

به کارگیری تکنیک های پردازش متن در دسته بندی خودکار کلیپ های ویدئویی

جلال حشمت پناه^۱، محمد جواد احمدی^۲

^۱ کارشناسی ارشد الکترونیک، دانشگاه آزاد اسلامی، واحد شهرری، گروه الکترونیک، تهران، ایران، jalal.panah@gmail.com

^۲ کارشناسی ارشد نرم افزار، mohamad.j.ahmadi@gmail.com

چکیده

در این مقاله روشی برای دسته بندی تصاویر ویدئویی ارائه شده است. هدف، ارائه مدلی است که با آن بتوان دسته ای که یک ویدئو بدان تعلق دارد را تخمین زد. برای این منظور از تکنیک کیف کلمات که از روش های پردازش متن می باشد استفاده شده است. با این تفاوت که به جای جملات با کلیپ های ویدئویی و به جای کاراکترها با هیستوگرام های فریم های ویدئو سروکار داریم. فریم های ویدئویی در این روش به رشته های کاراکتری طولانی تبدیل گشته و دسته بندی بر روی این رشته های کاراکتری انجام می پذیرد. آزمایشات عملی بر روی ویدئوهای ورزشی که در چهار دسته قرار گرفته اند انجام پذیرفت.

کلمات کلیدی

دسته بندی ویدئو، کیف کلمات، توصیفگر SIFT

۱- مقدمه

یکی از کاربردهای پردازش تصاویر بازشناسی اشیاء و سکانس در تصاویر ویدئویی و دسته بندی^۱ موضوعی ویدئو ها میباشد. دسته بندی تصاویر و تصاویر ویدئویی به شاخه های موضوعی، مسئله ای است که در حوزه های تحقیقاتی و کاربردی بسیاری مورد توجه قرار گرفته است. به عنوان مثال یک مجموعه آنلاین از تصاویر با استفاده از الگوریتم های دسته بندی به شاخه های "تصویر اشخاص"، "منظره" یا "حیوانات" دسته بندی می شود تا مرور تصاویر با کارایی بهتری صورت پذیرد. امروزه با گسترش ویدئوهای موجود در اینترنت و تلویزیون، آرشیوهای حجیم خانگی و سازمانی از ویدئوهای مختلف بوجود آمده است و برای جستجو در میان آرشیو بزرگی از ویدئوهای خبری، دسته بندی ویدئوها بر اساس موضوعات یا بر اساس سکانسی خاص (مثلا ملاقات اشخاص) یا اشیاء (مثلا ساختمان) می تواند باعث کارایی جستجو شود. در این مقاله با استفاده از الگوریتم های یادگیری ماشین و ترکیب تکنیک های مورد استفاده در پردازش متن و تصویر، روشی ارائه می دهیم که با استفاده از آن می توان دسته موضوعی یک کلیپ ویدئو را تخمین زد.

از یک منظر، تحقیقات در زمینه دسته بندی وقایع و موضوعی ویدئو در سه شاخه عمده: شناسایی وقایع غیرمتعارف در تصاویر ویدئویی [2] [1]، طبقه بندی فعالیت-

های انسان در تصاویر ویدئویی [3] و بازشناسی وقایع در ویدئو [4] قرار می گیرد. شناسایی فعالیت های مشکوک در سیستم های نظارتی و ردیابی ویدئویی، شاخه تحقیقاتی فعالی بوده و راه حل های گوناگونی برای آن ارائه شده است. بسیاری از این راه حل ها بر مبنای مدل های HMM یا شبکه های بیزی معرفی شده اند.

اخیراً برای دسته بندی تصاویر از روشی استفاده می گردد به نام کیف کلمات^۲ که به دلیل عملکرد و کارایی بالایی که از خود نشان داده است بسیار محبوب نیز گشته است. این روش در ابتدا برای پردازش زبان های طبیعی و بازیابی اطلاعات معرفی شده بود و از آن برای دسته بندی متن استفاده می شد. در حوزه دسته بندی تصاویر و یا شناسایی اشیاء نیز از این روش استفاده شده است. حتی با استفاده از روش کیف کلمات می توان کلیپ های ویدئویی با موضوعات مختلف را نیز دسته بندی نمود [5]. در [5] با استفاده از این روش و توصیف کلیپ های ویدئو بوسیله هیستوگرام متوالی متشکل از مراکز خوشه و نمایش هر یک از کلیپ های ویدئو بوسیله بردارهای وزن دار چند بعدی به دسته بندی کلیپ های ویدئویی بر اساس موضوعی که در بر داشتند پرداخته اند. آنها با معرفی فاصله تصحیح نیدل من واچ و با استفاده از ماشینهای برداری پشتیبان بر مبنای کرنل های رشته ای توانستند با دقت خوبی عملیات دسته بندی را پیاپی سازی کنند.