



دانشگاه سوادکوه

دانشکده مهندسی

گروه برق

پایان نامه کارشناسی

گرایش: الکترونیک

عنوان :

خوشه بندی شات های دنباله تصاویر ویدئویی

با استفاده از الگوریتم ژنتیکی و تجزیه ماتریسی مستطیلی

استاد راهنما: دکتر امیری

نگارش: هادی حامدی

مرداد ۸۸

فهرست مطالب

چکیده
۱- مقدمه ۱
۲- مروری بر کارهای انجام شده ۸
۲-۱- مقدمه ۸
۲-۲- رونوشت برداری ویدئویی ۹
۲-۳- خلاصه سازی ویدئویی ۱۰
۲-۴- خلاصه سازی ویدئویی براساس استخراج فریم کلیدی بوسیله خوشه بندی Delaunay ۱۵
۲-۵- خلاصه سازی اتوماتیک ویدئویی با استفاده از مدل گراف ۲۴
۲-۶- خلاصه سازی ویدئویی با استفاده از SVD ۳۰
۲-۷- الگوریتم ژنتیکی ۴۱
۲-۸- خوشه بندی ۵۷
۲-۹- ژنتیک الگوریتم و نرم افزار Matlab ۶۵
۳- رهیافت پیشنهادی ۸۳
۳-۱- ایجاد شات های مناسب ویدئویی ۸۳
۳-۲- ایجاد ماتریس خصیصه ی فریم ۸۴
۳-۳- تابع SVD ۸۵
۳-۴- سیستم خوشه بندی بردارهای تکین با الگوریتم ژنتیک ۸۸
۴- نتایج آزمایشات ۹۱
۴-۱- مجموعه داده های استفاده شده ۹۱
۴-۲- نتایج آزمایشات ۹۱
۴-۳- مقایسه با سیستم های قبلی ۹۴
۵- نتیجه گیری و پیشنهادات ۹۹

چکیده

شات در یک ویدئو دنباله ای از تصاویر است که به یکدیگر مرتبط بوده و از لحاظ معنایی کوچکترین جز معنادار ویدئو محسوب می شود. در یک سیستم پردازش ویدئویی برای استخراج اطلاعات از ویدئو لازم است شات های مشابه در گروه های یکسان طبقه بندی شوند تا بتوان یک پایگاه داده زمانی مکانی با مفهوم (معنادار) جهت استخراج اطلاعات از آن ایجاد نمود. در این پروژه هدف خوشه بندی شات های ویدئویی با استفاده از الگوریتم ژنتیکی و تجزیه ماتریسی است. پروژه دارای سه فاز است:

فاز اول تشخیص شات ها: این کار با استفاده از یک الگوریتم ساده مبنی بر تفاضل فریم صورت می گیرد.

فاز دوم استخراج ویژگی ها از شات: بر اساس تجزیه ماتریس های مستطیلی ویژگی های مناسبی از شات ها استخراج می کنیم، سپس این ویژگی ها را برای خوشه بندی استفاده خواهیم کرد.

فاز سوم خوشه بندی شات ها: در این فاز سعی خواهیم کرد با تغییر ساده روی الگوریتم های ژنتیکی، شات های ویدئو را در خوشه های مشابه طبقه بندی کنیم.

۱- مقدمه

تصاویر ویدئویی توزیع شده رایج در سیستم های کامپیوتری و شبکه ها همچنان ها و چالش ها را نمایش می دهند. ویدئو مهیج است زیرا صحنه های دنیای واقعی را واضح تر و پایدارتر بیان می کند. بررسی ویدئو چالش برانگیز است زیرا تصاویر ویدئویی بزرگ دارای افزونگی است و معنی کل آن در یک لحظه به دست نمی آید. از یک مجموعه داده های ویدئویی بزرگ پیدا کردن توالی ویدئویی مناسب و یا تسهیم مطلوب ویدئو یک عمل سخت می باشد. در اینترنت وضعیت بدتر می شود. بر اساس پخش همگانی تبلیغات و یا معرفی محصولات، با پهنای باند بسیار محدود برای اغلب کاربران خانگی، آنها دهها دقیقه برای دانلود تصاویر ویدئویی حجیم صرف می کنند که بعضی از آنها نامربوط می باشند. چرخش مجموعه ویدئویی به منابع اطلاعاتی با ارزش، تکنیک هایی که امکان جستجو براساس مفهوم را می دهند همچنین آسان کردن نمایش سریع نتایج تبدیل به چیزی واجت شده است.

با تعداد زیادی داده های ویدئویی بزرگ، نمایش دادن خلاصه ای از هر فایل ویدئویی به کاربر وظیفه پیدا کردن مورد مطلوب را بسیار آسان می کند. جستجوی مفهوم ویدئو و خلاصه سازی، دو تکنیک اساسی هستند که هر یک دیگر را پیاده سازی می کنند.

یک ساختار محتوای ویدئویی سلسله مراتبی با افزایش تدریجی از صحنه های ویدیویی و گروه های ویدئویی خوشه بندی شده به فریم های کلیدی ساخته می شود. نهایتاً یک طرح خلاصه ویدئویی سلسله مراتبی توسط اجرا کردن چندین خط مشی در سطوح مختلف سلسله مراتب ویدئویی تولید می شود تا بطور ساکن یا متحرک خلاصه ویدئویی ساخته شود.

بعلاوه کاهش هزینه دستگاه های ذخیره سازی، سرعت های انتقال بالاتر و تکنیک های مترکم سازی بهبود یافته، ویدئوهای دیجیتال به صورت افزایشی در دسترس قرار گرفته اند.

بهرحال روشی که در آن محتوای ویدئویی برای عملی مانند جستجو و بازیابی آماده شده، به یک وظیفه مهم چالش انگیز برای سیستم های با کاربرد های ویدئویی و هم برای کاربران می گردد. بعضی روش ها برای آماده سازی محتوای ویدئویی از طریق جستجوی سلسله مراتبی وجود دارد. کاربران می توانند با سرعت از طریق یک صحنه ویدئویی جستجو کنند، از یک قطعه به یک قطعه دیگر هدایت شوند تا یک

گذر سریع کلی از محتوی ویدئویی بگیرند و به سطوح مختلف جزئیات برای موقعیت یابی قطعه های دلخواه تمرکز کنند.

پژوهش ها نشان می دهد که بطور متوسط تقریباً ۲۰۰ شات برای یک ویدئو کلیپ ۳۰ دقیقه ای در میان انواع مختلف برنامه ها، از قبیل اخبار و نمایش وجود دارد. با قرض اینکه یک فریم کلیدی برای نمایاندن هر شات انتخاب شود نمایاندن ۲۰۰ فریم بار عمده را بر حسب پهنای باند و زمان تحمیل خواهد کرد. استفاده از تصاویر تقلیل یافته ویژه که بطور رایج تصاویر Thumbnail شناخته می شوند، چیزی است که می تواند اندازه را کاهش دهد، ولی اگر تمام شات ها برای یک جستجوی سریع محتوی، میبایست نشان داده شوند این روش هنوز گران تمام می شود. بنابراین یک استراتژی خلاصه سازی ویدئویی برای مهیا کردن یک خلاصه فشرده ویدئویی برای کاربران که تنها قسمتهایی از شات را نشان دهد، نیاز خواهد بود.

بطور کلی یک خلاصه ویدئویی بعنوان یک توالی از تصاویر راکد یا متحرک (با یا بدون صدا) تعریف می شود که محتوای یک ویدئو را آنچنان می نمایاند که برای گروه هدف مربوطه، بسرعت اطلاعات مختصر و مفید در مورد محتوا مهیا گردد در حالیکه پیام اساسی اولیه نگه داشته شود.

سه نوع خلاصه ویدئویی بطور معمول استفاده می گردد:

خلاصه تجسمی: یک مجموعه از تصاویر بی حرکت (ممکن است در اندازه متغیر باشند) است که بر حسب زمان مرتب شده اند. استفاده از تصاویر فشرده به کاربر این امکان را می دهد که یک دید کلی از ویدئو بطور مستقیم و کارآمد داشته باشد هرچند که اطلاعات صدا و حرکت ویدئو از دست می رود.

نمونه گیری ویدئو: یک مجموعه از کلیپ های ویدئویی (شات های ویدئویی) که در رده های زمانی مرتب شده اند. در مقایسه با خلاصه های تصویری بدون حرکت، نمونه گیری ویدئویی (کف گیری ویدئویی) برای کاربران بسیار جالب تر است زیرا که هم حرکت و هم صدا باقی نگه داشته میشوند. هرچند مقدار زمان لازم برای از نظر گذراندن یک نمونه گیری اظهار می دارد که یک ویدئوی نمونه گیری شده برای جستجوی سریع مناسب نمیشود.

نقشه توزیع اطلاعات: یک تصویر ترکیبی است که توزیع بعضی از اطلاعات معین را در پایگاه داده نشان می دهد. به سبب پیچیدگی ذاتی تصاویر و ویدئوها، استفاده از فریم های تصویری یا نمونه گیری ویدئویی، برای خلاصه کردن یک پایگاه داده ویدئویی ممکن است غیر عملی باشد. از اینرو نقشه توزیع داده، به نظر یک مکانیسم موثر برای تدارک یک دید کلی از تمام مجموعه داده ها (دیتاست) برای

کاربران می باشد. متأسفانه چنان نقشه توزیعی ممکن است به اندازه فریم های تصویری یا نمونه گیری قابل درک و حس نباشد، چونکه نقاط یا منحنی های ترکیبی امکان دارد برای کاربران بی تجربه مفهومی را نرساند.

وقتی برای کاربران یک خلاصه ویدئویی فشرده فراهم می شود آنها می توانند محتوای ویدئویی را با سرعت و جامع بدست آورند. علاوه بر این برتری خلاصه های بصری می تواند در بسیاری کاربردها مانند بایگانی های چند رسانه ای ، بازیابی ویدئو ، سرگرمی های خانگی و مجله های دیجیتال کمک کننده باشد.

موتورهای جستجوی ویدئویی یک مجموعه از ویدئوها را با معیاری معین برمیگرداند و خلاصه ساز های ویدئویی مفهوم ویدئویی خلاصه سازی شده را تولید می کند تا کاربران را قادر به فهم سریع معنی کل هر ویدئو بکند. موتور جستجو با یک فیلتر اطلاعاتی خاصی ویدئو های مناسب را از پایگاه داده برمی گرداند در حالی که یک خلاصه ساز ویدئویی کاشف اطلاعات می باشد تا کاربر بتواند ویدئوی مناسب را انتخاب کند. امروزه خلاصه سازی ویدئویی اساساً با استخراج فریم های کلیدی از ویدئوی اصلی و نمایش بخشی از آنها در یک صفحه انجام می شود. از معایب این روش این است که فریم های کلیدی مجموعه ای از تصاویر ایستا می باشند و وابسته به فضا و زمان نیستند. فرق یک برنامه ویدئویی با یک تصویر ویدئویی این است که برنامه ویدئویی نمایش پویا وابسته به زمان و فضا از صفحه دارد. یک مجموعه فریم های کلیدی ایستا بدون این خصوصیات یک ویدئو ، نمایش ضعیفی خواهد بود.

امروزه خلاصه سازی ویدئویی با استفاده از فریم های کلیدی استخراج شده از توالی ویدئویی اصلی انجام می شود. انرژی زیادی روی شکستن ویدئو به شات ها و سپس پیدا کردن تعداد ثابتی از فریم های کلیدی در هر شات متمرکز می شود. برخی از دانشمندان راه های مختلفی را برای استخراج فریم های کلیدی بکار می برند. (اولین فریم هر شات و ...)

یک عیب آشکار تعیین تعداد ثابت فریم کلیدی این است که شات های طولانی و شات های کوتاه بطور یکسان خلاصه سازی ما شوند. برای حل این مشکل آقای DeMenthon پیشنهاد کرد که تعداد فریم های متغیر برای هر شات را بر اساس درجه فعالیت آن شات در صحنه مورد نظر ، در نظر بگیریم. این متد یک توالی ویدئویی بعنوان یک مسیر منحنی در یک فضای ویژه پر اهمیت را نشان می دهد و از یک الگوریتم جداسازی منحنی دودویی بازگشتی برای پیدا کردن یک مجموعه از نقاط ادراکی مهم برای تخمین منحنی ویدئویی استفاده می کند. این تقریب ادامه پیدا می کند تا زمانی که خطای تخمین به زیر مقدار مشخص شده توسط کاربر برسد. فریم های حاصل از این نقاط ادراکی به عنوان فریم کلیدی

گروه برق آندامگاه پژوهش برق و الکترونیک دانشگاه زنجان و الکترونیک مهندسی گروه برق آندامگاه پژوهش برق و الکترونیک دانشگاه زنجان

برای خلاصه سازی استفاده می شود. این الگوریتم نقاط بیشتری را برای انحنای بیشتر تعیین می کند. این متد طبعاً فریم های کلیدی بیشتری برای شات های با تغییرات زیادتر تولید می کند.

فریم های استخراج شده ممکن است دارای تکرار و افزونگی باشند. مانند صحبت یک مجری تلویزیون، بکار بردن الگوریتم بالا در این شرایط ممکن است فریم هایی را استخراج کند که تعداد زیادی از آنها شبیه و یکسان هستند. برای حذف این افزونگی Yeung از هر شات یک فریم را انتخاب نموده، سپس عملیات کلاستر بندی سلسله مراتب را بر اساس شباهت و فاصله زمانی روی این فریم ها اجرا کرده، از هر کلاستر یک فریم را استخراج می کند. Boreczky و Girsensohn از تکنیک کلاستر بندی سلسله مراتب برای گروه بندی فریم های کلیدی به تعدادی کلاستر که تعداد آن توسط کاربر مشخص می شود استفاده کردند. برای هر کلاستر فریم های کلیدی انتخاب شده محدودیت هایی از قبیل توزیع فریم های کلیدی در طول ویدئو و حداقل فاصله بین فریم ها با هم تلاقی دارند.

در روش اخیر خلاصه سازی بر اساس کلاستر بندی، Dirk Farin پیشنهاد کرد که دامنه دانش را برای حذف فریم های انتخابی از شات های نامربوط یا غیر جالب به هم پیوندیم. به عنوان مثال، شاتهایی که محو می شوند. دانش دامنه ای به پردازش کلاستری پیوند داده می شود و این کار با نمایش بردارهای خصیصه صحنه های غیر جذاب به عنوان مراکز کلاستر اضافی از اینکه بعنوان فریم کلیدی انتخاب شوند، محروم می شوند.

تمام مقدمه های گفته شده، دارای محدودیت هایی می باشند. البته ماهیت تصاویر ویدئویی با عکس فرق می کند. یک مجموعه از فریم های کلیدی ایستا که این خصیصیات اساسی ویدئو را برآورده نمی کند، یک نمایش ضعیف از مفهوم کلی تصویری برنامه ویدئویی بدست می دهد. برای کاهش فقدان مفهوم زمانی-فضایی سبب شده توسط نمایش فریم های کلیدی، تکنیک به هم پیوستن تصاویر برای ساختن یک دید کلی از صحنه ضبط شده ی ورودی بوسیله هر شات بکار گرفته شود. برای یک شات ویدئویی داده شده، بکه های به هم پیوسته آن با پیچ و تاب دادن همه فریم های آن در یک سیستم و سپس بخیه زدن آنها به یکدیگر ایجاد می شود تا پس زمینه ی کامل برای هر فریم در تمام شات دوباره ساخته شود. اگرچه یک mosaic فقط هنگامی می تواند بصورت موفقیت آمیز ساخته شود که صحنه هدف دارای چندین شرط باشد: از قبیل اینکه صحنه نباید برجستگی های ساختار 3D را داشته باشد، نباید شامل

اشیا در حال حرکت باشد و غیره. این قبیل محدودیت ها یقیناً روی عملیات خلاصه سازی تاثیر منفی می گذارند.

مندی گروه برق آناگاه پژوهش برق دانشگاه زنجان، دانشکده مهندسی گروه برق آناگاه پژوهش برق دانشگاه زنجان

مندی گروه برق آناگاه پژوهش برق دانشگاه زنجان، دانشکده مهندسی گروه برق آناگاه پژوهش برق دانشگاه زنجان

تحقیقاتی انجام شده است که تلاش می کند تا خروجی خلاصه سازی ویدئویی با دیدگاه معنایی آن بهتر وفق داده شود. Cuevideo نام سیستمی است که توسط IBM ارائه شده است و روش کار آن به این صورت می باشد که تابع playback ویدئویی آن شات های ایستا و طولانی را با سرعت بیشتر (frame rate) و شات های کوتاه را با سرعت کمتر اجرا می کند. البته این سرعت متغیر سبب می شود که شات های ایستا بیشتر از شات های پویا و شات های پویا بیشتر از شات های ایستا نمایش داده شوند. بنابراین خصوصیات زمانی توالی ویدئویی تحریف نمایی می شود. به عبارت دیگر سیستم اطلاعاتی CMU، رها سازی ویدئویی را فراهم می کند که تلاش می کند تا بخش های تصویری ای که از نظر معنایی مهم می باشند را شناسایی و اجرا کند که این کار وابسته به عبارات صوتی برای فهم معنایی در یک توالی ویدئویی می باشند. هر بخش تصویری مهم با استفاده از یک مجموعه از قوانین اکتشافی که تا حد زیادی موضوعی و معنی دار می باشند اندازه گیری می شود.

این سیستم خلاصه سازی بر اساس قانون، یقیناً محدودیت هایی را در بررسی متنوع تصاویر ویدئویی ایجاد کرده است.

در سال های اخیر خلاصه سازی مفهومی بازی های ورزشی توجه محققان را در انجمن های رسانه ای به خود جلب کرده است. یک فیلم ورزشی دارای ساختار داخلی خوب می باشد که بصورت مناسب تعریف شده است از قبیل اینکه همه بازی ها بر اساس قانون بوده و قابل پیشگویی می باشند. سیستم های خلاصه سازی فیلم های ورزشی سطوح بالاتری از خلاصه سازی (فشرده سازی) مفهومی را با پیگیری خصوصیات دامنه ای خاص و دانش می توانند بدست دهند. Gong یک سیستم تحلیل بازی فوتبال را طراحی کرد که در آن هر شات ویدئویی فوتبال را بر اساس خصوصیات زمین فوتبال، موقعیت بازیکنان و حرکت آنها و حضور یا عدم حضور توپ برای سازماندهی گروه بندی هر شات صحنه استفاده شد. Xu پیشنهاد کرد که هر بازی فوتبال از دو حالت بازی یا توقف تشکیل شده است و سیستمی را طراحی کرد که یک فیلم فوتبال را به دو بخش بازی و توقف بر اساس سرعت چمن در فریم های ویدئویی تقسیم بندی کرد.

Chang و Zhong تمرکز خود را روی این حقیقت متمرکز کردند که تمام نقاط اوج در بازی بیس بال و یا تنیس زمانی می باشد که توپ توسط بازیکن زده می شود و گام های بعد آن می باشد و تلاش کردند که این صحنه ها را با گروه بندی هیستوگرام های رنگی فریم های کلیدی شات های مربوط به صحنه، در این دو نوع بازی آشکار سازی کنند (در تنیس زدن توپ و در بیس بال اوج توپ را آشکار سازی کردند).

Rui فرض کرد که صحنه های مهیج در بیس بال با نوع صحبت گوینده رابطه دارد که اغلب پس از زدن

ضربه آشکار می شود. بر اساس این فرضیات او نقاط بازی بیس بال را با استفاده از تحلیل صدای گوینده آشکار کرد.

برای دست یافتن به یک خلاصه ویدئویی مناسب برای کاربردهای گوناگون نام برده شده ، ما یک استراتژی خوشه بندی شات های ویدئو را معرفی می کنیم که از مراکز خوشه ها برای خلاصه سازی و دسته بندی شات ها استفاده می کند. در این روش ابتدا توالی های ویدئویی مناسب را از ویدئوهای مختلف انتخاب می کنیم ، سپس با استفاده از یک الگوریتم استخراج ویژگی، فریم های ویدئویی را با استفاده از پردازش خاصی به بردارهای ستونی تبدیل کرده و سپس با الحاق این بردارها به یکدیگر، ماتریسی بدست می آید که ماتریس خصیصه فریم نامیده می شود. ماتریس ویژگی را برای هر شات استخراج کرده و در مرحله بعد با اجرای دستور svd بر روی ماتریس خصیصه فریم، و در نهایت از طریق اعمال خوشه بندی توسط الگوریتم ژنتیک، مراکز خوشه های بهینه را بدست می آوریم.

در ادامه چند مفهوم ویدئویی و تعاریف آنها ذکر می شود:

یک فیلم ویدئویی معمولاً شامل چندین صحنه و هر صحنه شامل یک یا چندین شات می باشد.

تعریف ۱: یک طرح خلاصه سازی استاتیک، یک روشی است که نیازی به مشارکت کاربر برای تولید خلاصه ندارد، برای تمام کاربران یک ویدئو، این استراتژی یک خلاصه را ایجاد خواهد کرد...

تعریف ۲: یک طرح خلاصه سازی دینامیک، یک روشی است که اعمال کاربر ممکن است منجر به تغییراتی در خلاصه ایجاد شده گردد...

تعریف ۳: ساختار محتوای ویدئو، بعنوان یک سلسله مراتب از صحنه های خوشه بندی شده، صحنه های ویدئویی، گروه های ویدئویی، و شات های ویدئویی (فریم های کلیدی)، که بصورت تدریجی از بالا به پایین در آدرس دهی محتوای ویدئویی افزایش می یابند، تعریف می شود. یک ساختار محتوا می تواند در اکثر ویدئوها در زندگی روزانه ما پیدا شود.

تعریف ۴: یک شات ویدئو عنصر اساسی در ویدئوها است. آن فریم هایی را که از یک حرکت پیوسته واحد دوربین، از لحظه ای که شروع می شود تا لحظه ای که خاتمه پیدا می کند، ثبت می کند. در واقع بخش یک تکه و مداوم توالی فریم های ویدئویی با حرکت دوربین به صورت پیوسته و یا ایستا می باشد.

تعریف ۵: یک فریم کلیدی، فریمی است که نماینده برای توصیف کردن محتوای شات داده شده است. بسته به تغییرات محتوای شات، یک یا چندین فریم کلیدی می تواند استخراج گردد.

تعریف ۶: یک گروه بزرگ ویدئویی، شامل شات‌های مشابه از نظر بصری در تمام طول یک ویدئو است.

تعریف ۷: یک گروه ویدئویی، یک نهاد واسطه بین شات‌های فیزیکی و صحنه‌های معنایی است و شامل شات‌های مشابه از لحاظ بصری است که به یک واحد گزارشی یکسان متعلق باشند.

تعریف ۸: یک صحنه ویدئویی، یک مجموعه از گروه‌های وابسته از لحاظ معنایی و موقتا مجاور است که یک مفهوم یا گزارش سطح بالا را مجسم می‌کند و انتقال می‌دهد.

تعریف ۹: یک صحنه‌ی خوشه بندی شده، یک مجموعه از صحنه‌های ویدئویی مشابه از لحاظ بصری است.

سکده مهندسی گروه بین‌رشته‌ای پژوهش‌های انرژی‌های تجدیدپذیر و انرژی‌های نو (گروه بین‌رشته‌ای پژوهش‌های انرژی‌های تجدیدپذیر و انرژی‌های نو) دانشگاه تهران
سکده مهندسی گروه برق و انرژی‌های تجدیدپذیر و انرژی‌های نو (گروه بین‌رشته‌ای پژوهش‌های انرژی‌های تجدیدپذیر و انرژی‌های نو) دانشگاه تهران
گروه برق و انرژی‌های تجدیدپذیر و انرژی‌های نو (گروه بین‌رشته‌ای پژوهش‌های انرژی‌های تجدیدپذیر و انرژی‌های نو) دانشگاه تهران

دانشجویان محترم:

جهت دسترسی به متن کامل پایان نامه‌ها به کتابخانه دانشکده مهندسی و یا آزمایشگاه پروژه گروه برق مراجعه فرمایید.

مکان نام کارشناسی

۵- نتیجه گیری و پیشنهادات

۵-۱- نتیجه گیری

کاهش بعد یک پردازش داده‌ی ضروری برای پایگاه داده‌ها و کاربرد های استخراج داده از قبیل حافظه موثر و بازیابی داده‌های با ابعاد بالا می‌باشد.

در این پروژه روشی نو برای خلاصه سازی ارائه شد که بر پایه‌ی خصوصیات یکتای به دست آمده از تجزیه‌ی بردارهای تکین ماتریسی می‌باشد. با استفاده از تحلیل ریاضی توانستیم ماتریس حجیم خصیصه فریم به دست آمده از توالی ویدئویی را به یک آرایه تبدیل کنیم و همچنین یک بردار ستونی با ابعاد بسیار زیاد را به یک و تنها یک عدد تبدیل نماییم و در نتیجه با این تبدیلات توانستیم حجم داده‌های به دست آمده از توالی ویدئویی را به مقدار بسیار زیادی کاهش دهیم.

کاهش حجم اطلاعات به دست آمده ناشی از اعمال تجزیه‌ی بردارهای تکین SVD بر روی داده‌های استخراج شده از توالی ویدئویی باعث افزایش سرعت پردازش و همچنین سادگی در کارها شده است. پس از این می‌توان به راحتی و با استفاده از روش گفته شده حجم داده‌ها را برای افزایش پردازش و سادگی کار کاهش داد.

در بسیاری از برنامه‌های کاربردی، به دست آوردن نمونه‌ای که نماینده داده‌ها باشد بسیار گران و وقت گیر می‌باشد.

سیستم خلاصه سازی و خوشه بندی ویدئو، که با استفاده از روش SVD طراحی شد، دارای سرعت عمل بیشتر در مقایسه با سیستمهای مشابه می‌باشد و همچنین در بخش خوشه بندی ویدئو نیز دارای دقت

عمل خوبی می‌باشد. به این معنا که در فضای داده‌ای وسیع، سریعتر به جواب همگرا می‌شود.

