

تشخیص تغییر صحنه در استاندارد کدکننده ویدیویی پربازده

نوید درفشان^۱، محمد رضا رمضان پور^{۲*}^۱دانشگاه آزاد اسلامی، واحد نجف آباد، گروه مهندسی کامپیوتر، نجف آباد، ایران

navid.dorfeshan@gmail.com

^۲دانشگاه آزاد اسلامی واحد مبارکه، گروه مهندسی کامپیوتر، مبارکه، اصفهان، ایران

ramezanpour@mau.ac.ir

چکیده

آخرین استاندارد رمز گذاری ویدئو استاندارد کدگذاری ویدئویی پربازده نام دارد. یکی از چالش های موجود در این حوزه تشخیص تغییر صحنه است تاکنون تحقیقات بسیاری در زمینه تشخیص تغییر صحنه در حوزه استاندارد های مختلف ارائه شده است. که بیشتر آنها از آستانه های ثابت برای معیارهای شباهت و تصمیم گیری اینکه یک تغییر صحنه وجود دارد یا خیر استفاده می کنند. این آستانه ها با تجربی بدست آمده اند و یا باید قبل از تشخیص تغییر صحنه محاسبه شوند و پس از اتمام دنباله فیلم به دست آیند. عملکرد تشخیص تغییر صحنه کاهش قابل توجهی برای فیلم ها با پیچیدگی بالا و تنوع صحنه داشته است. در این مقاله ابتدا، استاندارد ویدئویی پربازده که آخرین استاندارد رمزگذاری تصویر نام دارد را مورد بررسی قرار می دهیم، این استاندارد دارای ساختار بلوکی در اندازه های 4×4 تا 64×64 می باشد که اندازه هر بلوک با توجه به جزییات یک تصویر مشخص می گردد. استاندارد ویدئویی پربازده اطلاعات بهتری درباره تصاویر فشرده کد شده نسبت به استاندارد H.265/AVC فراهم می کند. در ادامه به بررسی آثار قبلی در این زمینه و ارتباط بین ویژگی های آماری محلی را در مدت زمان اجرای صحنه و تغییر صحنه می پردازیم. سپس با استفاده از ویژگی های بلوک تبدیل و یک مدل آستانه دینامیک تشخیص تغییر صحنه را در استاندارد کدکننده ویدئویی پربازده پیاده سازی می کنیم. نتایج بدست آمده بر روی فیلم ها با فرمت های مختلف که توسط گروه JCT-VC ارائه شده است، عملکرد بسیار خوبی با بهبود قابل توجهی در میزان دقت و مقدار فراخوانی، پردازش ویدئویی با حداقل پیچیدگی را نشان می دهد.

واژه های کلیدی: استاندارد کدکننده ویدئویی پربازده، تشخیص تغییر صحنه، آستانه دینامیک

۱- مقدمه

همه روزه مقادیر عظیمی از داده های ویدئویی ویدئویی در سرتاسر جهان تولید می شود و پردازش کردن این داده ها، نیازمند منابع فراوانی همچون زمان، نیروی انسانی و سخت افزاری قدرتمند می باشد.

در قرن اخیر تکنولوژی ویدئوی دیجیتال با سرعت زیادی در حال رشد است و به سبب پیشرفت تکنولوژی، ذخیره مقدار عظیم داده

ویدئویی آسانتر شده است. بخش عمده محتوای ویدئوها شامل اخبار، فیلم، ورزش، اسناد و... می شود. به هر حال نیاز به جستجوی اطلاعات مهم در ویدئوها هم افزایش یافته است. بطور مثال دوربین های حفاظتی بطور ۲۴ ساعته در حال تولید اطلاعات ویدئویی می باشند که جستجو در این اطلاعات کار زمان بر و طاقت فرسایی می باشد. برای حل این مشکل راه حل های زیادی بیان شده است، در این بین روش های تشخیص تغییر صحنه برای خلاصه سازی ویدئو نقش مهمی را در این خصوص ایفا کرده و به کاربر برای بازیابی اطلاعات در ویدئو کمک فراوان می کند.

بسیاری از روش هایی که تاکنون برای تشخیص تغییر صحنه معرفی شده اند، از قوانین اکتشافی که معمولاً بر اساس تجربه فردی بدست می آیند، استفاده می کنند. از جمله این روش ها می توان به اعمال حدآستانه به صورت ثابت یا پویا [۱] و قوانین فازی [۲] اشاره کرد. اما در پژوهش های اخیر، بیشتر از روش های آماری نظیر خوشه بندی، مدل سازی احتمالی با توابع گوسی [۳] و شبکه عصبی [۴] استفاده گردیده است. در اغلب این پژوهش ها، علاوه بر تشخیص تغییر صحنه، نوع گذار نیز تعیین شده است. تشخیص تغییر صحنه ویدئو به شکل های مختلفی ساخته می شود. دو راه رایج برای تشخیص تغییر صحنه، روش استاتیکی و دینامیکی می باشد [۵]. بطور معمول در تشخیص تغییر صحنه دینامیکی ابتدا مرزهای شات ویدئو با استفاده از روش هایی همچون مقایسه توأم [۶] و با استفاده از ویژگی رنگ تصویر و یا تغییر نرخ لبه های تصویر [۷] و به کارگیری آستانه های وقتی و یا روش های دیگر شناسایی شده و با حذف صحنه های اضافی، ویدئو به قسمت های معنی دار تقسیم می شود. سپس با انتخاب قسمت های چند ثانیه ای از هر صحنه و حذف صحنه های یکسان تغییر صحنه دینامیک ویدئو تشخیص داده می شود [۸].

بطور معمول استخراج فریم کلیدی به دو صورت رایج مبتنی بر متن و مبتنی بر محتوا وجود دارد [۹]. در نوع اول، که در آن متن همراه ویدئو نیز گرفته می شود، کلمه و یا جمله ای دلخواه انتخاب شده، سپس فریمی که متناسب با آن متن است انتخاب می شود شکل (۱). بطور مثال در روش Refaey [۱۰] برای تولید خلاصه ویدئوهای ورزشی و نمایش رخداد های مهم، از متن موجود در صفحات وب برای آن ویدئو استفاده کرده است. در این روش با استفاده از تکنیک

* نویسنده مسئول