



ترکیب ویژگیهای سطح پایین برای طبقه بندی معنایی تصاویر

حسین نظام‌آبادی پور^۱ و احسان‌اله کبیر^۱

^۱ گروه الکترونیک، بخش مهندسی برق، دانشگاه تربیت مدرس، تهران، ایران
^۲ هسته پژوهشی پردازش تصویر، بخش مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران

چکیده

طبقه بندی تصاویر به گروههای معنایی بر اساس ویژگیهای دیداری سطح پایین، یکی از چالشهای مهم در بازیابی و طبقه بندی تصویر است. در این مقاله، اثر ترکیب ویژگیهای رنگ، شکل و بافت و نیز وزن دار کردن مولفه‌های بردارهای ویژگی در طبقه بندی تصویر بررسی شده است. همچنین اثر تعداد تصاویر پایگاه داده بر نرخ طبقه بندی مطالعه شده است. برای این منظور از یک پایگاه تصویر شامل ۱۰ گروه معنایی، هر گروه شامل ۱۰۰ تصویر، استفاده شده است. برای طبقه بندی از روش k همسایه نزدیکتر و برای تخمین نرخ آن از روش "یکی را کنار بگذار" استفاده شده است. تعیین وزن مناسب برای هر یک از مولفه‌های یک نوع ویژگی و نیز تعیین وزن‌های بهینه در ترکیب انواع متفاوت ویژگیها با الگوریتم وراثتی انجام شده است.

کلمات کلیدی: بازیابی تصویر بر اساس محتوی، ویژگی سطح پایین، ویژگی معنایی، ترکیب ویژگیها، طبقه بندی تصویر، الگوریتم وراثتی

۱- مقدمه

در سالهای اخیر برای ایجاد روشهایی کارآمد در بازیابی تصویر بر پایه محتویات^۱ آن تلاشهای زیادی شده است [۲ و ۱]. این تلاشها منجر به ساخت سیستمهای متفاوتی در بازیابی تصویر شده است [۳]. اساس کار این سیستمها بر این استوار است که تصاویر را بر پایه مجموعه‌ای از ویژگیهای دیداری سطح پایین^۲ آنها مثل شکل، بافت، رنگ و چیدمان رنگ^۴ ارائه و نمایه سازی کنند. بازیابی با مقایسه ویژگیهای تصویر پرس و جو با نمایه‌های تصاویر پایگاه و تعیین نزدیکترین تصاویر انجام می‌شود. عموماً کاربران هنگام پرس و جوی تصویر به ویژگیهای سطح پایین تصویر فکر نکرده و بدنبال ویژگیهای سطح بالا^۵ یا همان ویژگیهای معنایی^۶ هستند. در این موارد اغلب سیستمهای بازیابی تصویر، عملکرد ضعیفی دارند. این ضعف بخاطر فاصله‌ای است که بین ویژگیهای دیداری سطح پایین و ویژگیهای معنایی سطح بالا وجود دارد و به شکاف معنایی^۷ مشهور است [۲]. برای رفع این مشکل، محققان بر آن شده‌اند که از بازخورد ارتباط و طبقه بندی معنایی تصاویر استفاده کنند [۴ و ۵].

در استفاده از بازخورد ارتباط، کاربر و کامپیوتر تا رسیدن به تصاویر تعامل دارند. این تعامل باعث می‌شود سیستم رفته رفته سلیقه کاربر را پیدا کند و تصاویر مورد علاقه او را مشخص کند.

استفاده از طبقه بندی معنایی تصاویر پایگاه داده نیز می‌تواند بطور قابل توجهی بازدهی سیستم بازیابی را افزایش دهد. برای حل این مشکل باید آزمون پذیری را به صورت یکی از معیارهای طراحی سطح بالا در نظر گرفت. دسته بندی تصاویر به گروههای معنایی با استفاده از ویژگیهای دیداری سطح پایین مسئله‌ای حاد در بازیابی تصویر است [۶]. برای روشن شدن سختی کار و مشاهده فاصله زیاد بین ویژگیهای دیداری و ویژگیهای معنایی می‌توان از شکل ۱ کمک گرفت. همانگونه که در این شکل دیده می‌شود، تصاویر الف، ب، ج و دال، دارای معنای مشترک "مردم" هستند، در حالیکه ویژگیهای دیداری آنها بسیار متفاوت هستند. همچنین می‌توان مثالهایی ارائه کرد که دارای ویژگیهای دیداری نزدیک هم بوده در حالیکه معنای آنها با یکدیگر متفاوت است. 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 [۷].

در بعضی از تحقیقات [۸ و ۹ و ۱۰]، عمل طبقه بندی با استفاده از بازخورد ارتباط و اطلاعات کاربر انجام شده است.

چهار گروه بافتی، غیر بافتی، گراف و عکس تقسیم می‌شوند [۱۳]. در هر یک از این گروهها برای نمایه سازی از ویژگیهای خاصی استفاده می‌شود. در این روش ابتدا گروه تصویر پرس و جو تعیین شده، سپس ویژگیهای مناسب از آن استخراج می‌شوند و بازیابی در گروه مربوطه انجام می‌شود.

در مرجع [۱۵]، برای طبقه بندی معنایی تصاویر به چهار گروه داخل خانه، مناظر شهری، مناظر طبیعی دور و نزدیک، از توزیع فرامحلی جهت‌های غالب محلی^۱ استفاده شده است. در این تحقیق از طبقه بند k همسایه نزدیکتر استفاده شده و پایگاه شامل ۴۷۰ تصویر خاکستری است. بهترین نرخ طبقه بندی برای چهار کلاس ۸۸/۷٪ گزارش شده است.

مرجع [۱۶] با استفاده از اطلاعات آماری مربوط به چیدمان رنگ تصاویر، سعی در طبقه بندی معنایی تصویر دارد. پایگاه تصویر استفاده شده در این تحقیق شامل ۳۵۷ تصویر از ۱۰ گروه معنایی متفاوت بوده که ۹۱ تصویر برای آموزش و ۲۶۶ تصویر برای آزمایش بکار رفته است. نرخ طبقه بندی با استفاده از روش پیشنهادی در این مرجع ۷۰/۷٪ گزارش شده است.

با مطالعه تحقیقات انجام شده در می‌یابیم که نکات مبهم بسیاری در زمینه طبقه بندی معنایی تصویر وجود دارد. سوالات فراوانی که پاسخ به هر کدام از آنها، نیازمند یک تحقیق جامع و مفصل است. همانگونه که ملاحظه شد، در هیچ یک از تحقیقات ذکر شده به نحوه ترکیب انواع ویژگیها در طبقه بندی اشاره نشده یا اینکه مورد توجه نبوده است. تعداد تصاویر پایگاه داده و تاثیر آن بر نرخ طبقه بندی نیز مطالعه نشده است. همچنین انتخاب مولفه‌های ویژگی و وزن دار کردن آنها و تاثیر آن بر نرخ طبقه بندی تا کنون مطالعه نشده است. در تحقیقات انجام شده در این زمینه، گروه‌های معنایی انتخاب شده دارای مفاهیم گسترده‌ای بوده‌اند (مثل گروه‌های بافتی و غیر بافتی [۱۳]، گروه‌های داخل خانه و بیرون خانه [۱۱] و گروه‌های مناظر شهری و مناظر طبیعی [۱۲]). بررسی گروه‌هایی با مفاهیم معنایی دقیق‌تر مثل شیرها، پلنگ‌ها و اسب‌ها که ویژگیهای سطح پایین آنها به یکدیگر نزدیک هستند یا گروه‌های بسیار متنوع مثل مردم چندان مورد توجه نبوده است. هدف این تحقیق، بررسی اثر ترکیب ویژگیهای رنگ، شکل و بافت و نیز وزن دار کردن مولفه‌های ویژگی در طبقه بندی تصویر است. همچنین اثر تعداد تصاویر پایگاه داده بر نرخ طبقه بندی بررسی شده است. برای این منظور از یک پایگاه تصویر شامل ۱۰ گروه معنایی، هر گروه شامل ۱۰۰ تصویر، استفاده شده است. گروه‌های معنایی بر اساس مفاهیم دقیق‌تر نسبت به تحقیقات مشابه انتخاب شده‌اند. سعی شده است تا گروه‌ها به گونه‌ای باشند که بعضی از آنها از نظر ویژگیهای سطح پایین به یکدیگر نزدیک و بعضی دارای تنوع بسیار در ویژگیهای سطح پایین باشند. برای طبقه بندی از روش k همسایه نزدیکتر استفاده شده است. برای تخمین نرخ طبقه بندی، روش "یکی را کنار بگذار" بکار گرفته شده است. تعیین وزن مناسب برای هر یک از مولفه‌های یک نوع ویژگی و نیز تعیین وزن‌های بهینه در ترکیب انواع متفاوت ویژگیها با الگوریتم وراثتی انجام شده است.

در بخش بعدی، پایگاه تصویر بکار رفته در این تحقیق معرفی خواهد شد. برای طبقه بندی از ترکیب سه نوع ویژگی متفاوت استفاده شده است که در بخش سوم معرفی می‌شوند. نحوه بکار گیری الگوریتم وراثتی برای ترکیب انواع ویژگیها و وزن دار کردن مولفه‌های ویژگی در بخش چهارم آمده است. روش طبقه بندی و نتایج آزمایش در بخش پنجم ارائه می‌شود. در بخش آخر جمع بندی این مقاله آمده است.

۲- پایگاه تصویر

از یک پایگاه شامل ۱۰۰۰ تصویر استفاده شده است. فرض شده است که ویژگیهای معنایی موجود در پایگاه تصویر مشخص و از پیش تعیین شده اند. این پایگاه از ۱۰ کلاس ۱۰۰ تصویری تشکیل شده است. تمام تصاویر متعلق به یک

در این روش عموماً برای هر تصویر، یک یا چند کلاس معنایی پیش فرض در نظر گرفته می‌شود. این کلاس پیش فرض، با توجه به نظرات کاربران در جریان فرآیند بازیابی تصحیح می‌شود. در زمینه طبقه بندی تصویر تحقیقات قابل توجهی انجام شده است که در ادامه به تحقیقاتی که به کار گزارش شده در این مقاله نزدیک است، اشاره می‌شود.

در یک تحقیق، ویژگیهای هیستوگرام رنگ، ویژگیهای MSAR^۸، ویژگیهای استخراج شده از حوزه DCT و نیز ترکیب این ویژگیها در طبقه بندی دو گروه از تصاویر داخل خانه و بیرون خانه با یکدیگر مقایسه شده‌اند [۱۱]. البته به چگونگی ترکیب ویژگیها اشاره نشده است. برای طبقه بندی از k همسایه نزدیکتر و روش "یکی را کنار بگذار" استفاده شده است. در این تحقیق، از یک پایگاه شامل ۱۳۴۶ تصویر استفاده شده است و نرخ طبقه بندی حدود ۹۰٪ گزارش شده است.

در تحقیق مشابهی، دسته بندی تصاویر به گروه‌های معنایی با استفاده از ویژگیهای سطح پایین آزموده شده است [۱۲]. در این تحقیق نیز از طبقه بند k همسایه نزدیکتر با روش "یکی را کنار بگذار" استفاده شده است. تصاویر پایگاه داده به دو گروه شهر و منظره طبیعی طبقه بندی می‌شوند. همچنین کارآیی چندین ویژگی برای طبقه بندی با یکدیگر مقایسه شده است. در این تحقیق از یک پایگاه شامل ۲۷۱۶ تصویر استفاده شده و بیشترین نرخ طبقه بندی صحیح ۹۳/۹٪ گزارش شده است.



(ب)



(الف)



(د)



(ج)

شکل ۱ - چهار تصویر نمونه از گروه معنایی مردم. با اینکه این تصاویر متعلق به یک گروه معنایی هستند، ویژگیهای دیداری آنها تفاوت زیادی دارد

در تحقیق دیگری، از طبقه بند بیز باینری^۹ برای طبقه بندی سلسله مراتبی تصاویر به گروه‌های معنایی استفاده شده است [۱۶]. برای دسته بندی تصاویر به زیرگروهها از ویژگیهای متناسب با خصوصیات آن گروه استفاده شده است. در این تحقیق، تصاویر ابتدا به دو گروه داخل خانه و بیرون خانه دسته بندی می‌شوند. تصاویر بیرون خانه به دو گروه شهر و مناظر طبیعی و تصاویر مناظر طبیعی به سه گروه غروب آفتاب، کوهستان و جنگل طبقه بندی می‌شوند. در کار دیگری، تصاویر به

۳- توصیف ویژگیها

در این مطالعه از سه نوع ویژگی رنگ، شکل و بافت استفاده شده است. از آنجایی که هدف این تحقیق بررسی ترکیب ویژگیها و در نظر گرفتن اثر وزن دار کردن مولفه‌های ویژگی بوده است؛ در انتخاب نوع ویژگیها حساسیت خاصی اعمال نشده است. به همین دلیل سعی شده است از ویژگیهایی استفاده شود که در بسیاری از سیستمهای بازیابی تصویر معمول هستند. از هیستوگرام رنگ و هیستوگرام جهت لبه‌ها به ترتیب برای ویژگیهای رنگ و شکل و از فیلترهای گابور برای ویژگی بافت استفاده شده است. در ادامه به توصیف این ویژگیها می‌پردازیم.

۱-۳ هیستوگرام رنگ

رنگ ویژگی‌ای است که اغلب محققان در بازیابی تصویر از آن استفاده کرده‌اند و در موارد متعدد هیستوگرام رنگ به عنوان اصلی‌ترین ویژگی در بازیابی مورد توجه بوده است. فوایدی همچون پایداری، موثر بودن، سادگی پیاده سازی، سادگی محاسباتی و حجم پایین بردار حاصل برای نمایه سازی، استفاده از هیستوگرام رنگ را توجیه می‌کند.

فضای رنگ RGB^{۱۳} با سیستم بینایی انسان و روشی که انسان تصویر را درک می‌کند، همخوانی ندارد و برای نمایش تصویر توسط مانیتور و تلویزیون مناسب است. در این فضا تفاوت دو نقطه با توجه به مختصاتشان، متناسب با درک انسان نیست و فاصله اقلیدسی، تفاوت بین دو رنگ را به شکلی که برای انسان معنی‌دار باشد، بیان نمی‌کند. ما در این تحقیق از فضای رنگ HSV^{۱۴} استفاده کرده‌ایم. از این فضای رنگ در تحقیقات مشابه به دفعات استفاده شده است [۱۷-۲۳]. در این تحقیق، برای تشکیل هیستوگرام رنگ، فضای HSV بطور خطی کوانتیزه می‌شود. در این فضا بنا بر اهمیتی که مولفه H نسبت به سایر مولفه‌ها دارد، این مولفه به شش بازه و دو مولفه دیگر هر کدام به سه بازه کوانتیزه می‌شوند [۱۷-۱۹]. بنابراین فضای رنگ به $6 \times 3 \times 3 = 54$ بازه کوانتیزه می‌شود. بعد از کوانتیزه کردن رنگ نقاط یک تصویر، هیستوگرام رنگ آن با شمردن نقاطی که در هر بازه قرار می‌گیرند محاسبه می‌شود. برای حل مسئله تفاوت اندازه تصاویر، هیستوگرام رنگ به تعداد کل نقاط آن تصویر نرمالیزه می‌شود. بردار حاصل از نمایه سازی هر تصویر با استفاده از ویژگی رنگ، ۵۴ بعدی است.

۲-۳ هیستوگرام جهت لبه‌ها

بسیاری از محققان برای ارائه ویژگی شکل یک تصویر از هیستوگرام جهت لبه‌های آن تصویر استفاده کرده‌اند [۲۴]. علت انتخاب این ویژگی، توانایی لبه‌ها در ارائه اطلاعات مفید برای نمایش شکل است. در این تحقیق زاویه گرادیان برای هر نقطه لبه در تصویر خاکستری با استفاده از لبه‌یاب سیل محاسبه شده و هیستوگرام جهت لبه با ۳۶۰ عنصر تشکیل می‌شود. این هیستوگرام بعد از هموارسازی و کوانتیزاسیون خطی به ۳۶ عنصر، به تعداد نقاط لبه در تصویر نرمالیزه می‌شود. بنابراین بردار حاصل از نمایه سازی هر تصویر با استفاده از ویژگی راستای لبه‌ها، ۳۶ بعدی است.

۳-۳ ویژگی بافت

فیلترهای گابور به علت خاصیت تفکیک پذیری خوب و چندگانگی در حوزه‌های مکان و فرکانس به عنوان ابزاری قوی در تحلیل بافت استفاده می‌شوند [۱۴، ۲۵-۳۰]. تابع یک فیلتر دو بعدی گابور متقارن زوج و تبدیل فوریه آن مطابق روابط ۱ و ۲ است.

$$f(x, y) = \frac{1}{2\pi s_x s_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{s_x^2} + \frac{y^2}{s_y^2}\right)\right) \cos(2\pi u_0 x) \quad (1)$$

کلاس معنایی دارای ویژگی معنایی واحد هستند، اگرچه ممکن است ویژگیهای سطح پایین آنها متفاوت باشند. انتخاب کلاسها بر اساس ویژگی معنایی آنها است. تصاویر این پایگاه در ابعاد 256×384 یا 384×256 و به فرمت JPEG هستند. این تصاویر از پایگاه تصویر کورل^{۱۱} و از مجموعه تصاویر پروژه سیمپلیستی^{۱۲} انتخاب شده‌اند [۱۳]. گروههای معنایی استفاده شده عبارتند از: مردم، شیرها، ساختمانها، اتوبوسها، داخل خانه، فیله‌ها، گلها، اسبها، کوهستان و غذاها. شکل ۲ از هر گروه معنایی یک تصویر را به عنوان مثال به همراه شماره گروه نشان می‌دهد. گروه‌های مردم، ساختمانها، کوهستان و غذاها دارای تنوع بسیار در ویژگیهای سطح پایین هستند، در حالیکه گروههای شیرها، فیله‌ها و اسبها ویژگیهای سطح پایین نزدیک بهم دارند.

	
گروه ۲: شیرها	گروه ۱: مردم
	
گروه ۴: اتوبوسها	گروه ۳: ساختمانها
	
گروه ۶: فیله‌ها	گروه ۵: داخل خانه
	
گروه ۸: اسبها	گروه ۷: گلها
	
گروه ۱۰: غذاها	گروه ۹: کوهستان

شکل ۲- نمونه‌ای از تصاویر گروههای معنایی ده‌گانه در پایگاه تصویر

که در آن $I(u,v)$ تبدیل فوریه تصویر $I(x,y)$ و $F_{mn}(u,v)$ تبدیل فوریه $f_{mn}(x,y)$ است. برای حذف حساسیت فیلترها به مقدار قدر مطلق شدت روشنایی تصویر، $F_{mn}(0,0)$ برابر صفر در نظر گرفته شده است.

در این تحقیق از چهار مقیاس و شش جهت ($k=4$ و $s=6$)، در مجموع از ۲۴ فیلتر استفاده شده است. این فیلترها در ابعاد 256×256 پیکسل طراحی شده‌اند. در مرحله پیش پردازش، تصویر رنگی به خاکستری تبدیل و به ابعاد 256×256 نرمالیزه می‌شود. سپس زیر-تصویرهای فیلتر شده آن از رابطه ۶ محاسبه می‌شوند. از هر تصویر I_{mn} دو پارامتر میانگین، m_{mn} ، و انحراف معیار، s_{mn} ، مطابق رابطه ۷ به عنوان ویژگی استخراج می‌شوند [۲۵].

$$m_{mn} = \iint |f_{mn}(x,y)| dx dy \quad (7)$$

$$s_{mn} = \sqrt{\iint (|f_{mn}(x,y)| - m_{mn})^2 dx dy}$$

در نهایت یک بردار ویژگی با استفاده از m_{mn} و s_{mn} ساخته می‌شود که در این تحقیق ۴۸ بعدی و مطابق رابطه ۸ است.

$$\bar{f} = [m_{00} \ s_{00} \ m_{01} \ \dots \ m_{35} \ s_{35}] \quad (8)$$

۴- الگوریتم وراثتی

الگوریتم وراثتی یکی از بهترین روشهای موجود برای جستجوی بهینه است [۳۱]. این الگوریتم برای انتخاب ویژگیهای برتر ابزار مفید و کارآمدی است [۳۲-۳۶]. انتخاب ویژگی با الگوریتم وراثتی در مرجع [۳۴] ارائه شده است. در این روش از الگوریتم وراثتی برای یافتن بردار باینری بهینه ای استفاده می‌شود که هر یک از بیتهای آن متناظر یک ویژگی است. اگر i امین بیت این بردار یک باشد، i امین ویژگی در طبقه بندی شرکت کرده و در صورتیکه صفر باشد، ویژگی متناظر آن حذف می‌شود. در این روش، هر یک از زیر مجموعه های ویژگی با توجه به نرخ طبقه بندی درست آن روی مجموعه آزمایش با استفاده از طبقه بند k همسایه نزدیکتر ارزیابی می‌شوند. این روش در مرجع [۳۶] به صورت ترکیب خطی ویژگیها توسعه پیدا کرد. در این روش، تک بیت متناظر هر یک از ویژگیها با ضرایبی با مقادیر حقیقی جایگزین شد، در حالیکه مانند روش قبل امکان حذف ویژگیها هم وجود داشت.

در یک الگوریتم وراثتی، یک جمعیت از افراد طبق مطلوبیت آنها در محیط بقا می‌یابند. افرادی با قابلیتهای بالاتر، شانس ازدواج و تولید مثل بیشتری را خواهند یافت. بنابراین بعد از چند نسل، جمعیتی با کارایی بهتر بوجود می‌آید. در الگوریتم وراثتی هر فرد از جمعیت بصورت یک کروموزوم معرفی می‌شود. کروموزومها در طول چندین نسل کاملتر می‌شوند. در هر نسل کروموزومها ارزیابی می‌شوند و متناسب با ارزش خود امکان بقا و تکثیر می‌یابند. تولید نسل در بحث الگوریتم وراثتی با عملگرهای انتخاب^{۱۷}، همبری^{۱۸}، جهش^{۱۹} و انتقال^{۲۰} صورت می‌گیرد.

والدین برتر بر اساس یک تابع برازندگی^{۲۱} انتخاب می‌شوند. در عمل ازدواج، کروموزومهایی که برازندگی آنها بیشتر باشد، شانس بیشتری برای انتخاب شدن دارند. این عمل بصورت اتفاقی ولی بر اساس مقدار تابع برازندگی صورت می‌گیرد.

در این تحقیق، برای پیدا کردن وزنه‌های بهینه در ترکیب انواع ویژگی برای طبقه بندی، از الگوریتم وراثتی استفاده شده است. وزنه‌های بهینه، وزنه‌هایی هستند که به ازای آنها بیشترین نرخ طبقه بندی حاصل می‌شود. به عبارت دیگر برازندگی کروموزومها توسط نرخ طبقه بندی مشخص می‌شود. علاوه بر آن تاثیر وزن دار کردن مولفه‌های هر نوع ویژگی بر نرخ طبقه بندی نیز با استفاده از الگوریتم وراثتی بررسی شده است. روش انتخاب شده در این مقاله، روش "جمعیت ثابت"

$$F(u,v) = \frac{1}{2} \left(\exp \left(-\frac{1}{2} \left(\frac{(u-u_0)^2}{s_u^2} + \frac{v^2}{s_v^2} \right) \right) + \exp \left(-\frac{1}{2} \left(\frac{(u+u_0)^2}{s_u^2} + \frac{v^2}{s_v^2} \right) \right) \right) \quad (2)$$

که در آن u_0 فرکانس مرکزی موج سینوسی را در راستای محور x ، مشخص می‌کند. s_x و s_y به ترتیب انحراف معیار تابع پوش گوسی در راستای محورهای x و y و $s_u = \left(\frac{1}{2ps_x} \right)$ و $s_v = \left(\frac{1}{2ps_y} \right)$ به ترتیب انحراف معیار تابع گوسی در راستای محورهای u و v هستند. مجموعه‌ای از فیلترهای گابور از طریق گسترش و چرخش تابع $f(x,y)$ که فیلتر مادر نامیده می‌شود، با تابع مولد رابطه ۳، ایجاد می‌شوند.

$$f_{mn}(x,y) = a^{-m} f(x',y') \quad , \quad a \geq 1 \quad (3)$$

$$x' = a^{-m}(x \cos q + y \sin q), \quad y' = a^{-m}(-x \sin q + y \cos q)$$

$$q = n\pi/k, \quad n = 0,1,\dots,k-1, \quad m = 0,1,\dots,s-1$$

که در آن f_{mn} مقیاس شده و چرخش یافته‌ای از فیلتر مادر $f(x,y)$ ، a ضریب مقیاس، n اشاره‌گر مربوط به جهت^{۱۵} فعلی، k تعداد کل جهتها، m اشاره‌گر مربوط به مقیاس^{۱۶} فعلی، s تعداد کل مقیاسها و x' و y' مختصات مقیاس شده و چرخش یافته هستند. ضریب a^{-m} برای مستقل کردن انرژی فیلتر از مقیاس m است. طراحی این فیلترها به گونه‌ای است که محل تقاطع آنها در صفحه نصف توان آنها باشد. به این ترتیب اطمینان حاصل می‌شود که خروجی فیلترها دارای حداکثر اطلاعات مفید و حداقل اطلاعات زائد هستند. چنانچه u_h و u_l به ترتیب برای نمایش فرکانس مرکزی حداکثر و حداقل بکار روند، برای رسیدن به طرح مورد نظر، باید پارامترهای بانک فیلتر با توجه به رابطه ۴ محاسبه شوند.

$$s_v = \frac{\tan \left(\frac{p}{2k} \right) u_h - \left(\frac{s_u^2}{u_h} \right) (2 \ln 2)}{\sqrt{2 \ln 2 - (2 \ln 2)^2 \left(\frac{s_u^2}{u_h^2} \right)}} \quad (4)$$

$$a = \left(\frac{u_h}{u_l} \right)^{\frac{1}{s-1}} \quad s_u = \frac{(a-1)u_h}{(a+1)\sqrt{2 \ln 2}}$$

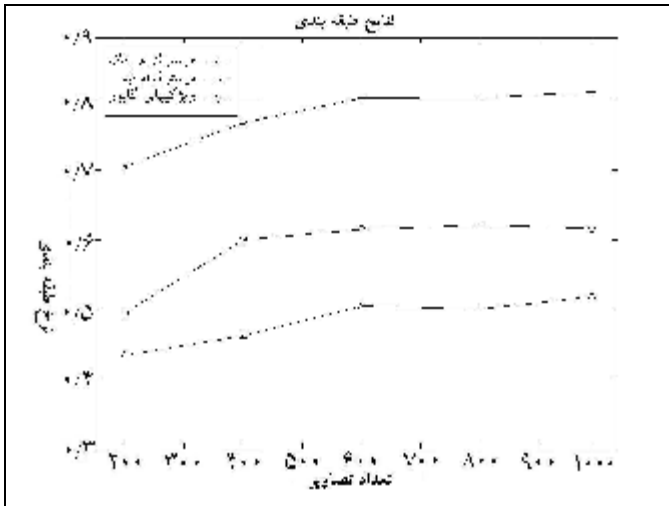
بنا بر پیشنهاد مرجع [۲۸]، u_h برابر $2 \sqrt{128}$ سیکل بر پهنای تصویر و u_l برابر $2 \sqrt{16}$ سیکل بر پهنای تصویر انتخاب شده‌اند. بنابراین مقدار a برابر ۲ خواهد بود. برای تصویر $I(x,y)$ تبدیل موجک گابور با استفاده از رابطه ۵ محاسبه می‌شود.

$$I_{mn}(x,y) = \iint I(x_1,y_1) f^*(x-x_1, y-y_1) dx_1 dy_1 \quad (5)$$

از آنجایی که بدست آوردن زیر-تصویرهای I_{mn} از طریق رابطه ۵ وقت گیر است، این زیر-تصویرها با استفاده از رابطه ۶ در فضای تبدیل فوریه محاسبه می‌شوند.

$$I_{mn}(x,y) = f^{-1} [I(u,v) F_{mn}(u,v)] \quad (6)$$

که در آن ضرایب w_c ، w_e و w_g ، به ترتیب وزنه‌های تعیین کننده نقش ویژگی‌های هیستوگرام رنگ، هیستوگرام جهت لبه و ویژگی‌های گابور است. ضرایب h_i^c ، h_i^e و h_i^g به ترتیب وزنه‌های تعیین کننده نقش مولفه i ام هیستوگرام رنگ، هیستوگرام جهت لبه و ویژگی‌های گابور است. همچنین x_i^c ، x_i^e و x_i^g و y_i^c و y_i^e و y_i^g به ترتیب مولفه i ام از بردارهای ویژگی هیستوگرام رنگ، هیستوگرام جهت لبه و ویژگی‌های گابور، از تصاویر x و y هستند. در رابطه ۱۱، برای ترکیب انواع ویژگی از ترکیب خطی آنها استفاده شده است که متداول ترین روش بکار رفته در زمینه ترکیب ویژگی‌ها است [۴۰-۴۲].



شکل ۳- نرخ طبقه بندی بر اساس هر یک از ویژگی‌های هیستوگرام رنگ، هیستوگرام لبه و ویژگی‌های گابور بر حسب تعداد تصاویر پایگاه

در طبقه بندی با ترکیب انواع ویژگی، ضرایب w و h توسط الگوریتم وراثتی محاسبه می‌شود. هر کروموزوم از کنار هم چیدن بیت‌های مربوط به ضرایب مختلف ایجاد می‌شود. طول هر کروموزوم با توجه به تعداد بیت در نظر گرفته شده برای ضرایب w و h تعیین می‌شود (شکل ۴).

طبقه بندی تصویر در وضعیتی که پایگاه شامل ۲۰۰ تصویر است، برای حالت‌های مختلف انجام شده است که در ادامه به آن پرداخته می‌شود. نتایج این طبقه بندی در جدول ۱ آمده است.

- **حالت یک:** برای ضرایب w ، هر کدام ۱۰ بیت و برای ضرایب h ، هر کدام یک بیت در نظر گرفته شده است. بنابراین طول هر کروموزوم ۱۶۸ بیت است.
- **حالت دو:** برای ضرایب w ، هر کدام ۱۰ بیت و برای ضرایب h ، هر کدام سه بیت در نظر گرفته شده است. بنابراین طول هر کروموزوم ۴۴۴ بیت است.
- **حالت سه:** در این حالت، ضرایب در دو مرحله بهینه شده‌اند.

مرحله اول: فقط ضرایب w یعنی ضرایب تعیین کننده نقش نوع ویژگی، بهینه شده‌اند. برای هر یک از ضرایب w ، ۱۰ بیت در نظر گرفته شده است. در این حالت طول هر کروموزوم ۳۰ بیت است.

مرحله دوم: ضرایب h ، با ثابت نگهداشتن مقادیر w در مقادیر بدست آمده در مرحله اول، بهینه سازی شدند. مرحله دوم تحت دو شرط پیاده سازی شده است. یک مرتبه، حالتی که برای بهینه سازی هر کدام از ضرایب h ، یک بیت در الگوریتم وراثتی در نظر گرفته شده است که آنرا الف می‌نامیم. در این حالت طول هر کروموزوم ۱۳۸ بیت است. بار دیگر برای حالتی که برای بهینه سازی هر کدام از ضرایب h ، سه بیت در الگوریتم وراثتی در نظر گرفته شده است که آنرا ب می‌نامیم. در این حالت طول هر کروموزوم ۴۱۴ بیت است.

- **حالت چهار:** در این حالت ضرایب در دو مرحله بهینه شده‌اند.

است، یعنی کروموزوم‌های والد، پس از ازدواج و تولید فرزند از جمعیت حذف می‌شوند. همچنین برای عملکرد انتخاب از روش چرخ گردان استفاده شده و احتمال عملگرهای همبری، جهش و انتقال به ترتیب برابر ۰/۰۹۵، ۰/۰۰۵ و ۰/۰۱۹ در نظر گرفته شده است [۳۱]. در تمام آزمایش‌های انجام شده، تعداد کروموزوم‌های جمعیت اولیه ۵۰ و تعداد تکرار الگوریتم ۲۰ در نظر گرفته شده است.

۵- طبقه بندی، آزمایشها و بررسی نتایج

۱-۵ طبقه بندی

روش k همسایه نزدیکتر، متداول ترین روش برای طبقه بندی و بازیابی تصویر است [۳۸ و ۳۷، ۱۵، ۱۲، ۱۱]. در این روش، برای تعیین گروه یک تصویر، باید بردار ویژگی آن با نمایه‌های تمام تصاویر پایگاه تصویر که برچسب آنها مشخص است، مقایسه شود. طبق تحقیقات انجام شده در این زمینه، معیار فاصله نقش بسزایی در نرخ طبقه بندی دارد [۳۷-۳۹]. بنا بر تحقیقی که قبلاً در این زمینه انجام داده ایم [۳۷]، از بین ۱۱ فاصله مطالعه شده، فواصل C^2 ، L_2 و C^2 به ترتیب برای ویژگی‌های هیستوگرام رنگ، هیستوگرام جهت لبه‌ها و ویژگی‌های گابور، بهترین نرخ طبقه بندی را نتیجه می‌دهند. بنابراین در این تحقیق، برای ویژگی‌های ذکر شده از این فواصل استفاده شده است. فواصل C^2 ، L_2 ، برای دو بردار n بعدی $\underline{X}=(x_1, x_2, \dots, x_n)$ و $\underline{Y}=(y_1, y_2, \dots, y_n)$ به ترتیب مطابق روابط ۹ و ۱۰ تعریف می‌شوند. همچنین در تحقیق فوق الذکر [۳۷]، مقدار بهینه k برای طبقه بند k همسایه نزدیکتر پنج بدست آمده است که در این تحقیق از آن استفاده می‌شود.

$$C^2(\underline{X}, \underline{Y}) = \sum_{i=1}^n \left(\frac{x_i - y_i}{x_i + y_i} \right)^2 \quad (9)$$

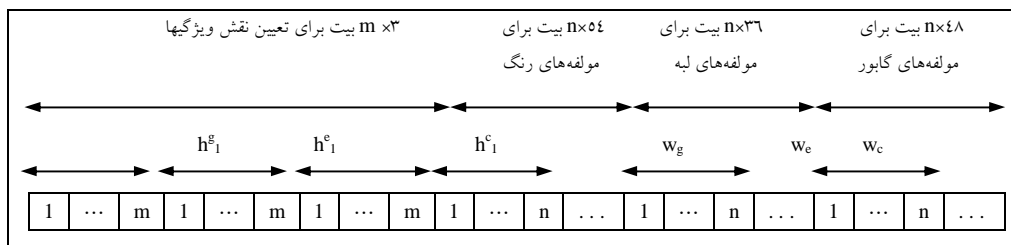
$$L_2(\underline{X}, \underline{Y}) = \left(\sum_{i=0}^n |x_i - y_i|^2 \right)^{\frac{1}{2}} \quad (10)$$

۲-۵ آزمایشها و نتایج

برای مشاهده اثر تعداد تصاویر بر نرخ طبقه بندی، نتایج طبقه بندی در پنج مرحله محاسبه شده است. در مرحله اول ۲۰۰ تصویر، از هر گروه معنایی ۲۰ تصویر، به تصادف انتخاب شده و طبقه بندی انجام شده است. در مراحل دو تا پنج، در هر مرحله ۲۰۰ تصویر به پایگاه داده اضافه شده است، یعنی اینکه در مراحل دو تا پنج تعداد تصاویر به ترتیب ۴۰۰، ۶۰۰، ۸۰۰ و ۱۰۰۰ است. به عبارت دیگر، در هر مرحله، به تصاویر هر کلاس معنایی نسبت به مرحله قبل ۲۰ تصویر اضافه شده است. برای بدست آوردن کارایی طبقه بندی 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 طبقه بند ارائه می‌شود و برچسب واقعی تصویر مذکور با برچسبی که الگوریتم طبقه بند به آن نسبت می‌دهد مقایسه می‌شود. این آزمایش برای تمام تصاویر انجام می‌شود و نرخ طبقه بندی محاسبه می‌شود.

ابتدا نرخ طبقه بندی را برای ویژگی‌های هیستوگرام رنگ، هیستوگرام جهت لبه و ویژگی گابور به تفکیک به ازای تعداد مختلف تصاویر پایگاه تصویر بدست آورده ایم که نتایج آن در شکل ۳ آمده است. در طبقه بندی با استفاده از ترکیب ویژگی‌ها، فاصله بردارهای ویژگی دو تصویر x و y با توجه به رابطه ۱۱ تعریف شده است.

$$D(\underline{X}, \underline{Y}) = w_c \left| \sum_{i=1}^{54} h_i^c \left(\frac{x_i^c - y_i^c}{x_i^c + y_i^c} \right)^2 \right| + w_e \left| \left(\sum_{i=0}^{36} h_i^e |x_i^e - y_i^e|^2 \right)^{\frac{1}{2}} \right| + w_g \left| \sum_{i=1}^{48} h_i^g \left(\frac{x_i^g - y_i^g}{x_i^g + y_i^g} \right)^2 \right| \quad (11)$$



شکل ۴- نحوه کنار هم قرار گرفتن ضرایب مختلف و تشکیل یک کروموزوم
 m و n به ترتیب نشاندهنده تعداد بیت‌های در نظر گرفته شده برای ضرایب w و h هستند

جدول ۱- نرخ طبقه بندی برای حالت‌های مختلف ترکیب ویژگیها در پایگاه شامل ۲۰۰ تصویر

مرحله دوم	چهار			سه		دو	یک	حالت
	مرحله اول			مرحله دوم				
	ویژگی گابور	ویژگی لبه	ویژگی رنگ	الف	ب			
	۰.۸۱/۵	۰.۵۲/۵	۰.۴۸	۰.۷۵	۰.۸۲	۰.۷۹	۰.۷۸	نرخ طبقه‌بندی

و برای تصاویر بیشتر ثابت می ماند. با توجه به جدول ۱، نتیجه می‌گیریم که فرآیند دو مرحله‌ای برای پیدا کردن مقادیر بهینه ضرایب w یعنی ضرایب تعیین کننده نقش نوع ویژگیها و h ها یعنی ضرایب تعیین کننده نقش مولفه‌های ویژگی موثرتر از فرآیند تک مرحله‌ای است. مقایسه حالت یک با حالت سه- مرحله دوم- الف، حالت دو با حالت سه- مرحله دوم- ب و حالت دو با حالت چهار- مرحله دوم، در جدول ۱ بیانگر این موضوع است. همچنین نتایج آزمایشها نشان می‌دهد که وزن دار کردن مولفه‌های ویژگی با در نظر گرفتن سه بیت برای هر یک از آنها، موثرتر از انتخاب ویژگی با در نظر گرفتن یک بیت برای هر یک از مولفه‌های ویژگی، به معنی بودن یا نبودن آن مولفه، است. مقایسه حالت‌های یک با دو و مقایسه حالت سه- مرحله دوم- الف با مرحله دوم- ب در جدول ۱ این موضوع را نشان می‌دهد. از دیگر نتایج قابل ذکر و مهم آزمایشها می‌توان به این نکته اشاره کرد که ترتیب بهینه کردن ضرایب h و w اثر چندانی در نتایج ندارد (مقایسه حالت سه- مرحله دوم- ب با حالت چهار- مرحله دوم در جدول ۱). با توجه به شکل ۵ نیز مشاهده می‌شود که اگر تعداد تصاویر پایگاه برای ۱۰ گروه معنایی استفاده شده

مرحله اول: فقط ضرایب h، ضرایب تعیین کننده نقش مولفه‌های ویژگی، بهینه شده‌اند.

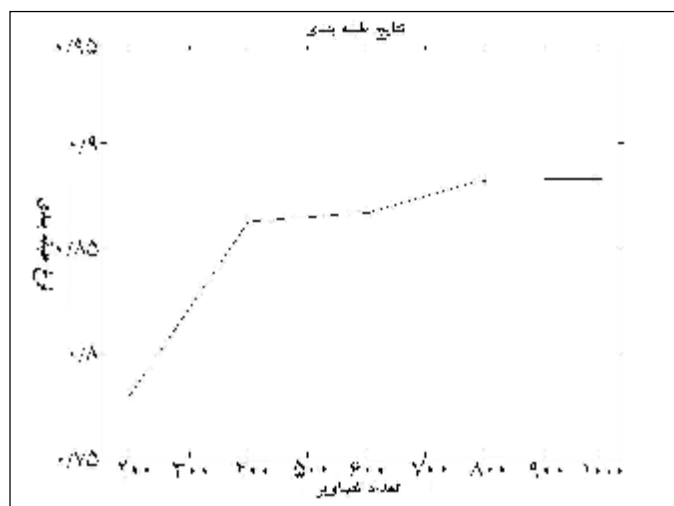
این مرحله در سه بخش مختلف انجام شده است. در بخش اول ضرایب h_i^c تعیین شده‌اند. برای تعیین این ضرایب، در طبقه بندی فقط از ویژگی رنگ استفاده شده است و دو ویژگی دیگر کنار گذاشته شده‌اند. در این مرحله، برای هر یک از ضرایب در الگوریتم وراثتی سه بیت در نظر گرفته شده است.

طول کروموزوم ۱۶۲ بیت است. در بخش‌های دوم و سوم به ترتیب ضرایب h_i^e و h_i^c برای ویژگیهای لبه و گابور تعیین شده‌اند. در تعیین این ضرایب نیز مانند بخش اول دو ویژگی دیگر کنار گذاشته شده و طبقه بندی انجام شده است. در اینجا نیز برای هر ضریب، سه بیت در نظر گرفته شده است. طول کروموزوم برای بخش‌های دوم و سوم به ترتیب ۱۰۸ و ۱۴۴ بیت است.

مرحله دوم: ضرایب w، با ثابت نگهداشتن ضرایب h، در مقادیر بهینه مرحله اول تعیین شده‌اند. برای هر یک از ضرایب w، ۱۰ بیت در نظر گرفته شده است. طول کروموزوم در این حالت ۳۰ بیت است.

همچنین آزمایشهایی بر روی پایگاه‌های با اندازه‌های ۴۰۰، ۶۰۰، ۸۰۰ و ۱۰۰۰ تصویر، مشابه حالت سه- مرحله اول، انجام شده است؛ یعنی اینکه ضرایب w بهینه سازی شده‌اند. در این حالت تاثیر تمام مولفه‌های ویژگی در طبقه بندی یکسان در نظر گرفته شده است. نرخ طبقه بندی حاصل از این آزمایشها در گراف شکل ۵ آمده است. این گراف تاثیر تعداد تصاویر را در طبقه بندی نشان می‌دهد. جدول ۲ تعداد تصاویری را که در هر گروه به درستی یا به اشتباه طبقه بندی شده‌اند، برای حالتی که پایگاه شامل ۱۰۰۰ تصویر است؛ نشان می‌دهد.

درایه (i,j) این جدول نشان دهنده تعداد تصاویری از کلاس i ام است که به اشتباه به کلاس j ام نسبت داده شده‌اند. همچنین درایه (i,i) تعداد تصاویری از گروه i ام را نشان می‌دهد که بدرستی طبقه بندی شده‌اند. تصاویری که از گروه دو و گروه هفت به اشتباه طبقه بندی شده‌اند، به ترتیب در شکل‌های ۶ و ۷ آمده‌اند.



شکل ۵- نرخ طبقه بندی ترکیب ویژگیهای هیستوگرام رنگ، هیستوگرام لبه و ویژگیهای گابور بر اساس تعداد تصاویر پایگاه. بهینه سازی ضرایب طبق حالت سه مرحله اول انجام شده است

۳-۵ بررسی نتایج

شکل ۳ نشان می‌دهد که برای طبقه بندی معنایی تصاویر در ۱۰ گروه، با استفاده از یک نوع ویژگی، نرخ طبقه بندی با ۶۰ تصویر در هر گروه تقریباً اشباع می‌شود

جدول ۲- نتایج طبقه بندی تصاویر پایگاه در گروههای ۱۰ گانه معنایی برای پایگاه ۱۰۰۰ تصویری شماره گروه نسبت داده شده به تصویر

	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
۱	۶۴	۱۱	۳	۰	۶	۳	۱	۰	۲	۱۰
۲	۰	۸۹	۰	۰	۰	۶	۰	۴	۱	۰
۳	۲	۳	۸۵	۰	۲	۱	۲	۰	۴	۱
۴	۰	۰	۱	۹۹	۰	۰	۰	۰	۰	۰
۵	۰	۱	۱	۳	۹۲	۰	۰	۰	۲	۱
۶	۰	۳	۱	۰	۰	۹۱	۰	۳	۱	۱
۷	۰	۵	۰	۰	۰	۰	۹۰	۴	۰	۱
۸	۰	۰	۰	۰	۰	۰	۰	۱۰۰	۰	۰
۹	۴	۳	۷	۲	۰	۸	۱	۰	۷۵	۰
۱۰	۶	۱	۰	۱	۰	۰	۰	۲	۱	۸۹

هر یک از ویژگیها برقرار کرد و نیز اینکه الگوریتم وراثتی می تواند نقش مهمی در این زمینه ایفا کند.

مراجع

- [1] S. Antani, R. Kasturi and R. Jain, "A Survey on the Use of Pattern Recognition Methods for Abstraction, Indexing and Retrieval," *Pattern Recognition*, vol. 1, pp. 945-965, 2002.
- [2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-based Image Retrieval at the End of the Early Years," *IEEE Trans. on Pattern Analysis and Machine Intelligence* vol. 22, no. 12, pp. 1349-1380, 2000.
- [3] R. C. Veitkamp and M. Tanase, "Content-based Image Retrieval System: a Survey," Technical Report, UU-CS-2000-34, University of Utrecht, 2000, <http://www.cs.uu.nl/research/techreps/UU-CS-2000-34.html>.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] R. Zhang and Z. Zhang, "A Clustering Based Approach to Efficient Image Retrieval," *Proceedings of the 14th IEEE Int. Conf. on Tools with Artificial Intelligence (ICTAI'02)*, pp. 339-346, 2002.

از ۴۰۰ تصویر، یعنی ۴۰ تصویر برای هر گروه، بیشتر شود، نرخ طبقه بندی تغییر محسوسی نخواهد داشت.

با توجه به جدول ۲، ۱۱ تصویر از گروه شیرها به اشتباه به سایر گروهها نسبت داده شده است. از این تعداد، چهار تصویر به گروه اسبها و شش تصویر به گروه فیلهها و تنها یک تصویر به گروه کوهستان نسبت داده شده است (شکل ۶). همچنین نه تصویر از گروه فیلهها به اشتباه طبقه بندی شده اند که از این تعداد، سه تصویر به گروه اسبها و سه تصویر به گروه شیرها نسبت داده شده است (جدول ۲). این در حالیکه تمام تصاویر گروه اسبها به درستی طبقه بندی شده اند (جدول ۲). بنابراین در مجموع تعداد ۲۰ تصویر از این سه گروه به اشتباه طبقه بندی شده است. این نتایج نشان می دهد، با توجه به نزدیک بودن ویژگیهای سطح پایین این گروهها، با ترکیب مناسب ویژگیها می توان طبقه بندی صحیح این گروهها را در حد مطلوبی انجام داد. چنانچه این سه گروه معنایی با یکدیگر ادغام شده و گروه معنایی بزرگتری به نام "حیوانات" ایجاد کنند، میزان طبقه بندی اشتباه از ۲۰ تصویر به چهار تصویر کاهش پیدا خواهد کرد. بنابراین ادغام گروههای معنایی با مفاهیم خیلی دقیق که ویژگیهای سطح پایین آنها به یکدیگر نزدیک است و تشکیل گروههای معنایی با مفاهیم وسیعتر و طبقه بندی تصاویر این گروه به زیر-گروههای دقیق با استفاده از ترکیب مناسبتری از ویژگیها، طبقه بندی سلسله مراتبی، می تواند برای افزایش نرخ طبقه بندی موثر باشد.

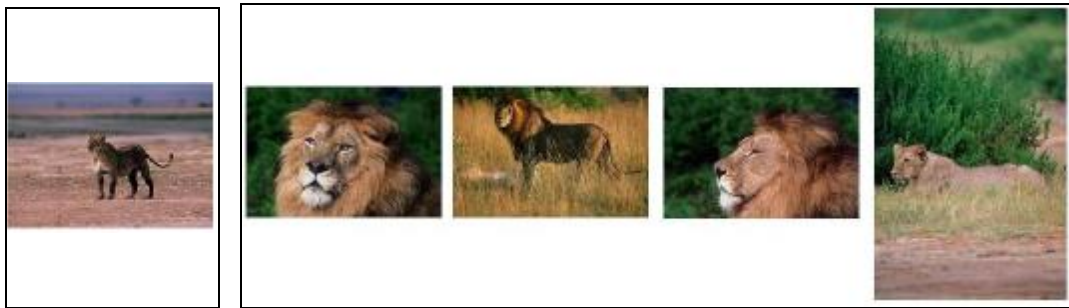
همچنین بررسی جدول ۲ نشان می دهد که بیشترین نرخ خطا در گروههای معنایی یک، سه و نه، به ترتیب گروههای مردم، ساختمانها و کوهستان، اتفاق می افتد. قبلاً نیز اشاره شد که این گروهها از تنوع بسیار برخوردارند. با توجه به جدول ۲ مشاهده می شود که تصاویر به اشتباه طبقه بندی شده این گروهها، در بین گروههای دیگر توزیع شده اند به گونه ای که تصاویر به اشتباه طبقه بندی شده گروههای مردم و ساختمان در هفت گروه و تصاویر به اشتباه طبقه بندی شده گروه کوهستان در شش گروه دیگر توزیع شده اند. این نتایج نیز تنوع بسیار این گروهها را تایید می کند. تقسیم گروههای متنوع از نظر ویژگیهای دیداری به گروههای کوچکتر می تواند راهکاری عملی برای کاهش خطای طبقه بندی باشد.

۶ - جمع بندی

نمایه سازی و بازیابی تصویر یکی از زمینه های مهم و فعال تحقیقاتی در بنیای ماشینی است. در سیستمهای بازیابی بر اساس محتوی، عموماً پرس و جوی کاربر بر پایه ویژگیهای معنایی انجام می شود. این در حالیکه که نمایه سازی معنایی تصاویر با استفاده از ویژگیهای سطح پایین کار دشواری است. طبقه بندی تصویر یکی از راههایی است که با استفاده از آن می توان به ویژگی معنایی تصویر پرس و جوی برد و کاربر را در راه رسیدن به تصاویر دلخواهش کمک کرد.

در این مقاله، ترکیب سه نوع ویژگی متفاوت هیستوگرام رنگ، هیستوگرام جهت لبه و ویژگیهای گابور در طبقه بندی بررسی شد. برای ترکیب این ویژگیها از الگوریتم وراثتی استفاده شد. از طرفی اثر وزن دار کردن مولفه های هر نوع ویژگی بر نرخ طبقه بندی مطالعه شد. در مورد وزن دار کردن مولفه های ویژگی و تعیین ضرایب هر یک از سه نوع ویژگی، آزمایشهای مختلفی انجام شد (جدول ۱). نتایج این آزمایشها نشان می دهد که تعیین این ضرایب در دو مرحله، بهتر از تعیین آنها در یک مرحله است و نیز اینکه الویت بهینه سازی آنها در فرآیند دو مرحله ای تاثیر چندانی بر نرخ طبقه بندی ندارد. همچنین اثر تعداد تصاویر پایگاه بر نرخ طبقه بندی مطالعه شد. نتایج این بررسی نشان می دهد که اگر تعداد تصاویر پایگاه برای ۱۰ گروه معنایی استفاده شده، از ۴۰۰ تصویر، یعنی ۴۰ تصویر برای هر گروه، بیشتر شود نرخ طبقه بندی تغییر محسوسی نخواهد داشت (شکل ۵).

نتایج این تحقیق نشان می دهد که می توان ارتباط معنی داری بین ویژگیهای سطح پایین و ویژگیهای سطح بالا از طریق بهینه سازی ضرایب تعیین کننده نقش



(ب)

(الف)



(ج)

شکل ۶- تصاویری از گروه دوم، شیرها، که به اشتباه طبقه بندی شده اند، (الف) این تصاویر در گروه هشتم، اسبها، طبقه بندی شده اند، (ب) این تصویر در گروه نهم، کوهستان، طبقه بندی شده است، (ج) این تصاویر در گروه ششم، فیلها، طبقه بندی شده اند



(الف)



(ج)

(ب)

شکل ۷ - تصاویری از گروه هفتم، گلها، که به اشتباه طبقه بندی شده اند، (الف) این تصاویر در گروه دوم، شیرها، طبقه بندی شده اند، (ب) این تصاویر در گروه هشتم، اسبها، طبقه بندی شده اند، (ج) این تصویر در گروه دهم، غذاها، طبقه بندی شده است

- [20] H. W. Yoo, S. H. Jung, D. S. Jang and Y. K. Na, "Extraction of Major Object Features Using VQ Clustering for Content-based Image Retrieval," *Pattern Recognition*, vol. 35, pp.1115-1126, 2002.
- [21] M. G. Albanesi, S. Bandelli and M. Ferretti, "Quantitative Assessment of Qualitative Color Perception in Image Database Retrieval," *IEEE Int. Conf. on Image Analysis and Processing*, pp. 410-415, 2001.
- [22] H. W. Yoo, D. S. Jang, S. H. Juang and J. H. Park, "Visual Information Retrieval System via Content-based Approach," *Pattern Recognition*, vol. 35, pp. 749-769, 2002.
- [23] J. R. Smith, S. F. Chang, "Tools and Techniques for Color Image Retrieval," *Storage and Retrieval for Image and Video Database IV, Proc. SPIE*, vol. 2670, 426-437. 1996.
- [24] A. K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233-1244, 1996.
- [25] B. S. Manjunath and W. Y. Ma, "Texture Feature for Browsing and Retrieval of Image Data," *IEEE PAMI*, vol. 18, no. 8, pp. 837-842. 1996.
- [26] Y. Zhu, T. Tan and Y. Wang, "Font Recognition Based on Global Texture Analysis," *IEEE Transaction on PAMI*, vol. 23, no. 10, 2001.
- [27] A. Kumar and G. K. H. Pang, "Defect Detection in Texture Materials using Gabor Filters," *IEEE Transaction on Industry Applications*, vol. 38, no. 2, 2002.
- [28] 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.
- [۲۹] ف. میرزاپور و ح. قاسمیان، " آشکارسازی یافت تصاویر ماهواره با استفاده از موجکهای گابور، " دومین کنفرانس ماشین بینایی و پردازش تصویر ایران، جلد اول، ص ۱۴۰ تا ۱۴۵، ۱۳۸۱.
- [۳۰] ا. فرامرزی، ا. صیادیان و ع. احمدیان، " شاخص گذاری بر روی تصاویر با استفاده از موجکهای گابور و ممانهای لزاندر، " دومین کنفرانس ماشین بینایی و پردازش تصویر ایران، جلد دوم، ص ۲۱۸ تا ۲۲۵، ۱۳۸۱.
- [31] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
- [32] H. Vafaie and K. De Jong, "Robust Feature Selection Algorithms," *Proc. of the fifth Conference on Tools for Artificial Intelligence, Boston*, pp. 356-363, 1993.
- [8] G. D. Guo, A. K. Jain, W. Y. Ma and H. J. Zhang, "Learning Similarity Measures for Natural Image Retrieval with Relevance Feedback," *IEEE Trans. on Neural Networks*, vol. 13, no. 4, 2002.
- [9] X. He, O. King, W. Y. Ma, M. Li and H. J. Zhang, "Learning a Semantic Space from User's Relevance Feedback for Image Retrieval," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 13, no. 1, pp. 39-48, 2003.
- [10] B. Bhanu and A. Dong, "Concept Learning with Fuzzy Clustering and Relevance Feedback," *Engineering Application of Artificial Intelligence*, vol. 15, pp. 123-138, 2002.
- [11] M. Szummer and R. W. Picard, "Indoor-outdoor Image Classification," *IEEE Int. Workshop on Content-Based Access of Image and Video Database*, 1998.
- [12] A. Vailaya, A. K. Jain and H. J. Zhang, "On image Classification: City vs. landscape," *Pattern Recognition*, vol. 31, pp. 1921-1935, 1998.
- [13] 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.
- [14] Q. Iqbal and J. K. Aggarwal, "Using Structure in Content-based Image Retrieval," *IASTED Int. Conf. Signal and Image Processing*, pp. 129-133, 1999.
- [15] A. Guerin-Dugue and A. Oliva, "Classification of Scene Photographs from Local Orientations Features," *Pattern Recognition Letters*, vol. 21, pp. 1135-1140, 2000.
- [16] J. R. Smith and C. S. Li, "Image Classification and Querying Using Composite Region Templates," *Computer Vision and Image Understanding*, vol. 75, pp.165-174, 1999.
- [17] H. Nezamabadi-pour, and E. Kabir, "Image Retrieval Using Histograms of Unicolor and Bicolor Blocks and Directional Changes in Intensity Gradient," Accepted for Publication in *Pattern Recognition Letters*, 2004.
- [۱۸] ح. نظام آبادی پور، ا. کبیر و س. سریزدی، " بازیابی تصویر با استفاده از اطلاعات رنگ و شکل " ، دومین کنفرانس ماشین بینایی و پردازش تصویر ایران، جلد دوم، صفحات ۲۲۶ تا ۲۳۲، ۱۳۸۱.
- [۱۹] ح. نظام آبادی پور، ا. کبیر و س. سریزدی، " بازیابی تصویر بر اساس هم‌رخدای رنگها در بلوکهای لبه " ، هشتمین کنفرانس انجمن کامپیوتر ایران، ص ۳۵۹ تا ۳۶۴، ۱۳۸۱.

- 1 leave-one-out
- 2 Content-based image retrieval (CBIR)
- 3 Low-level visual features
- 4 Color layout
- 5 High-level features
- 6 Semantic features
- 7 Semantic gap
- 8 Multi-resolution Simultaneous autoregressive
- 9 Binary Bayesian classifier
- 10 Global distribution of the local dominant orientation
- 11 Corel
- 12 Simplicity
- 13 Red- Green- Blue
- 14 Hue- Saturation- Value
- 15 Orientation
- 16 Scale
- 17 Selection
- 18 Cross-over
- 19 Mutation
- 20 Transfer
- 21 Fitness function



حسین نظام آبادی پور کارشناسی و کارشناسی ارشد خود

را در مهندسی برق- الکترونیک به ترتیب از دانشگاه شهید باهنر کرمان و دانشگاه تربیت مدرس در سالهای ۱۳۷۷ و ۱۳۷۹ دریافت کرد. او سپس دوره دکترای خود را در مهندسی برق- الکترونیک از دانشگاه تربیت مدرس در سال ۱۳۸۳ دریافت کرد و اکنون استادیار بخش مهندسی برق دانشگاه شهید باهنر کرمان است. زمینه های پژوهشی مورد علاقه او بازیابی تصویر، بازشناسی الگو و کاربرد رایانش نرم در پردازش تصویر است.



احسان اله کبیر کارشناسی ارشد پیوسته خود را در

مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستمهای الکترونیک از دانشگاه اسکس در انگلستان، به ترتیب در سالهای ۱۳۶۴ و ۱۳۶۹، دریافت کرد. او اکنون دانشیار بخش مهندسی برق دانشگاه تربیت مدرس است. زمینه های پژوهشی مورد علاقه او بازشناسی الگو به ویژه بازشناسی متون چاپی و دستنویس و همچنین بینایی ماشین است.

[33] M. L. Raymer, W. F. Punch, E. Gooddman, L. A. Kuhn and A. K. Jain, "Dimentionalty Reduction Using Genetic Algorithm," *IEEE Transaction on Evolutionary Computation* vol. 4, no. 2, pp. 164-171, 2000.

[34] W. Siedlecki and J. Sklansky, "A Note on Genetic Algorithm for Large Scale Feature Selection," *Pattern Recognition Letters*, vol. 10, pp. 335-347, 1989.

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

[36] J. D. Kelly and L. Davis, "Hybridizing the Genetic Algorithm and the k-nearest Neighbor Classification Algorithm," *Proc. 4th Int. Conf. on Genetic Algorithms Application*, pp. 377-383, 1991.

[۳۷] ح. نظام آبادی پور و ا. کبیر، "ارزیابی معیارهای عدم شباهت در طبقه بندی تصویر،" **نهمین کنفرانس سالانه انجمن کامپیوتر ایران**، دانشگاه صنعتی شریف، ص ۳۶۲ تا ۳۷۰، ۱۳۸۲.

[38] Y. Rubner, J. Puzicha, C. Tomasi, and J. M. Buhmann, "Empirical Evaluation of Dissimilarity Measures for Color and Texture," *Computer Vision and Image Understanding*, vol. 84, pp. 25-43, 2001.

[39] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos, "A Novel Vector based Approach to color Image Retrieval using a vector angular-based distance measure," *Computer Vision and Image Understanding*, vol. 75, pp. 46-58, 1999.

[40] B. M. Mehtre, M. S. Kankanhalli and W. F. Lee, "Content-based Image Retrieval using a Composite Color-shape Approach," *Information Processing and Management*, vol. 34, no. 1, pp. 109-120, 1998.

[41] Y. Rui, T. S.Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: a Power Tool for Interactive Content-based Image Retrieval," *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 8, no. 5, 1998.

[42] H. W. Yoo, D. S. Jang, S. H. Juang and J. H. Park, "Visual Information Retrieval System via Content-based Approach," *Pattern Recognition*, vol. 35, pp. 749-769, 2002.