل برای بازشناسی ارقام دستنویس فارسی استفاده شده است. برای هر نقطه نمونه برداری شده بر روی کانتور شکل، توصیفگری براساس توزیع مکانی نقاط دیگر کانتور به‌دست می‌آید. برای تعیین میزان شباهت دو شکل، ابتدا بر اساس این توصیفگرها تناظری یک به یک بین نقاط نمونه برداری شده روی کانتور شکل اول با نقاط روی کانتور شکل دوم به‌دست می‌آید. جمع فواصل بین نقاط متناظر در دو شکل، معیاری برای عدم شباهت آنها است. سپس تبدیلی تعریف می‌شود که نقاط کانتور شکل اول را بر روی نقاط متناظر در کانتور شکل دوم قرار دهد. میزان پیچیدگی این تبدیل معیاری دیگر برای عدم شباهت دو شکل است.

با بهینه سازی پارامترهای مختلف این الگوریتم، از آن برای بازشناسی ارقام دستنویس فارسی استفاده شده است. از روش طبقه‌بندی نزدیکترین فاصله از نماینده یا نماینده‌های هر کلاس استفاده شده است. در آزمایشی برروی مجموعه ای شامل ۱۲۸۸رقم - که افراد مختلف نوشته‌اند - صحت بازشناسی ۸۹/۹٪ بوده است. این آزمایش بدون هیچگونه پس پردازشی انجام شده است.

کلید واژگان: تطابق شکل، کانتور، توصیفگر محلی، تبدیل هندسی، بازشناسی ارقام دستنویس.

۱- مقدمه

بازشناسی حروف و ارقام دستنویس همواره یکی از موضوعات مورد علاقه برای تحقیق بوده است [۱-۴]. در زمینه بازشناسی حروف و ارقام دستنویس عربی و فارسی نیز، کارهای زیادی صورت گرفته است [۵]. در یک تحقیق، از ویژگیهای گشتاوری و طبقه‌بندی کننده بیز برای بازشناسی حروف دستنویس فارسی استفاده شده است [۶]. در تحقیق دیگری، از ویژگیهای مکانهای مشخصه، طبقه‌بندی کننده بیز و زنجیر مارکف برای بازشناسی ارقام استفاده شده است [۷]. از روش 3-tuple نیز برای استخراج ویژگی از ارقام دستنویس استفاده شده است[۸]. مسروری از DTW برای بازشناسی ارقام استفاده کرده است [۹]. از منطق فازی نیز برای تقسیم حروف به چند دسته مختلف استفاده شده است [۱۰]. در تحقیق دیگری، بازشناسی ارقام دستنویس فارسی به کمک روشهای فازی انجام شده است [۱۱]. کار ما مبتنی بر توصیف شکل ارقام دستنویس است، به این صورت که از الگوریتم تطابق الگو [۱۲] برای توصیف و مقایسه دو رقم دستنویس استفاده و سعی کرده ایم با بهینه‌سازی این الگوریتم، آن را برای بازشناسی ارقام دستنویس فارسی اصلاح کنیم. در این روش ابتدا باید توصیفی از شکل مورد نظر داشته باشیم و در مرحله بعد معیاری را برای مقایسه توصیفگرها تعریف کنیم.

روشی که ما برای توصیف شکل استفاده می‌کنیم، مبتنی بر کانتور شکل است. در سالهای اخیر کارهای زیادی در زمینه توصیف شکل بر اساس کانتور انجام شده است. چنگ و یان [۱۳] ویژگیهایی را از کانتور ارقام استخراج کرده‌اند. این ویژگیها عبارتند از: توصیفگر فوریه کانتور، میانگین و واریانس فاصلهٔ پیکسلهای روی کانتور تا مرکز ثقل آن و همچنین طول و مساحت نرمالیزه شده کانتور. ارقام، بسته به اینکه چند کانتور دارند به چند دسته تقسیم می‌شوند و بر اساس این ویژگیها شناسایی می‌شوند. همچنین در روش دیگری ،چاکراوارتی و کمپلا بر اساس نقاط بحرانی، توصیفی ریاضی از کانتور حروف و ارقام ارائه دادند [۱۴]. نقطه بحرانی، نقطه‌ای است که مشتق اول تابع کانتور در آن صفر است. در این روش کانتور هر شکل بر اساس نقاط بحرانی به‌وسیله گرافی نمایش داده می‌شود. روش ما که بر کانتور شکل مبتنی است سه مرحله دارد [۱۲]:

۱. حل مسأله تناظر بین نقاط کانتور شکل اول با نقاط کانتور شکل دوم.
۲. استفاده از این تناظر برای تعریف تبدیل هندسی که دوشکل را بر روی هم قرار دهد.
۳. محاسبه فاصله بین دو شکل به صورت مجموع خطاهای تطابق نقاط متناظر دو شکل و جمله‌ای که میزان پیچیدگی این تبدیل هندسی را نشان می‌دهد.

در ادامه در بخشهای دوم و سوم به معرفی الگوریتم تطابق شکل می‌پردازیم و در بخش چهارم چگونگی استفاده از این الگوریتم را برای بازشناسی ارقام دستنویس فارسی توضیح می‌دهیم. در بخش پنجم این روش را با روشهای دیگر مقایسه می‌کنیم. بخش ششم به نتیجه گیری و ارائه پیشنهاد هایی برای بهبود بازشناسی اختصاص دارد.

۲- توصیف شکل ارقام

برای توصیف شکل رقم، ابتدا کانتور آن به‌دست آمده و سپس از تعدادی نقطه بر روی کانتور نمونه‌برداری می‌شود. برای هر نقطه، براساس فاصله نسبی با سایر نقاط روی کانتور، توصیفگری محلی تعریف می‌شود. برای مقایسهٔ رقم‌ها، ابتدا بر روی کانتور هر یک از آنها، نقاط لازم نمونه‌برداری می‌شود. سپس براساس معیار فاصله‌ای که تعریف می‌شود، تناظری یک به یک بین نقاط نمونه‌برداری شده دو رقم به‌دست می‌آید. در نتیجه فاصله بین دو رقم، به صورت جمع فاصله بین تک تک نقاط متناظر آنها قابل محاسبه خواهد بود. کانتور شکل ارقام به کمک لبه‌یاب کنی[1] مشخص می‌شود. تعداد نقاط نمونه‌برداری شده بر روی کانتور، بسته به دقت و کاربرد متفاوت است. البته مشخص است که هر چه تعداد این نقاط بیشتر باشد، توصیف دقیقتری از شکل به‌دست می‌آید. اما این نکته را هم باید درنظر داشت که با افزایش تعداد نقاط، حجم کار برنامه و در نتیجه زمان اجرای آن افزایش می‌یابد. بنابراین در این مورد باید انتخاب مناسبی داشته باشیم.

۲-۱- توصیفگر محلی برای نقاط شکل

اگر مجموعه بردارهایی را که از یک نقطه بر روی کانتور به تمام نقاط دیگر کانتور ترسیم می‌شود در نظر بگیریم، بردارهایی را داریم که توصیف تمام شکل را نسبت به نقطه مرجع نشان می‌دهند. براساس این بردارها، توصیفی را برای تک تک نقاط نمونه‌برداری شده می‌توان به‌دست آورد [۱۲].

برای تعریف این توصیفگر شکل ۱ را درنظر بگیرید. برای هر نقطه p_i بر روی شکل اول می‌توان هیستوگرام h_i را براساس فاصله و زاویه نسبی پاره خط واصل آن نقطه با $n-1$ نقطهٔ دیگر روی کانتور تعریف کرد.

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in bin(k)\} \qquad (۱)$$

این هیستوگرام به عنوان توصیفگر اطلاعات جانبی نقطه p_i تعریف می‌شود. همانگونه که در شکل ۱

1. Canny Edge Detector

مشخص است، برای محاسبه این هیستوگرام از مختصات لگاریتمی قطبی استفاده می‌شود زیرا هدف آن است که توصیفگر محاسبه شده در رابطه (۱) نسبت به نقاطی که در نزدیک آن قرار دادند، حساستر باشد.

در این مختصات، عناصر از نظر زاویه‌ای به ۱۲ قسمت مساوی ۳۰ درجه تقسیم شده و از نظر شعاعی نیز، درجه بندی شعاع عناصر به صورت لگاریتمی بوده و از ۰/۱۲۵ تا ۲ تقسیم شده است. به کمک توصیفگرهای هر نقطه، تناظری یک به یک بین تمام نقاط روی کانتور شکل اول با تمام نقاط روی کانتور شکل دوم به‌دست می‌آوریم. برای محاسبه این تناظر باید از نوعی معیار فاصله بین این توصیفگرها استفاده کنیم، نقطهٔ p_i را روی شکل اول و نقطه q_j را بر روی شکل دوم در نظر بگیرید. اگر (p_i, q_j) فاصله این دو نقطه باشد آنگاه :

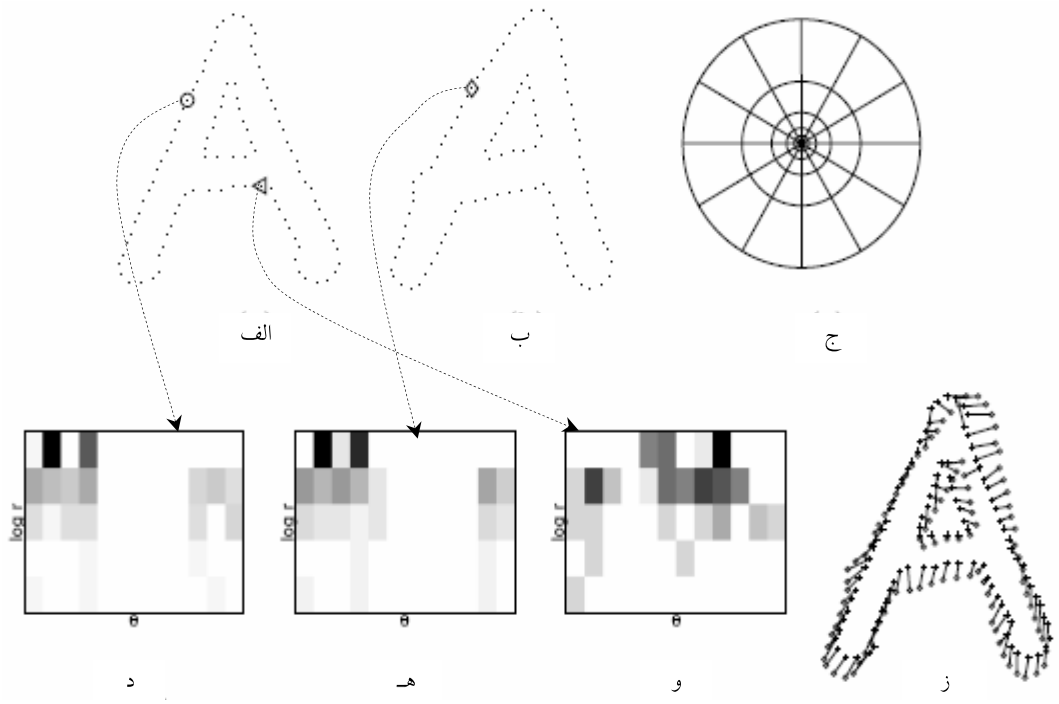
$$c_{ij} = \left(p_i, q_j\right) = \frac{1}{2}\sum_{k=1}^{K}\frac{\left[h_i(k) - h_j(k)\right]^2}{h_i(k) + h_j(k)} \qquad (۲)$$

که در این رابطه $h_i(k)$ و $h_j(k)$، معرف هیستوگرامهای نرمالیزه شدهٔ k عنصری مربوط به p_i و q_j هستند. پس از تعریف رابطه (۲) فاصله c_{ij} بین تمام نقاط روی شکل اول با تمام نقاط روی شکل دوم را به صورت ماتریس فاصلهٔ (۳) در اختیار خواهیم داشت:

$$Dis = \begin{bmatrix} c_{11} & \cdot & \cdot & c_{n1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ c_{n1} & \cdot & \cdot & c_{nn} \end{bmatrix} \qquad (۳)$$

اکنون هدف آن است که برای نقطه p_i بر روی شکل اول، نقطه $q_{\pi(i)}$ را بر روی شکل دوم به گونه‌ای به‌دست آوریم که فاصله کل بین دو شکل حداقل شود.

$$C_{context} = H(\pi) = \sum_i c(p_i, q_{\pi(i)}) \qquad (۴)$$



شکل ۱ الف وب) نقاط نمونه برداری شده از لبه دو شکل ج) دیاگرام مختصا ت لگاریتمی- قطبی(در این دیاگرام از ۵ عنصر شعاعی و ۱۲ عنصر زاویه‌ای استفاده شده است.) د، هـ و) هیستوگرام بدست آمده برای سه نقطه نشان داده شده در شکلهای الف و ب. ز) نقاط متناظر دو شکل الف و ب [۱۲].

در واقع π تابعی یک به یک است که به هر نقطه i بر روی شکل اول، نقطه $\pi(i)$ را بر روی شکل دوم به عنوان نزدیکترین نقطه از نظر تشابه نسبت می‌دهد.

این مسأله به عنوان یک مسأله ریاضی در نظر گرفته می‌شود که ورودی آن ماتریس Dis است و خروجی آن، نقاط $\pi(i)$ بر روی شکل دوم برای تک تک نقاط i بر روی شکل اول است.

برای حل این مسأله می‌توان از روشهای مجاری[1] [۱۵] یا یونکر [۱۶] استفاده کرد که ما از هر دو روش استفاده کردیم. نتایج مربوط در بخش ٦ آورده شده است. پس از اجرای این الگوریتم، تناظری یک به یک بین نقاط شکل اول و شکل دوم به‌دست می‌آید و رابطه(٤) مقدار حداقل خود را می‌گیرد. اکنون از این رابطه می‌توان به عنوان معیاری برای تعیین فاصله دو شکل استفاده کرد؛ یعنی هر چه این مقدار بیشتر باشد، شباهت دو شکل کمتر است. همانگونه که اشاره شد، رابطه(٤) یکی از دو معیاری است که در تعیین معیار کلی فاصله بین دو شکل به کار می‌رود. یک معیار دیگر، از محاسبه میزان پیچیدگی انجام تبدیل هندسیی که این دو شکل را بر روی هم قرار می‌دهد قابل محاسبه است و همانطور که بعداً اشاره می‌شود، مجموع وزنی این دو فاصله را می‌توان به عنوان معیار فاصله کلی بین دو شکل در نظر گرفت [۱۲].

۳- تبدیل اسپلاین صفحه نازک[2]

هنگامی‌که مجموعه‌ای از نقاط متناظر را بر روی کانتور دو رقم در اختیار داریم، می‌توانیم تبدیلی را طراحی کنیم که بر اساس آن، این نقاط را بر روی هم قرار دهیم. با محاسبه میزان پیچیدگی این تبدیل، معیاری برای میزان اختلاف دو رقم به‌دست می‌آید. تبدیلی که در این قسمت استفاده کرده‌ایم، جزو خانواده تبدیلات اسپلاین صفحه نازک است که برای نمایش تبدیل در مختصات قابل انعطاف مناسب است [۱۲]. برای تعریف این تبدیل به صورت زیر عمل می‌کنیم.

فرض کنید بردار V شامل مقادیر خروجی تابعی است که بر مجموعه نقاط بردار P شامل نقاط $P_i = (x_i, y_i)$ اعمال می‌شود. مجموعه نقاط بردار V را با (x'_i, y'_i) نشان می‌دهیم. تبدیل f(x,y) مجموعه نقاط (x_i, y_i) را به صورتی بر مجموعه نقاط (x'_i, y'_i) می‌نشاند که انرژی لازم برای انجام این تبدیل حداقل شود.

$$f(x,y) = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U\left(\left\|(x,y) - (x_i, y_i)\right\|\right) \quad (۵)$$

که تابع کرنل $u(r)$ به صورت زیر است:

$$U(r) = r^2 \log r^2 \quad (۶)$$

برای اینکه تابع $f(x,y)$ مشتقات مرتبهٔ دوم پیوسته داشته باشد باید داشته باشیم:

$$\sum_{i=1}^{n} w_i = 0 \text{ and } \sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i = 0$$

براساس تبدیل (۵) می‌توان یک سیستم خطی را به صورت زیر برای ضرایب TPS[1] در نظر گرفت:

$$\begin{pmatrix} k & p \\ p^T & o \end{pmatrix} \begin{pmatrix} w \\ a \end{pmatrix} = \begin{pmatrix} v \\ o \end{pmatrix} \quad (۷)$$

که در آن ماتریسهای K و P چنین است:

$$K_{ij} = U(\|(x_i, y_i) - (x_j, y_j)\|) \quad (۸)$$

ith row of $P = (1, x_i, y_i)$

و 0 ماتریس صفر ۳×۳ است.

a, w بردارهای ضرایب در رابطه (۵) هستند.

$$a = [a_1, a_x, a_y]^T \quad (۹)$$

$$w = [w_1, w_2, \ldots, w_n]^T, v = [v_1, v_2, \ldots, v_n]^T \quad (۱۰)$$

اکنون همانطور که بوکستین [۱۷] نشان داده اگر v_i ها را – که نقاط متناظر بر روی کانتور شکل دوم هستند – به صورت زیر در نظر بگیریم:

1. Hungarian
2. Thin Plate Spline

$$v_1=\left(x'_1,y'_1\right),v_2=\left(x'_2,y'_2\right)\ldots,v_n=\left(x'_n,y'_n\right) \quad (۱۱)$$

و بردار Y را به صورت زیر تعریف کنیم:

$$Y=\left(V\mid 000\right)^T \quad (۱۲)$$

که Y بردار ستونی با طول n+3 است. آنگاه مقادیر $W=(w_1,\ldots,w_n)^T$ و $A=(a_1,a_x,a_y)^T$ از رابطه زیر به‌دست می‌آیند:

$$L^{-1}Y=(w_1,a_1,a_x,a_y)^T \quad (۱۳)$$

که در آن برای L رابطه زیر را داریم:

$$L=\begin{bmatrix}K & P\\ P^T & 0\end{bmatrix} \quad (۱۴)$$

اکنون برای محاسبهٔ انرژی لازم برای این تبدیل از رابطه (۱۵) که به‌وسیله بوکستین نشان داده شده استفاده می‌کنیم [۱۷]:

$$C_{tps}=I_F=wkw^T \quad (۱۵)$$

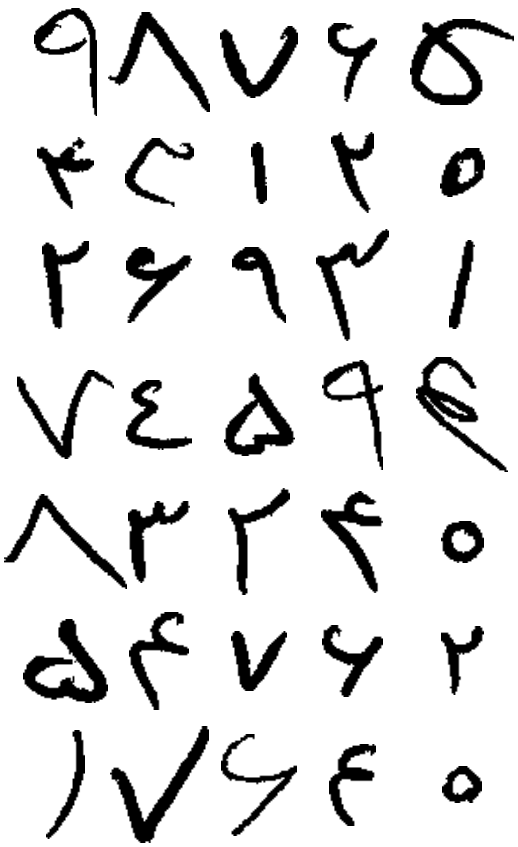
رابطهٔ (۱۵) معیاری را برای پیچیدگی تبدیل به‌دست می‌دهد. هر چه این مقدار بیشتر باشد، فاصله دو شکل بیشتر است. در نهایت به کمک دو معیار فاصله‌ای که به‌دست آوردیم، یعنی معیار فاصله براساس توصیفگرهای نقاط متناظر(رابطهٔ ٤) و معیار فاصله براساس پیچیدگی تبدیل لازم (رابطه ۱۵)، می‌توان معیاری کلی از ترکیب وزنی این دو به‌دست آورد که از آن در بازشناسی ارقام استفاده می‌شود.

٤- بازشناسی ارقام دستنویس فارسی

مجموعه داده ای که ما بازشناسی ارقام را بر روی آن انجام می‌دهیم، شامل ۱۲۸۸رقم دستنویس فارسی است که توسط افراد مختلف نوشته شده و در آن سعی شده تا گوناگونی ارقام فارسی، که در بین عموم شایع است، در آن حفظ شود. نمونه هایی از این ارقام در شکل۲ مشاهده می‌شود.

٤-۱- بازشناسی به کمک انتخاب نماینده

برای طبقه‌بندی ارقام دستنویس، از روش کمترین فاصله از نماینده‌های هر کلاس استفاده کرده ایم. در این روش بسته به گوناگونی هر رقم، تعداد مناسبی نمونه را که بتوانند نماینده‌های خوبی برای توصیف آن رقم باشند، انتخاب می‌کنیم. هر ورودی از مجموعه آزمایش را با هر یک از این نمونه‌ها مقایسه می‌کنیم و آن را به کلاسی نسبت می‌دهیم که کمترین فاصله را با آن دارد. روابط (٤) و (۱۵) به عنوان معیارهای فاصله استفاده شده‌اند.



شکل۲ نمونه هایی از ارقام مجموعه داده

بلونجی و همکاران از روش سه همسایه نزدیکتر، 3-NN، به عنوان معیار فاصله استفاده کرده‌اند [۱۲]. این روش نسبت به روش ما نتایج بهتری می‌دهد زیرا در آن، هر نمونه ورودی با تعداد زیادی از نمونه‌های ارقام مقایسه می‌شود، در نتیجه تصمیم دقیقتری اتخاذ می‌شود.

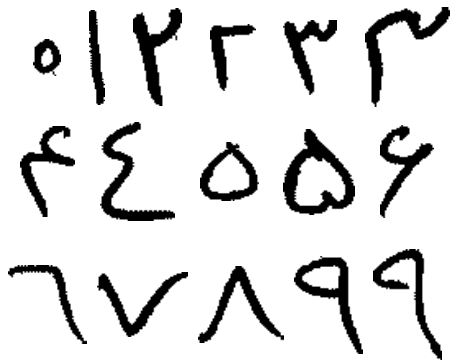
از طرف دیگر در هر بار اجرای الگوریتم، کلاسی انتخاب می‌شود که سه بار به عنوان انتخاب برتر، انتخاب شده است در نتیجه درجه اطمینان تصمیمی که گرفته می‌شود زیاد و خطا کم می‌شود. با وجود این، این روش زمان اجرای بسیار طولانی تری نسبت به روش کمترین فاصله از نماینده کلاس دارد. بنابراین از آنجا که تعداد کلاسهای ارقام مورد استفاده ما و همچنین گوناگونی ارقام در هر کلاس به گونه‌ای است که با انتخاب نماینده‌های مناسب، به کمک روش کمترین فاصله از نماینده کلاس می‌توانیم به نتایج مطلوبی دست یابیم، از این روش استفاده کرده‌ایم.

٤-٢- انتخاب اولیهٔ نماینده‌ها و تنظیم پارامترها

انتخاب نماینده برای ارقام دستنویس فارسی باید با توجه به گوناگونی آنها انجام شود. به عنوان مثال رقم ٤ نسبت به رقم ١ از گوناگونی بیشتری برخوردار است، در نتیجه نیاز به نماینده‌های بیشتری دارد. بنابراین با انتخاب نماینده‌های مناسب برای هر رقم می‌توان به نرخ شناسایی بالاتری برای آن رقم دست یافت. مسأله دیگر آن است که گر چه با افزایش تعداد نماینده‌ها برای هر رقم، احتمال آنکه بیشتر شکلهای ورودی مربوط به آن کلاس درست شناسایی شوند بیشتر می‌شود، اما از طرف دیگر احتمال آنکه این نماینده‌ها، شکلهای ورودی مربوط به ارقام دیگر را نیز به خود جذب کنند بالا می‌رود که این، موجب کاهش نرخ شناسایی می‌شود. بنابراین در این زمینه باید انتخابی بهینه داشته باشیم.

برای آنکه بتوانیم پارامترهای مختلف الگوریتم تطابق شکل را به طرز مناسبی انتخاب و نیز نماینده‌های مناسبی را برای بازشناسی ارقام پیدا کنیم، ابتدا ٣٢٠ رقم را از میان مجموعه ارقام به طور تصادفی انتخاب کردیم. هدف از انتخاب این نمونه‌ها آن است که ابتدا نماینده‌هایی را برای هر رقم انتخاب کنیم و در مرحلهٔ بعد با انتخاب ترکیبات مختلفی از پارامترهای الگوریتم تطابق شکل، بازشناسی را بر روی مجموعهٔ تمرین انجام دهیم. نمونه هایی که در این مرحله به عنوان نماینده انتخاب شد در شکل ٣ نشان داده شده است. در این مرحله برای ارقام ٠ و ١ با توجه به این که تنوع نگارشی کمتری دارند، یک نماینده و برای سایر ارقام دو نماینده را در نظر گرفته ایم.

در این قسمت هدف آن است که پارامترهایی را که به بهترین بازشناسی بر روی مجموعهٔ تمرین منجر می‌شوند، به عنوان پارامترهای تثبیت شده برای بازشناسی بر روی کل مجموعهٔ ارقام انتخاب کنیم. پارامترهای الگوریتم تطابق شکل که باید انتخاب شوند به صورت زیر است.



شکل ٣ نماینده‌های انتخاب شده در کلاسهای ده گانه ارقام برای تعیین پارامترهای الگوریتم.

٤-٢-١- تعداد عناصر زاویه‌ای مختصات لگاریتمی- قطبی

منظور از تعداد عناصر زاویه‌ای مختصات لگاریتمی- قطبی، آن است که در محاسبه هیستوگرام برای هر نقطه نمونه برداری شده بر روی کانتور، از چه تعدادی از عناصر زاویه‌ای استفاده می‌کنیم، برای مثال همانگونه که در شکل ١.ج پیدا است، تعداد عناصر زاویه‌ای برابر ١٢ است.

٤-٢-٢- تعداد عناصر شعاعی در مختصات لگاریتمی – قطبی

منظور از تعداد عناصر شعاعی مختصات لگاریتمی- قطبی آن است که در محاسبه هیستوگرام برای هر نقطه نمونه برداری شده بر روی کانتور، از چه تعدادی از عناصر

شعاعی استفاده می‌کنیم. برای مثال در شکل ۱.ج از پنج عنصر شعاعی استفاده شده است.

٤-۲-۳- روش پیدا کردن نقاط متناظر بین دو شکل

برای پیدا کردن تناظری یک به یک بین نقاط نمونه‌برداری شده در شکل اول و شکل دوم از دو روش ریاضی می‌توان استفاده کرد [۱۲]. روشهایی که به این منظور پیشنهاد شده، روش مجاری [۱۵] و روش یونکر [۱٦] است.

٤-۲-٤- انتخاب ضرایب وزنی فواصل

همانگونه که در بخش قبل اشاره شد، برای مقایسه دو شکل از دو نوع فاصله استفاده می‌کنیم.

۱. فاصله‌ای که از جمع فاصله های بین تک تک نقاط متناظر دو شکل به‌دست می‌آید. این فاصله را با $C_{context}$ نشان می‌دهیم (رابطهٔ ٤).

۲. فاصله‌ای که بر حسب انرژی لازم برای تبدیل اسپلاین صفحه نازک به‌دست می‌آید. این فاصله را با C_{tps} نشان می‌دهیم (رابطهٔ ۱۵).

$$C_{total} = w_{context} C_{content} + w_{tps} C_{tps} \quad (۱٦)$$

C_{total} فاصله‌ای است که از آن برای تعیین اختلاف دو رقم مورد نظر استفاده می‌کنیم. $w_{context}$ و w_{tps} ضرایب وزنی این دو فاصله است. بلونجی و همکاران به ضرایب وزنی $w_{context}=1.6$ و $w_{tps}=0.3$ برای بازشناسی نهایی بر روی مجموعه ارقام دستنویس MNIST رسیده‌اند [۱۲]. به همین دلیل ما در ابتدا برای آنکه انتخابی تقریبا درست برای این ضرایب وزنی داشته باشیم همین مقادیر را انتخاب کردیم.

پس از تعیین پارامترهای فوق، نماینده‌هایی را برای هر یک از ارقام انتخاب و با تغییر نمونه‌ها سعی کردیم نرخ بازشناسی را در مجموعه دادهٔ تمرین بیشتر کنیم. پس از آنکه این نمونه‌ها انتخاب شد، در مرحله بعدی سعی کردیم براساس این نمونه‌ها و تغییر پارامترهای الگوریتم، نرخ نمونه‌برداری را در مجموعه تمرین بهبود دهیم. آزمایشهای گوناگونی برای انتخاب بهترین پارامترهای الگوریتم فوق بر اساس نمونه هایی که در شکل ۳ نشان داده شده، انجام شد. تعداد عناصر زاویه ای مختصات لگاریتمی-قطبی که در این آزمایشها از آنها استفاده شد، برابر ۸، ۱۰، ۱۲، ۱۴ و ۱۶ بود. برای تعداد عناصر شعاعی نیز از انتخابهای ۳ تا ۶ استفاده کردیم. برای تعداد نقاط نمونه برداری شده انتخابهای ۵۰، ۶۰، ۷۵ و ۱۰۰ را در نظر داشتیم. پارامترهایی که در نهایت انتخاب شد، در جدول ۱ نشان داده شده است. نرخ بازشناسی که با این پارامترها و به کمک نمونه‌های انتخاب شده در شکل ۳ بر روی مجموعهٔ تمرین به‌دست آمد برابر ۸۸/۴٪ است.

جدول۱ پارامترهای نهایی الگوریتم تطابق شکل برای بازشناسی ارقام فارسی

تعداد عناصر زاویه‌ای مختصات لگاریتمی- قطبی	۱۲
تعداد عناصر شعاعی مختصات لگاریتمی– قطبی	۵
تعداد نقاط نمونه برداری شده بر روی هر کانتور	٦۰
روش پیدا کردن نقاط متناظر در دو شکل	یونکر
ضرایب وزنی فواصل	$W_{context} = 1.8$ $C_{TPS} = 0.2$

٤-۳- بازشناسی بر روی کل مجموعهٔ ارقام

پس از تثبیت پارامترهای الگوریتم تطابق شکل به کمک همان نماینده‌هایی که در بخش تمرین انتخاب شد، بازشناسی را بر روی کل مجموعهٔ ارقام انجام دادیم. نرخ بازشناسی این آزمایش برابر ۸۴/۷٪ شد. در مراحل بعد برای بهبود نرخ بازشناسی، به اصلاح نماینده‌های انتخاب شده پرداختیم. به این منظور نماینده‌های ۳ و ۶ را اصلاح و یک نماینده را به نماینده‌های قبلی ارقام ۴، ۷ و ۸ اضافه کردیم. همچنین دو نماینده را به نماینده‌های رقم ۵ اضافه نمودیم. این نماینده‌های جدید در شکل ۴ نمایش داده

شده است. در این مرحله به نرخ بازشناسی ٪۸۸/۴ رسیدیم.

شکل ۴ نماینده‌های جدیدی که به نماینده‌های قبلی ۵،۴، ۷ و ۸ اضافه شده‌اند.

در مرحله آخر با افزودن نماینده‌هایی به نماینده‌های ۱ و ۹، که در شکل ۵ نشان داده شده، به نرخ بازشناسی ٪۸۹/۹ رسیدیم. ماتریس کارایی این آزمایش در جدول ۲ نشان داده شده است. همان طور که در این جدول دیده می‌شود، ارقام ۷ و ۸ به دلیل تنوع کمتری که دارند بیشترین نرخ شناسایی را داشته‌اند (٪۹۵ و ٪۹۷) و رقم ۳ به دلیل تنوع زیاد و نزدیکی نحوهٔ نگارش آن با رقم ۲، کمترین میزان بازشناسی را دارد (٪۸۱).

شکل ۵ نماینده‌های جدیدی که به نماینده‌های قبلی ۱ و ۹ اضافه شده‌اند.

دلیل افزودن نمایندهٔ ۴ به سایر نماینده‌ها آن است، که این شکل از رقم ۴ به‌وسیله سایر نمونه هایی که قبلاً برای رقم ۴ داشتیم، پوشش داده نمی شد. دو نمونه رقم ۵ به این دلیل اضافه شد که چون به طور کامل بسته نشده‌اند، کانتوری کاملا متفاوت با نمونه های قبلی رقم ۵ دارند. در مورد ارقام ۷ و ۸، این دو نمونه با این هدف اضافه شد که نمونه های نامتقارن این ارقام را پوشش دهند.

۵- مقایسه با روشهای دیگر

در این بخش مقایسه ای داریم بین نتایج روشی که ما برای بازشناسی ارقام دستنویس فارسی از آن استفاده کرده‌ایم و سایر روشهایی که در سالهای اخیر انجام شده است. این مقایسه در جدول ۳ آورده شده است.

همانگونه که مشخص است، نرخ بازشناسی ارقام به کمک تطابق شکل، فقط از دو روش [۷] و [۱۸] کمتر است؛ اما باید توجه داشت که در بیشتر این روشها از مراحل تکمیلی پس پردازش یا پیش پردازش استفاده شده است که طبیعتاً نرخ بازشناسی را بالا خواهد برد.

جدول ۲ ماتریس کارایی آزمایش نهایی

	0	1	1	2	2	3	3	4	4	4	5	5	5	5	6	6	7	7	8	8	9	9	9	Sum	%
0	120	.	.	.	.	.	.	.	.	2	2	7	1	.	.	.	.	.	.	.	.	.	.	132	90.91
1	3	61	59	6	1	.	.	1	1	.	.	.	.	.	2	.	.	.	.	.	2	.	.	136	88.24
2	.	.	.	85	24	3	.	.	.	3	.	.	.	.	.	1	1	.	.	.	.	2	1	120	90.83
3	.	.	.	5	5	52	48	.	1	11	.	.	.	.	.	.	.	.	.	.	.	.	.	122	81.97
4	.	.	2	7	.	4	2	40	45	20	.	.	.	.	.	.	.	.	.	.	.	.	.	120	87.50
5	12	.	.	.	.	.	1	.	.	1	39	28	24	11	.	.	.	.	3	.	.	.	.	119	85.71
6	.	.	1	5	.	.	.	.	.	3	.	.	.	.	95	16	.	3	.	.	2	.	.	125	88.80
7	.	.	.	3	.	.	.	.	.	.	.	.	.	.	3	.	81	56	.	.	.	.	.	143	95.80
8	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	119	15	3	.	.	138	97.10
9	.	.	3	2	.	.	.	2	.	1	.	.	.	.	.	1	.	.	.	.	38	49	29	125	92.80

جدول۳ مقایسه نتایج روشهای مختلفی که برای بازشناسی ارقام ارائه شده است.
× روشهایی را نشان می‌دهد که از پس پردازش یا پیش پردازش استفاده کرده‌اند.

ردیف	مرجع	استخراج ویژگی	روش طبقه‌بندی	تعداد ارقام	نرخ بازشناسی	
۱	[۷]	روش مکانهای مشخصه مورب	پنج همسایه نزدیکتر	۱۲۷۷۸	٪۸۵/۰۷ تا ٪۹۵/۵	×
۲	[۹]	DTW	کمترین فاصله از نماینده‌های هر کلاس	------	٪۷۵	
۳	[۱۰]	گشتاورهای هندسی و تبدیل فوریه	شبکه عصبی	------	٪۶۰	
٤	[۱۸]	ویژگیهای ساختاری	درخت تصمیم باینری	۳۲۰۰	٪۹۳	×
۵	[۸]	3-tuple	کمترین فاصله از نماینده‌های هر کلاس	۵۰۰	٪۸۱	×
٦	[۱۱]	zoning	روش فازی	۵۰۰	٪۸۳	×
۷	[۱۹]	انتخاب ویژگی با الگوریتم‌های وراثتی	روش فازی	۵۰۰	٪۸٦/۲	
۸	روش تطابق شکل	به کمک اطلاعات جانبی حول نقاط کانتور	کمترین فاصله از نماینده‌های هر کلاس	۱۲۸۸	٪۸۹/۹	

برای مثال در پیش پردازش می‌توان کیفیت تصویری ارقام را بهبود داد [۱۸]، یا ارقامی مانند ۰ و ۱ را قبل از بازشناسی اصلی، از بقیهٔ ارقام جدا کرد [۸].

در پس پردازش نیز با توجه به اشتباهات متداولی که طبقه‌بندی اصلی پیش می‌آید، با انتخاب ویژگیهایی متفاوت از ویژگیهای اصلی، می‌توان بازشناسی را اصلاح کرد [۷-۸].

همچنین باید توجه داشت که در روش ما، فقط از ۱ تا ٤ نمونه به عنوان نمایندهٔ هر کلاس استفاده شده است، اگر همانطور که در بخش ٤-۱ ذکر شد، از روش طبقه‌بندی نزدیکترین همسایه استفاده شود، با هزینهٔ بیشتر، نتایج بهتری به دست خواهد آمد.

٦- نتیجه‌گیری و پیشنهادها

دراین مقاله از نوعی الگوریتم تطابق شکل [۱۲] برای بازشناسی ارقام دستنویس فارسی استفاده شده است. برای هر نقطهٔ نمونه برداری شده بر روی کانتور شکل، توصیفگری بر اساس توزیع مکانی نقاط دیگر کانتور به‌دست آوردیم. برای تعیین میزان شباهت دو شکل، ابتدا بر اساس این توصیفگرها تناظری یک به یک بین نقاط نمونه برداری شده روی کانتور شکل اول با نقاط روی کانتور شکل دوم به دست آورده و جمع فواصل بین نقاط متناظر دو شکل را به عنوان معیاری برای عدم شباهت دو شکل انتخاب کردیم. سپس تبدیلی را تعریف کردیم که نقاط کانتور شکل اول را بر روی نقاط متناظر در کانتور شکل دوم قرار می‌دهد. میزان پیچیدگی این تبدیل، معیار دیگری برای عدم شباهت دو شکل است. در نهایت با بهینه‌سازی پارامترهای مختلف این الگوریتم، از آن برای شناسایی ارقام دستنویس فارسی استفاده شده است.

در آزمایشی بر روی ۱۲۸۸ رقم که افراد مختلف

نوشته‌اند ۸۹/۹٪ بازشناسی درست حاصل شده که این نتیجه بدون هیچ پس پردازشی به‌دست آمده است. البته مشخص است که تغییر هر یک از پارامترهایی که در جدول ۱ آمده است ، برکارایی الگوریتم تطابق شکل تأثیر خواهد داشت، اما آنچه مهم است، ترکیب مناسب این پارامترها می‌باشد. مقدار هر یک از این پارامترها بر اساس ویژگیهایی که مجموعه ارقام دستنویس مورد استفاده دارند، دارای محدودیتهایی خواهند بود. برای مثال چون ارقام دستنویس به‌کار رفته از نظراندازه نرمالیزه نیستند، تعداد نقاط روی مرز آنها برای ارقام مختلف متفاوت است.

برای بالا بردن نرخ بازشناسی می‌توان ارقام صفر را برحسب تعداد نقاط روی مرز آنها بازشناسی کرد و در نتیجه الگوریتم تطابق شکل را برای ارقام غیرصفر به کار برد. در این صورت می‌توان از تعداد نقاط نمونه برداری شدهٔ بیشتری برای توصیف ارقام استفاده کرده و در نتیجه توصیف قوی تری از ارقام به دست آورد.

مشکل دیگری که در این روش بازشناسی وجود دارد آن است که با وجود آنکه برای هر رقم، سعی کردیم از تعداد مناسبی نماینده استفاده کنیم - و همچنین برای انتخاب نماینده‌های بهینه نیز آزمایشهای گوناگونی را پس از تثبیت پارامترهای الگوریتم انجام دادیم و تا حد امکان نیز به نماینده‌های بهینه‌ای برای بازشناسی کلی دست یافتیم - اما با این حال به دلیل تنوع نحوهٔ نگارش ارقام دستنویس فارسی، این ۲۳ نماینده نمی‌توانند تمام حالات موجود را پوشش دهند. در زمینهٔ انتخاب نماینده و تعداد نماینده‌های هر کلاس، هنوز جای کار وجود دارد.

۷- منابع

[1] O.D.Trier; A.K.Jain; T.Taxt; "Feature Extraction Methods for Character Recognition- A Survey"; *Pattern Recogniton*; Vol. 29; No. 4; 1996; pp. 641–662.

[2] R.Plamondon; S. Srihari; "On-line and Off-line Handwritten Recognition: A Comprehensive Survey"; *IEEE Trans. on Pattern Analaysis and Machine Intelligence*; Vol. 22; No. 1; January 2000; pp. 63-83.

[3] N. Arica; F.T. Yarman-Vural; "An overview of caracter recognition focused on off-line handwritting"; *IEEE Trans. on Systems; Man and Cybernetics-Part C: Application and Reviews*; Vol. 31; No. 2; May 2001; pp. 216-233.

[4] C. Suen; M. Berthod; "Automatic Recognition of Handprinted Characters- The state of the Art"; *Proceedings of the IEEE*, Vol 68; No 4; April 1980; pp. 469-487.

[5] A. Amin; "Off-Line Arabic Caracter Recognition: The State of the Art"; *Pattern Recognition*; Vol. 31; No. 5; 1998; pp. 517-530.

[۶] ر.عزمی؛ ا.کبیر؛ "شناسایی آماری حروف دستنویس فارسی" ؛ دومین کنفرانس مهندسی برق ایران؛ ۱۳۷۳؛ ص ۲۷۷- ۲۸۵.

[۷] ح. ر. نفیسی؛ ا. کبیر؛ " شناسایی ارقام دستنویس فارسی"؛ دومین کنفرانس مهندسی برق ایران؛ ۱۳۷۳؛ ص ۲۹۵- ۳۰۴.

[۸] س. م. رضوی؛ ا. کبیر؛ " خواندن اتوماتیک فرمهای انتخاب درس"؛ سومین کنفرانس بین المللی سالانه انجمن کامپیوتر ایران؛ ۱۳۷۶؛ ص ۶۴- ۶۹.

[۹] ک. مسروری؛ و. پور محسنی خامنه؛ " شناسایی ارقام دستنویس با استفاده از الگوریتم DTW"؛ چهارمین کنفرانس بین المللی انجمن کامپیوتر ایران؛ ۱۳۷۷؛ ص ۱- ۷.

[10] S. Mansoory; H. Hassibi; F. Rajabi; "A Heuristic Persian Handwritten Digit Recognition with Neural Network"; *6th Iranian Conference on Electrical Engineering*; Vol. 3; 1998; pp. 131-135

[۱۱]و. جوهری مجد؛ س. م. رضوی؛ "بازشناسی فازی ارقام دستنویس فارسی"؛ اولین کنفرانس ماشین بینایی و پردازش تصویر ایران؛ ۱۳۷۹؛ ص ۱۴۴- ۱۵۱.

[12] S. Belongie; J. Malik; J. Puzicha; "Shape Machting and Object Recognition Using Shape Context." *IEEE transaction on Pattern Analysis and Machine Intelligence*; Vol. 24; No. 4; 2002; pp. 509-522.

[13] D. Cheng; H. Yan; "Recognition of Handwritten Digits Based on Contour Information". *Pattern Recognition*; Vol. 31; No. 3; 1999; pp. 235-255.

[14] S. Chakravarthy; B. kompella; "The Shape of handwritten characters"; *Pattern Recognition Letters*; Vol. 24; 2003; pp. 1901-1913.

[15] C. Papadimitriou; k. Steiglitz; *Combinatorial Optimization: Algorithm and Complexity*; Prentice- Hall, 1982.

[16] R. Jonker; T. Volgenant; "A Shortest Augmenting Path Algorithm for Dense and Sparse Linear Assignment Problems"; *Computing*; Vol. 38; 1987; pp. 325-240.

[17] Bookstein; F.L.; "Principal Warps: Thin Plate Splines and Decompositions of Deformations"; *IEEE Trans. Pattern Analysis and Machine Intelligence*; Vol. 11; No. 6; 1989; pp. 567-585.

[۱۸] ح. کتابدار؛ " شناسایی ارقام دستنویس فارسی به روش ساختاری"؛ ششمین کنفرانس مهندسی برق ایران؛ ۱۳۷۷؛ ص ۴-۱۹ و ۴-۲۴.

[۱۹] س. م. رضوی؛ ه. صدوقی یزدی؛ ا. کبیر؛ "انتخاب ویژگی برای بازشناسی ارقام دستنویس فارسی به کمک الگوریتمهای وراثتی"؛ هفتمین کنفرانس سالانه انجمن کامپیوتر ایران؛۱۳۸۰؛ ص ۲۸۵- ۲۹۲.