

## فیلترهای چند جمله‌ای مدوله شده با دوره محدود و کاربرد آن‌ها در طبقه‌بندی تصویر

سلیمه بامری، کارشناس ارشد، سعید سریزدی، دانشیار، حسین نظام‌آبادی‌پور، دانشیار

بخش مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران

nezam@mail.uk.ac.ir و saryazi@mail.uk.ac.ir, salime\_bameri@yahoo.com

چکیده: فیلترهای گابور، موجک‌های گوسی مدوله شده هستند که به علت توانایی در استخراج ویژگی‌های تصویر در جهت‌ها و فرکانس‌های متفاوت به‌طور وسیع مورد استفاده قرار می‌گیرند. اما به دلیل دوره نامحدود مکانی و تابع تبدیل غیرکسری در حوزه فرکانس پیاده‌سازی دقیق این فیلترها امکان‌پذیر نیست و پیاده‌سازی تقریبی آن‌ها، منجر به افت کارایی آن‌ها می‌گردد. در این مقاله یک موجک مدوله شده بر اساس فیلترهای چند جمله‌ای با دوره محدود ارائه می‌شود که خصوصیتی نزدیک به موجک گابور دارد و به علت دارا بودن دوره محدود، علاوه بر امکان پیاده‌سازی دقیق، از سرعت بالاتری نسبت به گابور برخوردار است. به منظور ارزیابی فیلترهای ارائه شده، از آن‌ها در دو کاربرد متفاوت طبقه‌بندی شامل طبقه‌بندی معنایی به منظور بازیابی تصویر و بازشناسی قلم فارسی استفاده شده است. نتایج مقایسه‌ای با فیلترهای گابور، بهبود توأم سرعت و کارایی در این کاربردها را تایید می‌کند.

واژه‌های کلیدی: فیلتر گابور، فیلتر با دوره محدود، استخراج ویژگی، بازیابی تصویر، طبقه‌بندی معنایی تصویر، بازشناسی قلم‌های فارسی.

## Polynomial modulated filters with compact support and their application in image classification

S. Bameri, S. Saryazdi and H. Nezamabadi-pour  
Shahid Bahonar University of Kerman, Kerman, Iran

**Abstract:** Gabor filters are Gaussian modulated wavelets that due to their ability to extract features in different directions and frequencies in an image are used widely in different applications. However, because of their infinite support, an exact realization is not practically possible and an approximated realization results in losing performance. In this paper, a new modulated wavelet with compact support and close to Gabor wavelet is proposed. Due to its compact support, an exact and fast implementation is possible. To evaluate the performance of the proposed filters a comparative study between them and Gabor filters in two classification application containing semantic image classification and Farsi font recognition is given. The obtained results confirm the performance improvement of the proposed filters in both speed and precision.

**Keywords:** Gabor filters, Compact support filter, Feature extraction, Image retrieval, Semantic image classification, Farsi font recognition.

## ۱- مقدمه

در بسیاری از زمینه‌های پردازش تصویر استخراج ویژگی‌های تصویر یک مرحله اجتناب‌ناپذیر است. از آنجایی که خصوصیات مورد نظر در تصویر مقیاس‌ها و جهت‌های گوناگونی دارند، لذا استخراج اطلاعات و ویژگی‌های جهت‌دار در مقیاس‌های مختلف از تصویر یکی از گام‌های اساسی پیش‌پردازش می‌باشد. امروزه فیلترهای گابور به علت خواص مناسبی که دارند برای رسیدن به این منظور به‌طور وسیع مورد استفاده قرار می‌گیرند.

گابور در سال ۱۹۴۶ [۱] اصل عدم قطعیت را برای اطلاعات از روی روابطی در مکانیک کوانتوم استنتاج کرد. طبق این اصل دقت توام یک سیگنال در دو حوزه زمان و فرکانس (حاصل ضرب پهنای باندهای زمانی و فرکانسی آن) توسط یک حد پایینی محدود می‌شود. سپس او گروهی از توابع یک بعدی را معرفی کرد که به حد پایین اصل عدم قطعیت و به عبارتی مینیمم رزلوشن توام در زمان و فرکانس دست می‌یابند و آن‌ها را سیگنال‌های (توابع) بنیادی نامید. داگمن در سال ۱۹۸۰ با الهام از کار گابور روابط عدم قطعیت را در دو بعد بیان و خانواده‌ای از توابع دو بعدی را که به مینیمم مقدار در اصل عدم قطعیت دست می‌یابند، معرفی کرد و آن‌ها را توابع گابور نام گذاشت. تابع گابور دو بعدی از حاصل ضرب تابع گوسی دوبعدی در تابعی سینوسی در جهات مختلف فضای دو بعدی به دست می‌آید [۲].

این توابع به عنوان فیلتر به علت خواص بسیار مفیدشان کاربرد زیادی در زمینه‌های مختلف بینایی ماشین نظیر تحلیل بافت [۳ و ۴]، طبقه‌بندی [۵ و ۶]، بازیابی تصویر [۷]، تشخیص قلم [۸] و ... پیدا کرده‌اند. از جمله این خواص می‌توان به سادگی، تمرکز توانان بهینه در حوزه مکان و فرکانس و امکان انتخاب جهت و فرکانس برای استخراج اطلاعات تصویر اشاره کرد.

هر چند فیلترهای گابور دارای خصوصیات بسیار مناسب و منحصر به فرد هستند، متأسفانه پیاده‌سازی دقیق آن‌ها امکان‌پذیر نیست. زیرا از یکسو دارای دوره مکانی نامحدود هستند و بنابراین پیاده‌سازی دقیق آن‌ها در حوزه مکان امکان‌پذیر نیست [۵، ۹ و ۱۰] از سوی دیگر، از آنجا که مشخصه فرکانسی این فیلترها تابعی کسری<sup>۱</sup> نیست، پیاده‌سازی دقیق آن‌ها توسط یک ساختار تکراری<sup>۲</sup> نیز امکان‌پذیر نمی‌باشد. از طرف دیگر پیاده‌سازی تقریبی فیلترهای گابور علاوه بر این که باعث اتلاف مقداری از اطلاعات می‌شود، سبب از دست رفتن بعضی خواص مفید این فیلترها نیز می‌شود.

فیلترهای گوسی علاوه بر کاربرد در فیلترهای گابور در زمینه‌های دیگری همچون تجزیه و تحلیل در فضای مقیاس [۱۱]، آشکارسازی لبه [۱۲ و ۱۳]، مدل‌سازی تصاویر بر اساس معادلات دیفرانسیل پاره‌ای و انتشار و ... مورد توجه قرار گرفته‌اند. تابع گوسی دارای خصوصیتی است که آن را منحصر به فرد ساخته است. از جمله این خصوصیات می‌توان به موارد ذیل اشاره کرد [۱۴]:

- تابع گوسی از دیدگاه اصل عدم قطعیت بهینه است.
- این تابع جواب معادله انتشار حرارت است که این معادله در پیاده‌سازی تصویر مورد استفاده قرار می‌گیرد [۱۴].
- تنها تابع دوبعدی تفکیک‌پذیر با تقارن دایره‌ای است.
- کانولوشن دو گوسی خود تابعی گوسی است.
- بینهایت بار مشتق‌پذیر است.

به لحاظ اهمیت تابع گوسی، تلاش‌های فراوانی برای رفع مشکل پیاده‌سازی دقیق آن در تحقیقات علمی گزارش شده است. دسته اول مقالات تقریب تابع گوسی توسط یک فیلتر FIR را پیشنهاد می‌کنند [۱۲]. دسته دوم تقریب این تابع در حوزه فرکانس توسط توابع کسری را مطرح می‌سازند [۱۵]. در دسته سوم مقالات [۱۴ و ۱۶] به جای تقریب تابع گوسی، سعی در یافتن توابعی با دوره محدود و با خصوصیات شبیه گوسی می‌شود.

به منظور برطرف کردن مشکل پیاده‌سازی فیلترهای گابور، در این مقاله یک گروه فیلتر جدید با دوره محدود معرفی می‌شود که بر اساس ایده دسته سوم مقالات و مشخصاً کار قبلی ما در زمینه هسته‌های چند جمله‌ای با دوره محدود (PKCS)<sup>۳</sup> [۱۴] حاصل شده‌اند. بدین منظور، ابتدا یک موجک مدوله شده بر اساس فیلترهای چند جمله‌ای با دوره محدود ارائه می‌شود که خصوصیات نزدیک به موجک گابور دارد و به علت دارا بودن دوره محدود، علاوه بر امکان پیاده‌سازی دقیق، از سرعت بالاتری نسبت به گابور برخوردار است، سپس با گسسته‌سازی جهت و فرکانس مرکزی این موجک، مشابه با فیلترهای گابور، فیلترهای چند جمله‌ای مدوله شده با دوره محدود را تعریف می‌کنیم و به منظور ارزیابی آن‌ها، این فیلترها را در دو کاربرد متفاوت طبقه‌بندی شامل طبقه‌بندی معنایی تصویر و تشخیص قلم فارسی به کار می‌بریم. ساختار این مقاله بدین ترتیب است: بخش دوم به توصیف هسته‌های چند جمله‌ای با دوره محدود و بخش سوم به یادآوری تبدیل فیلتر گابور اختصاص دارد. در بخش چهارم فیلتر جدید را معرفی می‌کنیم. در بخش پنجم به معرفی دو کاربرد نوعی فیلترهای حوزه مکان-فرکانس شامل طبقه‌بندی معنایی تصویر و بازشناسی قلم‌های فارسی خواهیم پرداخت. سپس، در بخش ششم به منظور نتایج کار آن را با نتایج حاصل از فیلترهای گابور در دو کاربرد معرفی شده مقایسه می‌کنیم. در انتها و در بخش هفتم به جمع‌بندی و نتیجه‌گیری خواهیم پرداخت.

## ۲- فیلتر گابور

تبدیل گابور، در دسته تبدیل‌های موجک با پنجره مدوله شده قرار می‌گیرد. این تبدیل به علت دسترسی به حد پایین اصل عدم قطعیت، که سبب رسیدن به حداکثر دقت توام مکان-فرکانس می‌گردد، بسیار مورد استقبال محققین در پردازش تصویر قرار گرفته است. با استفاده

در این رابطه  $C_\gamma$  ثابت نرمالیزاسیون است و باید طوری انتخاب شود که حجم زیر هسته برابر واحد گردد یعنی  $\iint \rho_{\sigma,\gamma}(x,y) dx dy = 1$  شود.

هسته‌های KCS با حفظ برخی ویژگی‌های هسته گوسی [۱۶] به طور دقیق نیز قابل پیاده‌سازی هستند اما معایبی نیز دارند: الف) پیچیده هستند و ثابت نرمالیزاسیون  $C_\gamma$  به‌طور تحلیلی قابل محاسبه نیست و باید با روش‌های عددی محاسبه شود. ب) واریانس KCS (متغیر مقیاس) تعریف نشده است و نمی‌توان آن را به صورت تابعی از  $\lambda, \gamma$  به دست آورد. ج) در معادله انتشار که از ویژگی‌های مهم هسته گوسی است صدق نمی‌کند [۱۶].

در کار دوم [۱۴] یک گروه جدید از KCS توسط سریزدی و شریعت معرفی شده است که معایب هسته KCS را ندارند. روند تولید هسته جدید با نام PKCS مشابه KCS است. از آنجایی که انتخاب نگاشت W که فضای  $\mathbb{R}^2$  را به داخل دایره واحد تصویر می‌کند منحصر به فرد نیست در این کار از نگاشت دیگری استفاده شده است:

$$\rho_{\lambda,\gamma}(x,y) = \begin{cases} \frac{1}{C_\gamma} e^{\gamma \ln(1 - (x^2+y^2)/\lambda^2)} & \text{if } x^2 + y^2 < \lambda^2 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

پس از اعمال نگاشت فوق به هسته گوسی نرمالیزه دو بعدی (در فرم قطبی) پاسخ ضربه هسته جدید بدین صورت به دست می‌آید:

$$\varphi_{\lambda,\gamma}(r) = \begin{cases} \frac{\gamma+1}{\pi \lambda^{2\gamma+2}} (\lambda^2 - (r^2))^\gamma & \text{if } (r^2) < \lambda^2 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

که در این رابطه  $\gamma = \frac{1}{2\sigma^2}$  است، همچنین ثابت نرمالیزاسیون  $C_\gamma$  برابر با  $\frac{\pi}{\gamma+1}$  می‌باشد و واریانس آن به صورت تابعی از  $\lambda, \gamma$  قابل بیان است (رابطه (۶)):

$$\sigma^2 = \frac{\lambda^2}{2(\gamma+2)} \quad (6)$$

هسته PKCS علاوه بر آن که مانند هسته KCS دارای دوره محدود بوده و به طور دقیق قابل پیاده‌سازی است و برخی از معایب هسته KCS را برطرف می‌نماید [۱۴]: الف) هسته PKCS به فرم ساده چند جمله‌ای است و ثابت نرمالیزاسیون  $C_\gamma$  به راحتی محاسبه می‌شود. ب) واریانس PKCS به صورت تابعی از  $\lambda, \gamma$  قابل بیان است. ج) هسته PKCS جواب تقریبی معادله انتشار حرارت می‌باشد.

#### ۴- معرفی فیلتر جدید

هدف اصلی این تحقیق طراحی فیلترهایی با دوره زمانی (مکانی) محدود با خصوصیات شبیه و نزدیک به فیلترهای گابور است. مشابه روش طراحی هسته‌های KCS و PKCS از روی هسته گوسی، که در بخش ۳ به تفصیل تشریح شد، فیلترهای جدید، از روی فیلترهای

از تبدیل موجک دو بعدی گابور می‌توان ویژگی‌های جهت‌دار تصویر را در مقیاس‌های مختلف استخراج کرد. قسمت حقیقی موجک گابور دو بعدی و تبدیل فوریه آن مطابق روابط (۱) و (۲) است:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \cos(2\pi u_0 x) \quad (1)$$

$$G(u,v) = \frac{1}{2} \left( \exp\left(-\frac{1}{2}\left(\frac{(u-u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right) + \exp\left(-\frac{1}{2}\left(\frac{(u+u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right) \right) \quad (2)$$

که در آن  $u_0$  فرکانس موج سینوسی را در راستای محور  $x$  و فرکانس مرکزی تبدیل فوریه را مشخص می‌کند.  $\sigma_x$  و  $\sigma_y$  به ترتیب انحراف معیار پوش گوسی در راستای محورهای  $x$  و  $y$  و  $\sigma_u = \frac{1}{2\pi\sigma_x}$  و  $\sigma_v = \frac{1}{2\pi\sigma_y}$  به ترتیب انحراف معیار تابع گوسی تبدیل فوریه در راستای محورهای  $u$  و  $v$  هستند.

تبدیل موجک گابور به عنوان ابزاری برای استخراج ویژگی در زمینه‌هایی همچون طبقه‌بندی، بازیابی تصویر، تشخیص قلم و تحلیل بافت و ... مورد استفاده قرار می‌گیرد. اما از آنجایی که موجک گابور، گوسی مدوله شده است، پیاده‌سازی دقیق آن به همان دلایل ذکر شده در مورد تابع گوسی امکان‌پذیر نمی‌باشد. در این مقاله هسته‌ای جدید با دوره محدود با خواصی شبیه هسته گابور معرفی می‌گردد.

#### ۳- هسته‌های چندجمله‌ای با دوره محدود (PKCS)

همان گونه که اشاره شد، فیلترهای گوسی به علت داشتن خواصی مناسب که آن‌ها را منحصر به فرد می‌سازد به‌طور گسترده در زمینه‌های مختلف پردازش تصویر مورد استفاده قرار می‌گیرند. اما به دلیل دوره نامحدود، پیاده‌سازی دقیق آن‌ها در حوزه مکان امکان‌پذیر نیست. برای رفع این نقطه ضعف در دو تحقیق انجام شده [۱۴ و ۱۶] هسته‌های جدید با دوره محدود با خواصی نزدیک به گوسی تعریف شده‌اند. در ادامه به اختصار به مرور این دو کار پرداخته خواهد شد.

در کار اول که توسط شریعت و رمکی انجام شده است [۱۶]، یک گروه فیلتر جدید با نام KCS<sup>۱</sup> با پاسخ ضربه‌ای که دارای دوره محدود است از فیلترهای گوسی مشتق شده است. به این ترتیب که ابتدا توسط یک نگاشت، فضای  $\mathbb{R}^2$  به داخل یک دایره واحد<sup>۵</sup> تصویر می‌شود. این تبدیل باعث فشرده شدن تمام اطلاعات گوسی در این دایره واحد می‌شود. پاسخ ضربه این فیلترها در خارج دایره واحد صفر است. از این رو محاسبه دقیق کانولوشن امکان‌پذیر است. هسته تک پارامتری ( $\gamma$ ) KCS با دوره  $\lambda$  در مختصات دکارتی طبق رابطه (۳) تعریف می‌شود:

$$\rho_{\lambda,\gamma}(x,y) = \begin{cases} \frac{1}{C_\gamma} e^{\frac{\gamma \lambda^2}{(x^2+y^2)-\lambda^2+\gamma}} & \text{if } x^2 + y^2 < \lambda \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

$$\Phi_{\sigma_x, \sigma_y}(x, y) = \frac{1}{\lambda_x \lambda_y} \Phi_{\sigma} \left( \frac{x}{\lambda_x}, \frac{y}{\lambda_y} \right) = \begin{cases} \frac{1}{C_{\gamma} \lambda_x \lambda_y} (1 - (\frac{x^2}{\lambda_x^2} + \frac{y^2}{\lambda_y^2})^{\gamma}) \times \cos \left( 2\pi \cos \left( \tan^{-1} \left( \frac{y \lambda_x}{x \lambda_y} \right) \right) \sqrt{-\ln \left( 1 - (\frac{x^2}{\lambda_x^2} + \frac{y^2}{\lambda_y^2}) \right)} \right) \frac{x^2 + y^2}{\lambda_x^2 + \lambda_y^2} < 1 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

که در این رابطه،  $\lambda_x$  و  $\lambda_y$  به ترتیب مشخص کننده شعاع مقطع بیضی شکل فیلتر در راستای محورهای X و Y بوده و مشابه روابط مرجع [۱۴] تعریف می‌شود:

$$x = r \cos \varphi \quad (15)$$

$$\lambda_y = \sigma_y \sqrt{2(\gamma + 2)} \quad (16)$$

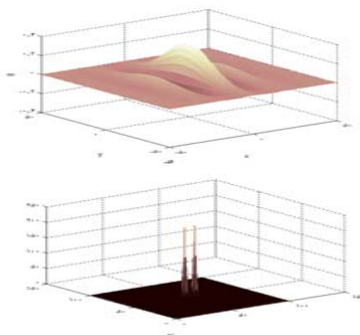
در این روابط،  $\sigma_x$  و  $\sigma_y$  انحراف معیار فیلتر گابور رابطه (۱۷) با مقطع بیضی شکل هستند که فیلترهای جدید مشابه با آن و با دوره محدود ساخته شده‌اند:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left( -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right) \times \cos(2\pi u_0 x) \quad (17)$$

از آنجاکه فیلترهای جدید از حاصل ضرب یک چند جمله‌ای در تابع کسینوسی تشکیل شده‌اند برای آن‌ها نام "فیلترهای چند جمله‌ای مدوله شده با دوره محدود" را انتخاب می‌کنیم. در شکل (۱) پاسخ ضربه فیلتر گابور به همراه تبدیل فوریه‌اش و در شکل (۲) پاسخ ضربه فیلتر چند جمله‌ای مدوله شده با دوره محدود و تبدیل فوریه‌اش نشان داده شده است.

## ۵- استخراج ویژگی در سیستم بازیابی تصویر بر اساس محتوا و بازشناسی قلم فارسی

در این بخش به منظور ارزیابی عملکرد فیلترهای پیشنهاد شده به معرفی دو کاربرد مهم این فیلترها در زمینه بینایی ماشین پرداخته می‌شود. سپس در بخش بعد، نتایج آزمایش‌ها ارائه خواهد شد



شکل (۱): (الف) فیلتر گابور با پارامترهای  $\sigma_x = 1/59$ ,  $\sigma_y = 2/1$  (ب)

تبدیل فوریه‌اش

گابور با نگاشت فضای  $\mathbb{R}^2$  به داخل دایره واحد حاصل شده، سپس با تغییر مقیاس در جهت‌های X و Y دوره دایره‌ای شکل فیلترها به دوره بیضی شکل با اندازه دلخواه تبدیل می‌شود. برای ساخت فیلترهای جدید از روی فیلتر گابور، ابتدا فیلتر گابور دو بعدی نرمالیزه با مقطع دایره‌ای  $(\sigma_x = \sigma_y = \sigma)$  را در مختصات دکارتی در نظر می‌گیریم:

$$g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \cos(2\pi u_0 x) \quad (7)$$

اکنون با استفاده از روابط (۸) و (۹) فیلتر گابور رابطه‌ی (۷) را به فرم قطبی بازنویسی می‌کنیم (رابطه (۱۰)):

$$x = r \cos \varphi \quad (8)$$

$$y = r \sin \varphi \quad (9)$$

$$g_{\sigma}(r, \varphi) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{r^2}{2\sigma^2} \right) \cos(2\pi u_0 r \cos \varphi) \quad (10)$$

در این روابط دامنه تغییرات r از صفر تا  $+\infty$  است که باید به داخل دایره واحد محدود شود. برای این منظور از همان نگاشت مرجع [۱۴] با رابطه (۲) استفاده می‌شود. با اعمال این نگاشت به فیلتر رابطه (۱۰)، فیلتر با دوره محدود به دایره واحد رابطه (۱۱) به دست می‌آید:

$$\Phi_{\sigma}(r, \varphi) = \begin{cases} \frac{1}{C_{\gamma}} e^{\gamma \ln(1-r^2)} \cos \left( 2\pi u_0 \cos \varphi \sqrt{-\ln(1-r^2)} \right) & r^2 < 1 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

که نمایش آن در مختصات دکارتی برابر است با:

$$\Phi_{\sigma}(x, y) = \begin{cases} \frac{1}{C_{\gamma}} (1 - (x^2 + y^2))^{\gamma} \times \cos \left( 2\pi u_0 \cos \left( \tan^{-1} \left( \frac{y}{x} \right) \right) \sqrt{-\ln(1 - (x^2 + y^2))} \right) & x^2 + y^2 < 1 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

در رابطه بالا  $\gamma = \frac{1}{2\sigma^2}$  و  $C_{\gamma}$  ثابت نرمالیزاسیون است و به گونه‌ای انتخاب می‌شود که حجم زیر تابع  $\Phi_{\sigma}(x, y)$  برابر واحد شود و با

محاسبه انتگرال رابطه (۱۳) به دست می‌آید:

$$C_{\gamma} = \iint_{|x^2+y^2| \leq 1} (1 - (x^2 + y^2))^{\gamma} \cos \left( 2\pi u_0 \cos \left( \tan^{-1} \left( \frac{y}{x} \right) \right) \sqrt{-\ln(1 - (x^2 + y^2))} \right) dx dy = \int_0^{2\pi} \int_0^1 (1 - r^2)^{\gamma} \cos \left( 2\pi u_0 \cos \varphi \sqrt{-\ln(1 - r^2)} \right) r dr d\varphi \quad (13)$$

از آنجا که فیلتر به صورت گسسته پیاده‌سازی می‌شود مقدار ثابت نرمالیزاسیون به راحتی با جمع درایه‌های ماسک معرف فیلتر به دست می‌آید. در نهایت فیلترهای دو بعدی با دوره محدود و مقطع بیضی شکل با اندازه دلخواه را با تغییر مقیاس زیر تعریف می‌کنیم:

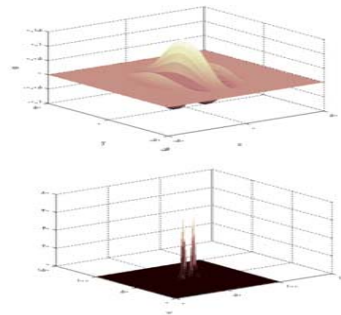
با سامانه ضروری است، اما ارائه ویژگی‌های جدید که بتوانند دقت سامانه بازیابی را افزایش دهد، همواره مورد توجه محققان در این زمینه بوده است. از سوی دیگر ارائه روش‌هایی که بتواند ویژگی‌ها را در یک زمان کوتاه‌تر نسبت به روش‌های موجود با دقتی بیشتر یا معادل با روش‌های موجود اخذ کند می‌تواند با کاهش زمان لازم برای نمایه‌سازی در پایگاه‌های بزرگ مورد توجه قرار گیرد.

استفاده از طبقه‌بندی معنایی تصاویر پایگاه داده می‌تواند به‌طور قابل توجهی بازدهی سیستم بازیابی را افزایش دهد. دسته‌بندی تصاویر به گروه‌های معنایی با استفاده از ویژگی‌های دیداری سطح پایین مسئله‌ای حاد در بازیابی تصویر است [۲۲]. اگر گروه معنایی مورد نظر کاربر مشخص شود، سیستم فقط آن گروه را برای رسیدن به تصاویر دلخواه کاربر جستجو می‌کند. این مسئله موجب افزایش کارایی و در عین حال افزایش سرعت سیستم بازیابی می‌شود [۲۳]. در این مقاله از ویژگی‌های مکان فرکانس برای استخراج ویژگی بافت در یک سامانه بازیابی تصویر به منظور طبقه‌بندی معنایی تصویر استفاده می‌شود. برای استخراج ویژگی، فیلترهای پیشنهادی و فیلترهای گابور مورد آزمون قرار می‌گیرند.

## ۵-۲- بازشناسی قلم‌های فارسی

یکی از اساسی‌ترین مراحل در تحلیل اسناد و تشخیص متن، بازشناسی قلم است که شامل مواردی مانند زبان متن، اندازه قلم و نوع قلم می‌باشد. روش‌های ارائه شده برای تشخیص قلم را می‌توان به دو دسته اساسی تقسیم کرد: الف) روش‌های مبتنی بر تصویر کلمات یا یک خط از متن (استفاده از ویژگی‌های موضعی) ب) روش‌های مبتنی بر تصویر یک صفحه از متن (استفاده از ویژگی‌های کلی). نوع قلم در روش‌های دسته اول با تحلیل ویژگی‌های کلمات و حروف مانند شیب، انحنا و پروفیل و تفاوت‌های آن‌ها در قلم‌های متفاوت و در روش‌های دسته دوم با استخراج ویژگی‌های قلم از کل متن تشخیص داده می‌شود [۲۴].

در مرجع [۲۵] با استفاده از ویژگی‌هایی مانند مقادیر پروفیل عمودی، چگالی پیکسل‌های سیاه و واریانس پروفیل افقی، استیل‌های معمولی، ایتالیک، رومن و پرنگ در سه نوع قلم تشخیص داده شده است. در این مرجع از یک طبقه‌بند بیز برای طبقه‌بندی قلم‌ها استفاده شده است. در [۲۶] از یک خط متن ۸ ویژگی با استفاده از پروفیل‌های عمودی، افقی و اجزاء پیوسته استخراج شده و ۱۰ نوع قلم در ۷ اندازه و ۴ نوع مختلف براساس این ویژگی‌ها طبقه‌بندی شده‌اند. در [۲۷] از ویژگی‌های نوشتاری برای تشخیص ۷ نوع قلم متداول انگلیسی با اندازه‌های ۷ تا ۱۸ استفاده شده است. هر تصویر به ناحیه صعودی<sup>۱۲</sup>، ارتفاعی<sup>۱۳</sup> و نزولی<sup>۱۴</sup> تقسیم و برای هر نوع قلم انواع سرفه‌ها در نواحی مختلف استخراج شده است. این ویژگی‌ها به یک طبقه‌بند شبکه عصبی داده می‌شود که در خروجی خود نوع قلم را مشخص می‌کند.



شکل (۲): الف) فیلتر جدید با پارامترهای  $\gamma = 3, \lambda_y = 6/64, \lambda_x = 5/03$  (ب) تبدیل فوریه‌اش

## ۵-۱- بازیابی تصویر بر اساس محتوا، CBIR

از اوایل دهه ۹۰، بازیابی تصویر بر اساس محتوا، یک زمینه فعال برای تحقیقات قلمداد می‌شود. تا به حال سامانه‌های بازیابی تجاری و تحقیقاتی فراوانی ایجاد و معرفی شده است. در همه این سامانه‌ها، ویژگی‌های سطح پایین تصاویر به‌طور خودکار استخراج شده و تصاویر با آن‌ها نمایه‌سازی می‌شوند. به‌طور کلی، سامانه‌های بازیابی تصویر در دو مرحله عمل می‌کنند. در مرحله اول، ویژگی‌های دیداری تصاویر پایگاه به‌صورت خودکار استخراج شده و تصاویر با آن‌ها نمایه‌سازی می‌شوند. در مرحله دوم، پس از دریافت تصویر پرس و جوی کاربر، ویژگی‌های سطح پایین<sup>۸</sup> یا ویژگی‌های دیداری آن استخراج شده و پایگاه ویژگی‌های دیداری برای یافتن نزدیکترین تصاویر به پرس و جوی جستجو می‌شوند.

استخراج ویژگی، کلیدی‌ترین بخش از یک سیستم بازیابی تصویر بر اساس محتوا است. ویژگی‌ها به دو دسته عمده ویژگی‌های عمومی<sup>۹</sup> و ویژگی‌های خاص<sup>۱۰</sup> تقسیم می‌شوند [۱۷ و ۱۸]. ویژگی‌های عمومی معمولاً شامل رنگ، بافت، شکل و روابط مکانی<sup>۱۱</sup> هستند [۲۱-۱۸]. ویژگی‌های خاص، وابسته به حوزه کاربرد هستند. این ویژگی‌ها، در یک سیستم بازیابی عمومی که برای تصاویر گوناگون به‌کارگرفته می‌شود، کارایی خوبی ندارند.

ویژگی بافت یکی از ویژگی‌های مهمی است که به دفعات در بازیابی تصویر از آن استفاده شده است. این ویژگی با استخراج اطلاعات فرکانسی تصاویر به نمایه‌سازی آن‌ها می‌پردازد. ویژگی بافت، اطلاعات مهمی درباره وضعیت ساختاری قسمت‌های مختلف یک تصویر و ارتباط آن‌ها با سایر قسمت‌ها در بر دارند. تحقیقات زیادی در زمینه تحلیل بافت به علت مفید بودنش در زمینه‌های بازشناسی الگو، بینایی ماشین و بازیابی تصویر انجام شده است. پر واضح است روش‌هایی که بتواند ویژگی‌های موثرتری از تصاویر استخراج کنند، قادر به افزایش دقت سامانه بازیابی در مراحل اولیه بازیابی خواهند بود. هر چند که استفاده از ویژگی‌های سطح پایین به تنهایی در یک سامانه بازیابی کافی نیست و استفاده از کاربر برای درک تصویر مورد نظر او در تعامل

## ۶-۱ پیاده سازی فیلترهای گابور

فیلترهای گابور را می‌توان از طریق گسترش و چرخش فیلتر مادر (رابطه (۱)) با تابع مولد رابطه (۱۸) به دست آورد:

$$\begin{aligned} g_{m,n}(x, y) &= a^{-m} g(x', y') \quad a > 1 \\ x' &= a^{-m} (x \cos \theta + y \sin \theta) \\ y' &= a^{-m} (-x \sin \theta + y \cos \theta) \\ \theta &= \frac{n\pi}{k} \quad n = 0, 1, \dots, k-1 \quad m = 0, 1, \dots, s-1 \end{aligned} \quad (18)$$

در این رابطه  $a$  پارامتر مقیاس،  $k$  تعداد کل جهت‌ها،  $s$  تعداد کل مقیاس‌ها و  $m$  و  $n$  به ترتیب مشخص کننده مقیاس و جهت هر فیلتر هستند. ضریب  $a^{-m}$  برای اطمینان از عدم وابستگی انرژی فیلتر به  $m$  است.

از آنجا که این فیلترها متعامد نیستند، در تصویرهای فیلتر شده مقداری افزونگی وجود دارد. برای کاهش این افزونگی و نزدیک شدن به تعامد، فیلترها طوری طراحی می‌شوند که علاوه بر پوشش کامل صفحه فرکانس توسط پاسخ فرکانسی فیلترها، میزان همپوشانی آنها در حوزه فرکانس کم باشد. ثابت شده است که برای رسیدن به این منظور باید محل تقاطع آنها در صفحه نصف توان باشد. در این صورت چنانچه  $u_h$  و  $u_l$  به ترتیب فرکانس‌های مرکزی حداکثر و حداقل مورد نظر برای فیلترها باشند، پارامترهای فیلتر به فرم روابط (۱۹) تا (۲۱) به دست می‌آیند [۳۳ و ۳۴].

$$a = \left( \frac{u_h}{u_l} \right)^{\frac{1}{s-1}} \quad (19)$$

$$\sigma_u = \frac{(a-1)u_h}{(a+1)\sqrt{2\ln 2}} \quad (20)$$

$$\sigma_v = \frac{\tan\left(\frac{\pi}{2k}\right) \left( u_h - \left(\frac{\sigma_u^2}{u_h}\right) (2\ln 2) \right)}{\sqrt{2\ln 2 - (2\ln 2)^2 \left(\frac{\sigma_u^2}{u_h}\right)}} \quad (21)$$

بنابر پیشنهاد مرجع [۳۵] برای یک تصویر با اندازه  $N \times N$ ، که  $N$  توانی از ۲ است فرکانس‌های مرکزی فیلترهای گابور  $1\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}, \dots, (N/4)\sqrt{2}$  سیکل بر پهنای تصویر انتخاب می‌شوند. این نحوه انتخاب فرکانس‌های مرکزی تضمین می‌کند که باند عبور فیلتر با بزرگترین فرکانس مرکزی، یعنی  $(N/4)\sqrt{2}$  سیکل بر پهنای تصویر، داخل آرایه تصویر قرار بگیرد. باید توجه داشت که از فیلترهای با فرکانس مرکزی بسیار کوچک (مانند  $1\sqrt{2}$  و  $2\sqrt{2}$  سیکل بر پهنای تصویر) صرف نظر می‌شود، زیرا این فیلترها تغییرات بسیار کم را از تصویر استخراج می‌کنند که مربوط به بافت نیستند [۳۵ و ۳۶]. در این تحقیق تعداد کل مقیاس‌ها ۴ انتخاب می‌شود به عبارت دیگر از ۴ فرکانس مرکزی بزرگتر برای فیلترها استفاده می‌گردد [۱۰]. با این

در مرجع [۲۸]، برای طبقه‌بندی ۶ نوع قلم چینی و ۸ نوع قلم انگلیسی، تمام متن به صورت یک تصویر بافت دار در نظر گرفته شده و برای استخراج ویژگی‌ها، تصویر پیش پردازش می‌شود. در پیش پردازش ارتفاع خطوط متن به یک اندازه ثابت نرمالیزه شده و زیرتصویری با ابعاد  $128 \times 128$  از تصویر متن نرمالیزه شده استخراج می‌گردد. ویژگی‌ها، میانگین و انحراف معیار استاندارد تصاویر خروجی از ۱۶ کانال فیلتر گابور در ۴ جهت و ۴ مقیاس می‌باشند که تولید یک بردار ویژگی ۳۲ بعدی می‌کنند. در مرجع [۲۹] برای تشخیص دو نوع قلم کراهی از مقایسه سرفیها استفاده شده است. برای تشخیص استیل‌های زیرخط دار، ایتالیک و پرنگ از ویژگی‌های استخراج شده از پروفیل کلمه استفاده شده است. در مرجع [۳۰] از ویژگی‌های گشتاورهای مرتبه سوم و چهارم به عنوان ویژگی‌های فرامحلی بافت برای طبقه‌بندی ۸ نوع قلم اسپانیایی در ۴ استیل مختلف استفاده شده است. در این مقاله تصویر خاکستری متن در یک بردار یک بعدی چیده شده، ۱۶ گشتاور مرتبه سوم و چهارم در ۴ جهت ۰ و ۴۵ و ۹۰ و ۱۳۵ درجه از هر تصویر استخراج شده است.

در زمینه بازشناسی قلم‌های فارسی تحقیقات کمتری انجام گرفته است. در مرجع [۳۱] روشی برای خوشه‌بندی ۲۳ قلم متداول فارسی ارائه شده است. در مرجع [۱۰] برای طبقه‌بندی ۲۰ قلم متداول فارسی از ۲۴ کانال فیلتر گابور در ۴ مقیاس و ۶ جهت استفاده شده است. در مرجع [۲۴] تصویر کل متن به عنوان یک تصویر بافت‌دار در نظر گرفته شده است که بافت متن وابسته به قلم می‌باشد. ضرایب خود همبستگی به عنوان ویژگی‌های مناسب برای طبقه‌بندی قلم‌های مختلف که ۲۰ قلم متداول فارسی می‌باشد استخراج شده، و با استفاده از طبقه‌بند  $k$  همسایه نزدیکتر طبقه‌بندی قلم‌ها انجام گرفته است. در مرجع [۳۲] مجموعه‌ای از ویژگی‌های جدید مبتنی عملگرهای سوبل و روبرت برای بازشناسی قلم فارسی در سطح کلمه پیشنهاد شده است. در یک سیستم بازشناسی قلم، بعد از دو سطحی‌سازی تصویر ورودی، با توجه به این که از چه رویگری در بازشناسی استفاده می‌شود، پیش پردازش‌هایی خاص صورت می‌گیرد. سپس ویژگی‌های مناسب از تصویر متن استخراج و به طبقه‌بند سپرده می‌شود. در این مقاله از ویژگی‌های مکان فرکانس در رویکرد مبتنی بر صفحه متن برای تشخیص قلم استفاده می‌شود. برای استخراج ویژگی، فیلترهای پیشنهادی و فیلترهای گابور مورد آزمون قرار می‌گیرند.

## ۶-۲ ارزیابی فیلترهای جدید

در این بخش برای ارزیابی فیلترهای معرفی شده در بخش قبل، از آن‌ها جهت استخراج بردار ویژگی در زمینه‌های طبقه‌بندی معنایی تصویر و بازشناسی قلم‌های فارسی استفاده کرده و نتایج را با نتایج حاصل از فیلترهای گابور مقایسه می‌کنیم. ابتدا نحوه پیاده‌سازی فیلترهای گابور و فیلترهای جدید با دوره محدود تشریح می‌شود.

گروه‌های معنایی استفاده شده عبارتند از: مردم، شیرها، ساختمان‌ها، اتوبوس‌ها، داخل خانه، فیل‌ها، گل‌ها، اسب‌ها، کوهستان و غذاها. شکل (۴) از هرگروه معنایی یک تصویر به عنوان مثال به همراه شماره گروه نشان می‌دهد.

### ۶-۳-۲ طبقه بندی و نتایج

قبل از استخراج ویژگی ابتدا تصویر رنگی به خاکستری تبدیل شده و اندازه آن به  $512 \times 512$  نرمالیزه می‌شود. در نهایت ویژگی‌های مناسب از این زیر تصویر استخراج شده و به طبقه بند  $k$  همسایه نزدیکتر سپرده می‌شود. خروجی طبقه‌بند گروه معنایی تصویر ورودی را مشخص می‌کند. روش  $k$  همسایه نزدیکتر، متداول‌ترین روش برای طبقه‌بندی و بازیابی تصویر است. در این روش برای تعیین گروه یک تصویر، باید بردار ویژگی آن با نمایه‌های تمام تصاویر پایگاه تصویر که بر چسب آن‌ها مشخص است، مقایسه شود [۱۷].

در طبقه‌بندی معنایی تصویر به علت کم بودن داده‌ها برای به دست آوردن کارایی طبقه‌بندی از روش "یکی را کنار بگذار" استفاده شده است. در این روش، هر یک از تصاویر پایگاه به نوبت به عنوان تصویر برچسب نخورده به الگوریتم طبقه‌بند ارائه می‌شود و برچسب واقعی تصویر مذکور با برچسبی که الگوریتم طبقه‌بند به آن نسبت داده است مقایسه می‌شود. این آزمایش برای تمام تصاویر انجام می‌شود و نرخ طبقه‌بندی محاسبه می‌گردد.

برای تعیین فاصله بین بردارهای ویژگی از فاصله  $\chi^2$  استفاده شده است. این فاصله برای دو بردار  $n$  بعدی  $\bar{X} = (x_1, x_2, \dots, x_n)$  و  $\bar{Y} = (y_1, y_2, \dots, y_n)$  به صورت رابطه ۲۴ تعریف می‌شود.

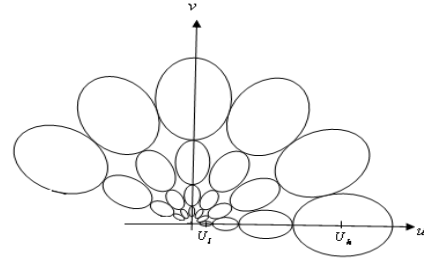
$$\chi^2(\bar{X}, \bar{Y}) = \sum_{i=1}^n \left( \frac{x_i - y_i}{x_i + y_i} \right)^2 \quad (24)$$

در طبقه‌بندی تعداد همسایه‌ها ۷ در نظر گرفته شده است (k=۷) [۳۸]. نتایج حاصل از این دو روش در جدول (۱) خلاصه شده است. با آزمایش، نرخ طبقه‌بندی صحیح ۶۶/۳٪ و زمان اجرای ۱۳۰ دقیقه با هسته جدید به دست آمده که بیانگر افزایش نرخ طبقه‌بندی و کاهش زمان اجرا است که مناسب بودن روش پیشنهادی را تایید می‌کند.

جدول (۱): مقایسه فیلتر گابور و فیلتر پیشنهادی بر اساس نرخ طبقه‌بندی

درست تصاویر و زمان استخراج ویژگی		
درصد طبقه‌بندی صحیح	مدت زمان اجرا (دقیقه)	
۶۲	۱۲۲۰	فیلتر گابور
۶۶/۳	۱۳۰	فیلتر جدید

شیوه انتخاب فرکانس‌های مرکزی و تعداد ۴ مقیاس، مقدار  $u_l$  و  $a$  مشخص می‌شود. به عنوان مثال، برای تصاویر با ابعاد  $512 \times 512$ ،  $u_l = u_0 = 128\sqrt{2}$  و  $u_h = 16\sqrt{2}$  سیکل بر پهنای تصویر و  $a = 2$  خواهد بود. فیلترهای گابور با ۶ جهت مختلف انتخاب می‌شوند  $(k = 6, \theta = 0, 30, 60, 90, 120, 150)$ . به این ترتیب مجموعه فیلتر گابور شامل ۲۴ فیلتر است که به خوبی کل صفحه فرکانسی را پوشش می‌دهند (شکل (۳)). دوره فیلترهای گابور برای پیاده‌سازی در حوزه مکان برابر با  $6\sigma_x \times 6\sigma_y$  انتخاب می‌شود [۱۶].



شکل (۳): مقطع عرضی ۲۴ فیلتر گابور که صفحه فرکانسی را پوشش داده‌اند.

### ۶-۲ پیاده‌سازی فیلترهای چندجمله‌ای مدوله شده با دوره محدود

فیلترهای جدید، از روی فیلتر مادر رابطه (۱۴) توسط تابع مولد (۲۲) ساخته می‌شوند:

$$\phi_{\gamma, \lambda_x, \lambda_y; m, n}(x, y) = a^{-m} \phi_{\gamma, \lambda_x, \lambda_y}(x', y') \quad a > 1 \quad (22)$$

که:

$$\begin{aligned} x' &= a^{-m}(x \cos \theta + y \sin \theta) \\ y' &= a^{-m}(-x \sin \theta + y \cos \theta) \\ \theta &= \frac{n\pi}{k} \quad n = 0, 1, \dots, k-1 \quad m = 0, 1, \dots, s-1 \end{aligned} \quad (23)$$

برای معنادار بودن مقایسه فیلترهای جدید با فیلترهای گابور، پارامترهای مشترک بین فیلترهای جدید و فیلترهای گابور را یکسان انتخاب می‌کنیم. پارامتر  $\gamma = 3$  انتخاب می‌شود. لازم به ذکر است که فیلترهای جدید دارای دوره‌ی محدود و برابر با  $\lambda_x \times \lambda_y$  هستند.

### ۶-۳ طبقه بندی معنایی تصویر

#### ۶-۳-۱ پایگاه داده

پایگاه داده‌ای که مورد استفاده قرار می‌گیرد شامل ۱۰ گروه ۱۰۰ تصویری است. بنابراین، در مجموع دارای ۱۰۰۰ تصویر می‌باشد که از پایگاه داده کورل انتخاب شده‌اند [۳۷]. فرض بر این است که ویژگی‌های معنایی موجود در این پایگاه تصویر مشخص و از پیش تعیین شده‌اند. تمام تصاویر متعلق به یک گروه معنایی دارای ویژگی معنایی واحدی هستند، اگر چه ممکن است ویژگی‌های سطح پایین آنها متفاوت باشند. انتخاب کلاس‌ها براساس ویژگی معنایی آنها است.

## ۴-۶ بازشناسی قلم‌های فارسی

## ۱-۴-۶ پایگاه داده

در این تحقیق از پایگاه داده‌ای شامل ۲۰ قلم متداول فارسی استفاده شده است. نام قلم‌های به کار رفته در این تحقیق به همراه شماره مربوط به گروه آن‌ها و یک زیر تصویر  $512 \times 512$  از هر کدام در شکل (۵) آمده است. برای هر کدام از مجموعه‌های یادگیری و آزمایش ۱۰۰۰ زیر تصویر، شامل ۵۰ زیر تصویر  $512 \times 512$  برای هر قلم و در مجموع ۲۰۰۰ زیرتصویر تهیه شده است. پهنای قلم از پیش تعیین شده که تصاویر متن به آن نرمالیزه می‌شوند، در مجموعه یادگیری و آزمایش ۶ پیکسل در نظر گرفته شده است. تصاویر متن با اندازه قلم‌های ۱۲ و ۱۴ و با چاپگر لیزری HP 1100 چاپ و با درجه تفکیک 300dpi و 400dpi با یک روبشگر HP ScanJet 6300 c به صورت دو سطحی رویش شده‌اند [۱۰].

## ۲-۴-۶ استخراج ویژگی

با توجه آنچه گفته شد، قبل از استخراج ویژگی ابتدا یک مرحله پیش‌پردازش روی تصاویر دو سطحی متن ورودی انجام می‌گردد [۱۰]. سپس یک زیر تصویر با ابعاد  $512 \times 512$  از تصویر پیش پردازش شده، به طور اتفاقی جدا می‌شود و در نهایت ویژگی‌های مناسب از این زیر تصویر استخراج شده و به طبقه‌بند  $k$  همسایه نزدیکتر سپرده می‌شود. خروجی طبقه‌بند نوع قلم تصویر ورودی را مشخص می‌کند. پیش پردازش در چند مرحله انجام می‌شود که عبارتند از: پیدا کردن پهنای قلم تصویر متن ورودی، نرمالیزه کردن تصویر متن به پهنای قلم از پیش تعیین شده، پیدا کردن خطوط متن و فاصله بین آن‌ها و تنظیم فواصل بین آن‌ها به اندازه پهنای قلم از پیش تعیین شده، حذف خطوطی از متن که پهنای آن‌ها کمتر از ثلث پهنای تصویر است، پیدا کردن بخش‌های همپوشان عمودی در تصویر هر خط متن و کاهش فاصله بخش‌هایی که فاصله آن‌ها از دو برابر پهنای قلم بیشتر است و در نهایت پرکردن فضاهای خالی - که در انتهای هر خط متن ایجاد می‌شود - توسط تصویر کلمات ابتدای همان خط [۱۰].

برای هر تصویر  $I(x,y)$  تبدیل موجک گابور و تبدیل موجک هسته جدید به ترتیب به صورت روابط ۲۵ و ۲۶ تعریف می‌شود:

$$w_{mn}(x,y) = \iint I(x_1,y_1) g_{mn}^*(x-x_1,y-y_1) dx_1 dy_1 \quad (25)$$

$$\varphi_{mn}(x,y) = \iint I(x_1,y_1) \phi_{mn,\lambda_x,\lambda_y}^*(x-x_1,y-y_1) dx_1 dy_1 \quad (26)$$

که در آن  $g_{mn}$  و  $\phi_{mn,\lambda_x,\lambda_y}$  به ترتیب معرف هسته گابور و هسته جدید در مقیاس  $m$  و جهت  $n$  هستند و "\*" بیانگر مزدوج مختلط است. تعریف بالا معادل با حاصل ضرب کانولوشن است که مفهوم فیلتر کردن را دارد. بنابراین با توجه به انتخاب ۴ مقیاس و ۶ جهت برای پیاده‌سازی فیلترها، هر تصویر از ۲۴ فیلتر گابور و همچنین ۲۴ فیلتر جدید عبور کرده و ۲۴ تصویر فیلتر شده  $w_{mn}$  و  $\varphi_{mn}$  به دست می‌آید.

از هر تصویر فیلتر شده دو پارامتر میانگین،  $\mu_{mn}$ ، و انحراف معیار،  $\sigma_{mn}$ ، مطابق روابط (۲۷) و (۲۸) به عنوان ویژگی استخراج می‌شوند [۶ و ۳۹].

$$\mu_{mn} = \iint |w_{mn}(x,y)| dx dy \quad (27)$$

$$\sigma_{mn} = \sqrt{\iint (|w_{mn}(x,y)| - \mu_{mn})^2 dx dy} \quad (28)$$

در نهایت برای هر تصویر یک بردار ویژگی با استفاده از  $\mu_{mn}$  و  $\sigma_{mn}$  تشکیل می‌شود که در این تحقیق ۴۸ بعدی و مطابق رابطه (۲۹) است:

$$\mu_{mn} = \iint |w_{mn}(x,y)| dx dy \quad (29)$$

این بردار ویژگی برای طبقه‌بندی و تعیین نوع تصویر ورودی به مرحله طبقه‌بندی سپرده می‌شود.

## ۳-۴-۶ طبقه بندی و نتایج

برای طبقه‌بندی از روش  $k$  همسایه نزدیکتر استفاده می‌کنیم. برای تعیین فاصله بین بردارهای ویژگی از فاصله  $L_2$  طبق استفاده شده است. در طبقه‌بندی تعداد همسایه‌ها ۷ در نظر گرفته شده است [۳۸].

نتایج حاصل از این دو روش در جدول (۲) خلاصه شده است.

با توجه به نتایج، زمان استخراج ویژگی توسط فیلترهای جدید هشتصد و هشتاد و پنج دقیقه بوده در حالی که این زمان برای فیلترهای گابور تقریباً هفت برابر است. همان طور که مشاهده می‌کنیم علاوه بر کاهش چشمگیر مدت زمان پردازش، درصد طبقه‌بندی نیز کمی افزایش داشته است که بیانگر کارایی فیلترهای جدید پیشنهاد شده است.

از بین کل خطای ایجاد شده، میزان ۱/۱۸٪ (۱۸ نمونه تصویر) از ۲/۷٪ (از ۲۷ تصویر) میانگین طبقه‌بندی ناصحیح فیلتر گابور و ۱/۱۳٪ (۱۳ نمونه) از ۲/۵٪ (۲۵ نمونه) میانگین طبقه‌بندی ناصحیح فیلتر جدید در مورد تشخیص نادرست قلم شماره ۲۰ (قلم زر) می‌باشد که با قلم شماره ۱۵ (قلم عربی سنتی) اشتباه شده است. با توجه به شکل (۵) دلیل این امر شباهت زیاد این دو قلم می‌باشد. بنابراین در صورت حذف این قلم نرخ طبقه‌بندی صحیح افزایش می‌یابد. نکته جالب توجه دیگر این است که ۰/۲٪ (دو نمونه تصویر) خطا در تشخیص قلم ۱۵ (عربی سنتی) در هر دو نوع ویژگی ایجاد شده که به اشتباه، قلم شماره ۲۰ (زر) شناسایی شده‌اند. بنابراین در مجموع میزان ۲٪ از خطای فیلترهای گابور و ۱/۵٪ از خطای فیلترهای پیشنهادی به جابجایی در تشخیص قلم‌های زر و عربی سنتی بوده است. حال با توجه به این نکته، چنانچه قلم شماره ۲۰ از مجموعه خارج گردد نرخ طبقه‌بندی صحیح در این سامانه بر مبنای فیلترهای پیشنهادی به ۹۹/۴۷٪ افزایش خواهد یافت. در شکل (۶) نمونه‌هایی از دو قلم زر و عربی سنتی که به اشتباه طبقه‌بندی شده‌اند، ارائه شده است. شباهت این دو قلم در این تصویر به روشنی قابل مشاهده است.





شکل (۴): نمونه‌ای از تصاویر گروه معنایی ده‌گانه در پایگاه تصویر

گروه ۱: مردم	گروه ۲: شیرها	گروه ۳: ساختمانها	گروه ۴: اتوبوسها	گروه ۵: داخل خانه
گروه ۶: فیله‌ها	گروه ۷: گلها	گروه ۸: اسبها	گروه ۹: کوهستان	گروه ۱۰: غذاها
گروه ۱: Arial	گروه ۲: Compset	گروه ۳: Elham	گروه ۴: Farnaz	گروه ۵: Homa
گروه ۶: Jadid	گروه ۷: Kamran	گروه ۸: Koodak	گروه ۹: Lotus	گروه ۱۰: Mitra
گروه ۱۱: Nasim	گروه ۱۲: Nazanin	گروه ۱۳: Roya	گروه ۱۴: Sina	گروه ۱۵: Traditional Arabic
گروه ۱۶: Titr	گروه ۱۷: Trafic	گروه ۱۸: UN	گروه ۱۹: Yagut	گروه ۲۰: Zar

شکل (۵): نام قلم‌های به کار رفته در این آزمایش به همراه شماره گروه و یک زیر تصویر ۵۱۲×۵۱۲ از هر کدام

فیلترهایی با کارایی قوی‌تر از گابور یافت می‌شود. به عنوان مثال خسروی و کبیر [۳۲] از فیلترهای سوبل و روبرت به منظور استخراج ویژگی در کاربرد تشخیص قلم مبتنی بر خطوط متن استفاده کرده‌اند. در این مقاله نویسندگان ابتدا یک خط از متن را جدا و پس از نرمالیزه کردن با اعمال فیلترهای فوق‌الذکر در چهار جهت به ۱۶ زیر بلوک از

لازم به ذکر است که در این مقاله ما به دنبال اثبات کارایی فیلترهای گابور نیستیم. زیرا نقش فیلترهای گابور به عنوان ابزاری قوی و کارا در تشخیص الگو تثبیت شده است. بلکه به دنبال افزایش سرعت این فیلترها با حفظ کارایی آن‌ها هستیم. فیلترهای گابور همانند هر ابزار دیگر دارای محاسن و محدودیت‌هایی هستند و در بعضی کاربردها،

اطلاعات و از دست رفتن خواص این فیلترها می‌شود. در این تحقیق برای رفع این مشکل، موجک مدوله شده جدیدی منتج از هسته‌های چندجمله‌ای با دوره محدود ارائه و بر اساس آن فیلترهای جدیدی با خصوصیات شبیه به فیلترهای گابور معرفی شده‌اند که دارای دوره محدود بوده و به راحتی قابل پیاده‌سازی هستند. سپس فیلترهای جدید در دو کاربرد متفاوت طبقه‌بندی شامل طبقه‌بندی معنایی تصویر و بازشناسی قلم فارسی به کار گرفته شدند. نتایج آزمایش‌ها بیانگر بهبود زمان استخراج ویژگی بین ۷ تا ۱۰ برابر در روش پیشنهادی نسبت به فیلترهای گابور با حفظ کارایی است.

تصویر متن اقدام به استخراج ویژگی به تشخیص قلم فارسی می‌نمایند. نشان داده شده است که برای چنین کاربردی و تصاویر ورودی با درجه تفکیک 100dpi، این فیلترها قادر به تشخیص قلم با نرخ متوسط ۹۴/۲٪ می‌باشند که در مقایسه با فیلترهای گابور از کارایی و سرعت بسیار بهتری برخوردار هستند [۳۲].

### ۷- نتیجه‌گیری

فیلترهای گابور به علت خواص منحصر به فرد، کاربردهای بسیار زیادی در زمینه‌های مختلف پردازش تصویر پیدا کرده‌اند. اما به دلیل این که دوره آن‌ها نامحدود است و در حوزه فرکانس نیز دارای تابع تبدیل کسری نمی‌باشند، به طور دقیق قابل پیاده‌سازی نیستند. بنابراین، در پیاده‌سازی آن‌ها از تقریب استفاده می‌شود که این امر سبب اتلاف

جدول (۲): مقایسه فیلتر جدید و فیلتر گابور بر اساس نرخ طبقه‌بندی صحیح قلم ورودی و زمان استخراج ویژگی

شماره نوع قلم تصویر	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
درصد طبقه بندی صحیح فیلتر گابور	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۹	۹۶	۱۰۰
درصد طبقه بندی صحیح فیلتر جدید	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۱۰	۹	۱۰۰	۱۰۰
شماره نوع قلم تصویر	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	میانگین	زمان اجرا (دقیقه)
درصد طبقه بندی صحیح فیلتر گابور	۱۰	۱۰	۱۰	۹۶	۱۰	۱۰	۱۰	۱۰	۵	۹۷/۳	۵۷۷۲
درصد طبقه بندی صحیح فیلتر جدید	۱۰	۱۰	۱۰	۹۶	۱۰	۱۰	۱۰	۹۸	۶	۹۷/۵	۸۸۵

به تعدادی از آنها اشاره می‌رژمی مناسب در شبکه در یک تحقیق نقاط بیشینه با با یک شبکه عصبی معجزه‌آسا اصلی با هم ترکیب صورت گرفته است.	ی شبکه از تابع انرژی استخراج شده و ناحیه بندی در، ساویر قرمز، سبز و آبی برای که هاپفیلد با نورون برای جاب شده بعد از آنالیز هیستو	به تعدادی از آنها اشاره می‌جدا سازی تصاویر خاکستر شده است. وزنهای شبکه از سبز و آبی استخراج شده و ج ناحیه بندی تصاویر قرمز،	ی عصبی از عناصر محاسباتی که کامل هر درجه ای از غیر بدل کرد. زمان آموزش در بکه های عصبی مصنوعی ن که در ادامه به تعدادی از
(الف)	(ب)	(ج)	(د)
اهش زمان محاسبات در داده استفاده شده است. در این اجرا ناحیه بندی با ترکیب دورا نظم با بعد کم، تصویر می‌اده کرد. در مرجع کنند؛ که یاله جدا سازی روش از حد	ی با کمینه کردن یک تابع اابع انرژی استخراج می‌شوا ناحیه بندی در هر یک از آ، سبز و آبی برای تشکیل تصورون برای جدا سازی رن	تصویر و تعداد حوسه ها دند، مساله آنالیز هیستو مشخص می‌شود. در تخینه محلی پیدا می‌کند. در راه حل کلی مناسب از یه و لبه پای انجام می‌شود	شبکه عصبی در یک تحقیق شینه هیستوگرامهای هاپی گیرد. نتایج قرمز، سبز ه دیگری از یک ناحیه بندهای، نگ انتخاب شبکه
(ح)	(ز)	(ت)	(ی)

شکل (۶): نمونه‌ای از تصاویر مجموعه آزمایش که به اشتباه شده‌اند. تصاویر (الف، ب، ج، د، ه، ز) متعلق به قلم زر هستند که به اشتباه به عنوان قلم عربی سنتی شناسایی شده‌اند. تصاویر (ت، ی) متعلق به قلم عربی سنتی هستند که به اشتباه به عنوان قلم زر شناسایی شده‌اند. نتایج ارائه شده در این شکل بر مبنای فیلترهای پیشنهادی حاصل شده‌اند.

## سپاسگزاری

این تحقیق از پشتیبانی مالی مرکز تحقیقات مخابرات ایران در چارچوب قرار داد شماره ۵۰۰/۴۸۲۴/ت مورخ ۱۳۸۸/۳/۲۷ بین دانشگاه شهید باهنر کرمان و آن مرکز محترم برخوردار بوده است.

## مراجع

- [13] G. Deng and L. W. Cahill, "An Adaptive Gaussian Filter For Noise Reduction and Edge Detection", IEEE Conference Record. Nuclear Science Symposium and Medical Imaging, vol. 3, pp. 1615-1619, 1994.
- [14] S. Saryazdi and M. Cheriet, "PKCS:a plomial kernel family with compact support for scale space image processing", IEEE Transaction on Image Processing, vol. 16, pp. 2299-2308, 2007.
- [15] I. T. Young and L. J. van Vliet, "Recursive implementation of the Gaussian filter", Elsevier, Signal Processing 44, pp. 139-151, 1995.
- [16] L. Remaki, M. Cheriet, "KCS-new kernel family with compact support in scale space:formulation and impact", IEEE Transaction on Image Processing, vol. 9, no. 6, pp. 970-981, 2000.
- [۱۷] ح. نظام‌آبادی‌پور، "کاربرد ویژگی‌های وابسته به نوع در یک سامانه بازیابی تصویر رنگی"، رساله دکتری، دانشگاه تربیت مدرس، بخش مهندسی برق، تابستان ۱۳۸۳.
- [18] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions and open issues", Journal of Visual Communication and Image Representation, vol. 10, pp. 39-62, 1999.
- [19] S. K. Bhattacharjee, "A computational approach to image retrieval", M. Sc. thesis, Michigan state university, USA, 1999.
- [20] Y. Li, X. Wan and C. C. J. Kuo, "Introduction to content-based image retrieval-overview of key techniques", in Image Database: Search and Retrieval of Digital Imagery, Edited by Bergman and Castelli, John Wiley & Sons, 2002.
- [21] J. P. Eakins, "Toward intelligent image retrieval", Pattern Recognition, vol. 35, pp. 3-14, 2002.
- [22] A. Vailaya, A. T. Figueiredo, A. K. Jain and H. J. Zhang, "Content-based hierarchical classification of vacation images", IEEE Int. Conf. on Multimedia Computing and Systems, ICMCS'99, Italy, pp. 518-523, 1999.
- [23] R. Zhang and Z. Zhang, "A clustering based approach to efficient image retrieval", In Proceedings of the 14th IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI'02), pp. 339-346, 2002.
- [۲۴] ع. راشدی، ح. نظام آبادی پور، س. سریزدی، "بازشناسی قلم‌های فارسی به کمک ضرایب خود هبستگی"، چهارمین کنفرانس پردازش تصویر و بینایی ماشین ایران، جلد اول، ص. ۴۸ تا ۵۶، مشهد، ایران، ۱۳۸۵.
- [25] A. Zramidni and R. Ingold, "Optical font recognition fram projection profiles", Electronic Publishing, vol. 6, no. 3, pp. 249-260, 1993.
- [26] A. Zramidni and R. Ingold, "Optical font recognition using typographical features", IEEE Trasactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 877-882, 1998.
- [1] D. Gabor, "Theory of communication" Proc. IEE, vol. 93, no. 26, pp. 429-457, 1946.
- [2] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters", J. Opt. Soc. Amer. A, vol. 2, Issue.7, pp.1160-1169, 1985.
- [3] G. M. Haley and B. S. Manjunath, "Rotation invariant texture classification using the modified Gabor filters", IEEE ICIP.95, pp.262-265, 1995.
- [4] An P. N. Vo, T. T. Nguyen and S. Oraintara, "Texture Image Retrieval Using Complex Directional Filter Bank", IEEE, pp. 5495-5498, 2006.
- [5] D. A. Clausi and M. Ed Jernigan, "Designing Gabor filters for optimal texture separability", Elsevier, Pattern Recognition 33, pp. 1835-1849, 2000.
- [6] A. Ahmadian, A. Mostafa, M. D. Abolhassani and N. Riahi Alam, "A method for texture classification of ultrasonic liver images based on Gabor Wavelet", IEEE signal processing, vol. 2, pp.971-974, 2004.
- [7] J. Han and K- Kuang Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval", Elsevier, Image and Vision Computing, vol. 25, Issue. 9, pp. 1474-1481, 2007.
- [8] B. Allier and H. Emptoz, "Font type extraction and character prototyping using gabor filters", IEEE ICDAR, pp. 799-803, 2003.
- [9] L. Tang, L. F. Tian, B. L. Steward and J. F. Reid, "Texture-Based Weed Classification Using Gabor Wavelets and Neural Network for Real-time Selective Herbicide Applications", Department of Agricultural Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, 1999.
- [۱۰] ح. نظام‌آبادی‌پور، س. سریزدی و ا. ابراهیمی، "بازشناسی قلم‌های فارسی بر پایه فیلترهای گابور"، نهمین کنفرانس انجمن کامپیوتر ایران، ص. ۳۷۱ تا ۳۷۸، ۱۳۸۲.
- [11] A. P. Witkin, "Scale-Space filtering", in Proc.8<sup>th</sup> Int. Joint Conf. Artificial Intelligence, Karlsruhe, Germany, pp. 14-45, 1983.
- [12] A. A. Farag, Y. Cao, D. M. Rose and E. J. Delp, "On Empirica Estimation of the Parameters of Edge Enhancement Filters", IEEE International Conference on Systems, Man and Cybernetics, vol. 1, pp. 346-350, 1992.

## زیرنویس‌ها

- 1- Rational
- 2- Recursive
- 3- Polynomial Kernel With Compact Support (PKCS)
- 4- Kernel with Compact Support (KCS)
- 5- Unit ball
- 6- Content based image retrieval (CBIR)
- 7- Query Image
- 8- Low Level Feature
- 9- General features
- 10- Domain specific features
- 11- Spatial relation
- 12- Ascender zone
- 13- Descender zone
- 14- Height zone

- [27] M. C. Jung , Y. C. Shin and S. N. Srihari, "Multifont classification using typographical attributes", IEEE Computer Society Washington, DC, USA, 1999.
- [28] Y. Zhu, T. Tan and Y. Wang, "Font recognition based on global texture analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 10, pp. 1192-1200, 2001.
- [29] S. H. Kim, H. K. Kwag and C. Y. Suen, "Word level optical font recognition using typographical features", Pattern Recognition, vol. 18, no. 4, pp. 541-561, 2004.
- [30] C. A. Cruz, R. R. Kuoppa, A. A. Gonzalez, and R. E. Perez, "High order statistical texture analysis-font recognition applied", Pattern Recognition Letters, vol. 26, vol. 2, pp. 135-145, 2005.

[۳۱] ا. ابراهیمی، ا. کبیر، ک. لوکس، "انتخاب ویژگی با الگوریتم وراثتی برای خوشه‌بندی قلم‌های فارسی با FCM"، اولین کنفرانس بین‌المللی فناوری اطلاعات و دانش، ص. ۲۷۴ تا ۲۷۹، دی ۸۲.

- [32] H. Khosravi and E. Kabir, "Farsi font recognition based on Sobel-Roberts features", Pattern Recognition Letters, vol. 31, no. 1, pp. 75-82, 2010.
- [33] B. S. Manjunath and W. Y. Ma, "Texture feature for browsing and retrieval of image data", IEEE PAMI, no. 8, vol. 18, pp. 837-842, 1996.
- [34] J. Han and K- Kuang Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval", Image and Vision Computing, vol. 25, no. 9, pp. 1474-1481, 2007.
- [35] A. K. Jain and S. K. Bhattacharjee, "Address block location on envelopes using Gabor filters", Pattern Recognition, vol. 25, no. 12, pp. 1449-1477, 1992.
- [36] K. Hammouda and P. E. Jernigan, "Texture segmentation using gabor filters", Tech. Rep., University of Waterloo, 2003.
- [37] J. Z. Wang, J. Li and G. Wiederhold, "SIMPLicity: semantic sensitive integrated matching for picture libraries", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 23, no. 9, pp. 947-963, 2001.

[۳۸] ح. نظام‌آبادی‌پور و ا. کبیر، "ارزیابی معیارهای عدم شباهت در بازیابی و طبقه‌بندی تصویر"، مجله علمی پژوهشی فنی مدرس، شماره ۲۲، صفحات ۸۹-۹۸، ۱۳۸۴.

- [39] M. Nabti and A. Bouridane, "An effective and fast iris recognition system based on a combined multiscale feature extraction technique", Pattern Recognition 41, pp. 868-879, 2008.