







دانشکده مهندسی کامپیوتر و فناوری اطلاعات

گروه هوش مصنوعی

## افزایش سرعت انتقال ویدیو در سیستم‌های آموزش از راه دور

علیرضا پاشا محمد نوری

استاد راهنما :

دکتر مرتضی زاهدی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

تابستان ۹۳



تقدیم بہ پدرم کہ صبر نہ کر د...

## تقدیر و شکر

بیچ حادثہ ای بی علت نیست و بیچ و قوعی بی دلیل. هر انسان در دلش آرزوهای بی شماری دارد که در گذر زمان ہمت آدمی بہ آنان جامہ سی عل می پوشاند. در این مسیر بیچ کس نمی تواند ادعا کند کہ بہ تنهایی بہ قله ای رسیده است. می دانم کہ تا قله مسیر سی طولانی در پیش است اما بہ یاد نیز دارم مکانی را کہ حال در آن ایستاده ام برایم روزی قلمگاہی بوده است. در قلم نہ قدرت بیان است و نہ تاب نوشتار الطاف بزرگان بہ من ناچیزہ خودم را موقوف می دانم تا آخرین دم زندگی، خود را میدیون و وام دار این عزیزان بدانم. پدرم کہ مرادوست می داشت و حسرت با او بودن تا بد در دلم مانده است. ہستی ام را میدیون بزرگواری و فداکاری ہایش می دانم. او برایم ہمہ چیز بوده و ہست و خواهد بود.

استاد بزرگوaram؛ جناب آقای دکتر علی اکبر پویان، سخنوری دانشمند، انسانی فرہنگی کہ ہمیشہ شیفتہ سی کلاشان می باشم. جناب آقای دکتر حمید حسن پور، محقق تلاسگر، اسادی دلوز کہ خود را شرمندہ ایشان می دانم. جناب آقای دکتر مرتضی زہدی، اسادی کہ مفہیم سخت و دشوار را، بچون جرعہ ای کواری ما نوشتانید. در طول مدت تحصیل ہمیشہ در کنار ما بود و راہنمایی ہایشان مسایل پیچیدہ را شرمسار دایت این عزیز می نمود.

من لم یسکر المخلوق لم یسکر الخالق

علیرضا پاشا محمد نوری

تابستان ۱۳۹۳



## چکیده

فشرده سازی تصویرهای ویدیویی یکی از مهمترین حوزه‌های فعالیت‌های تحقیقاتی و تجاری در علوم کامپیوتر می‌باشد. با ورود تصاویر ویدیویی در بستر وب، به علت محدود بودن پهنای باند ارتباطی، نیاز روز افزونی به فشرده سازی تصاویر ویدیویی احساس گردید. یکی از مهمترین حوزه‌های نقل و انتقال تصاویر ویدیویی بر بستر وب را می‌توان ویدیوهای آموزشی در حوزه‌ی آموزش از راه دور در نظر گرفت. این ویدیوها یا بصورت برخط و یا بصورت ضبط شده در بستر وب انتقال پیدا می‌کنند. با توجه به محدود بودن پهنای باند و حجم بالای تصاویر ویدیویی، فشرده سازی اینگونه تصاویر بسیار مورد توجه بوده است. با کاهش حجم یک تصویر ویدیویی می‌توان آن را براحتی بر پهنای وب انتقال داد و یا پهنای آزاد شده را به صوت و یا اطلاعات اضافی و مفید دیگری اختصاص داد. امروزه در حوزه‌ی فشرده سازی تصاویر ویدیویی رقابت بسیار شدیدی در ارائه‌ی کدک‌های جدیدتر که تصاویر را با حجم کمتر و کیفیت بالاتری فشرده می‌کنند وجود دارد. در این پایان نامه سعی شده است بنا به ماهیت تصاویر ویدیویی آموزش از راه دور که در حوزه‌ی تصاویر کم تحرک قرار می‌گیرند، یک کدک فشرده سازی ارائه گردد تا این تصاویر را با نرخ بالاتری فشرده سازی نمایند و به انتقال بهتر آن در پهنای وب کمک نمایند. این کدک در ابتدا سعی دارد تا تمامی افزونگی‌های زمانی یک ویدیو را حذف نماید. سپس با تقسیم بندی هر فریم تصویر بنا به موقعیت تغییرات تصویر افزونگی‌های مکانی تصویر را حذف نماید. این عمل باعث افزایش نرخ فشرده‌سازی ویدیو بدون آنکه به پیکسل‌های ارزشمند تصویر آسیبی وارد نماید و موجب کاهش کیفیت شود را ایجاد می‌نماید. می‌توان ادعا نمود فشرده سازی تصاویر ویدیویی آموزش از راه دور با این روش بسیار بهتر از فشرده سازی این تصاویر با فرمت‌های فشرده سازی دیگر می‌باشد که برای ویدیوهای پرتحرک نیز کاربرد دارند. هدف اصلی این پایان نامه ارائه روشی منحصر بفرد برای تصاویر کم تحرک آموزشی می‌باشد.



ادعا می‌شود این کدک برای اینگونه تصاویر ویدیویی از نرخ فشرده سازی بالاتری برخوردار بوده و این فشرده سازی، پیکسل های با اهمیت تصویر را دچار تغییر نخواهد نمود.

**کلمات کلیدی:** فشرده سازی ویدیو، آموزش از راه دور، انتقال ویدیو، هوش مصنوعی، کدک

**های جدید**

## لیست مقالات مستخرج از پایان نامه

1- Morteza zahedi, Alireza p. Nouri, "Introduce a new low motion video database,"  
*International Journal of Computer Application*, 2013.

## فهرست

فصل اول مقدمه .....	۲
۱-۱ مقدمه .....	۲
۲-۱ آشنایی با مفهوم ویدیو .....	۴
۳-۱ ضرورت فشرده سازی .....	۵
۴-۱ تعریف فشرده سازی .....	۷
۵-۱ تحرک در تصاویر ویدیویی .....	۸
۶-۱ کیفیت در فشرده سازی .....	۸
۱-۶-۱ اتلاف در فشرده سازی .....	۹
۲-۶-۱ معیارهای ارزیابی کیفیت .....	۱۱
۱-۲-۶-۱ معیار میانگین مربع خطا MSE .....	۱۱
۲-۲-۶-۱ معیار میانگین قدر مطلق خطا MAE .....	۱۲
۷-۱ ساختار بینایی انسان .....	۱۳
۸-۱ مدل‌های رنگی .....	۱۴
۱-۸-۱ مدل رنگی RGB .....	۱۴
۲-۸-۱ مدل رنگی Ycbcr .....	۱۶
۱-۲-۸-۱ فرمت های متفاوت نمونه برداری Ycbcr .....	۱۷
۹-۱ ساختار پایان نامه .....	۱۹
فصل دوم مروری بر فعالیت های انجام شده (روش ها و استانداردهای فشرده سازی) .....	۲۱
۱-۲ مقدمه .....	۳۶

۳۷	..... روشهای فشرده سازی
۳۸	..... ۱-۲-۲ افزونگی مکانی
۴۲	..... ۲-۲-۲ افزونگی زمانی
۴۳	..... ۱-۲-۲-۲ کدگذاری تفاضلی
۴۴	..... ۲-۲-۲-۲ تخمین حرکت و جبران آن بر پایه ی بلوک بندی
۴۷	..... ۳-۲-۲-۲ تخمین حرکت و جبران دو طرفه
۴۸	..... ۳-۲-۲ کدگذاری آنتروپی
۴۹	..... ۱-۳-۲-۲ کدگذاری هافمن
۵۱	..... LZ کدگذاری
۵۲	..... ۳-۲ مدل های کدک ویدئو
۵۳	..... ۴-۲ سازمانها و استانداردها
۵۴	..... ۱-۴-۲ سازمان ITU-T
۵۷	..... ۲-۴-۲ سازمان ISO MPEG
۵۷	..... MPEG-1 ۱-۲-۴-۲
۵۸	..... MPEG-2 ۲-۲-۴-۲
۶۰	..... MPEG-4 ۳-۲-۴-۲
۶۱	..... MPEG-7 ۴-۲-۴-۲
۶۲	..... JVT ۳-۴-۲
۶۴	..... ۱-۳ فصل سوم روش پیشنهادی
۶۵	..... ۲-۳ پیش پردازش
۶۵	..... ۱-۲-۳ تبدیل مدل رنگی RGB به مدل رنگی Ycbcr
۶۷	..... ۲-۲-۳

۷۲	..... ۱-۲-۲-۳ پیدا کردن حدآستانه برای سکانس بندی
۷۴	..... ۳-۲-۳ نتیجه گیری
۷۴	..... ۳-۳ جایگزینی سکانسهای بدون تغییر
۷۵	..... ۱-۳-۳ کدک شماره یک
۷۸	..... ۱-۴-۳ گریدبندی با سایز پنجره های یکسان
۸۰	..... ۲-۴-۳ محاسبهی سایز پنجره های گریدبندی
۸۰	..... ۳-۴-۳ محاسبه تغییرات در بلوکهای یک سکانس
۸۲	..... ۴-۴-۳ نتیجه گیری
۸۳	..... ۵-۳ تابع تخریب
۸۴	..... ۱-۵-۳ معرفی تابع تخریب
۸۷	..... ۶-۳ کدکهای زیربلوکهای دارای تحرک
۸۷	..... ۱-۶-۳ کدک شماره ی دو
۸۸	..... ۲-۶-۳ کدک شماره ی سه
۹۰	..... فصل چهارم محاسبه و نتایج
۹۰	..... ۱-۴ پایگاه داده تصاویر ویدیویی آموزش از راه دور
۹۳	..... ۲-۴ محاسبه عملکرد روش پیشنهادی بر روی پایگاه داده
۹۵	..... ۳-۴ استخراج بردار میانگین روشنایی تصویر
۹۹	..... ۴-۴ سکانس بندی تصاویر ویدیویی
۱۰۰	..... ۱-۴-۴ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۱
۱۰۱	..... ۲-۴-۴ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۳
۱۰۲	..... ۳-۴-۴ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۶
۱۰۳	..... ۳-۴-۴ نتیجه گیری

۱۰۴.....	۵-۴ ارزیابی کدک شماره ۱.....
۱۰۷.....	۴-۵-۱ نتیجه گیری.....
۱۰۸.....	۴-۶-۶ گرید بندی سکانس ها.....
۱۱۰.....	۴-۷-۷ تابع تخریب.....
۱۱۵.....	۴-۸-۸ ارزیابی نرخ فشرده سازی تصاویر ویدیویی.....
۱۱۷.....	۴-۹-۹ ارزیابی کیفیت تصاویر ویدیویی قبل و بعد از اعمال فشرده سازی.....
۱۱۹.....	۴-۹-۱-۱ نتیجه گیری.....
۱۲۰.....	فصل پنجم نتیجه گیری و کارهای آتی.....
۱۲۰.....	۵-۱-۱ نتیجه گیری.....
۱۲۲.....	۵-۲-۲ کارهای آتی.....
۱.....	مراجع.....

## فهرست شکل‌ها

- شکل ۱-۲: تصویر (a) فشرده سازی با ۸۴ درصد کمتر از اطلاعات تصویر اصلی، تصویر (b) فشرده سازی با ۹۲ درصد کمتر از اطلاعات تصویر اصلی، تصویر (c) فشرده سازی با ۹۸ درصد کمتر از اطلاعات تصویر اصلی ..... ۱۰
- شکل ۱-۳: پراکندگی رنگ‌ها در استاندارد RGB ..... ۱۵
- شکل ۱-۴: مدل پراکندگی رنگ‌ها در Ycber ..... ۱۶
- شکل ۱-۵: نوع فرمت ۴:۴:۴ ..... ۱۷
- شکل ۱-۶: نوع فرمت ۴:۲:۲ ..... ۱۸
- شکل ۱-۷: نوع فرمت ۴:۱:۱ ..... ۱۸
- شکل ۱-۸: نوع فرمت ۴:۲:۰ ..... ۱۹
- شکل ۱-۲: نمای کلی فشرده سازی یک فایل ویدیویی ..... ۳۸
- شکل ۲-۲: میزان همبستگی دو فریم مجاور ..... ۳۹
- شکل ۳-۲: نحوه پیشبینی مقادیر پیکسلها ..... ۴۰
- شکل ۴-۲: نحوه ی پیمایش ضرایب DCT برای بهینه کردن فشرده سازی ..... ۴۲
- شکل ۵-۲: مقدار تفاضلی دو فریم متوالی ..... ۴۳
- شکل ۶-۲: تصویر a, b دو صحنه ی متوالی شطرنج می باشند تصویر c نداشت صفحه بر روی بلوک‌ها می باشد و تصویر d بردار حرکت بلوک‌های شطرنج را نمایش میدهد ..... ۴۴
- شکل ۷-۲: تصویر بردارهای حرکت در یک فریم که میزان جابجایی هر بلوک را نسبت به فریم بعدی نشان میدهد ..... ۴۵
- شکل ۸-۲: فضای جستجوی یک فریم برای پیدا کردن بلوک مشابه ..... ۴۶
- شکل ۹-۲: روش تخمین حرکت و جبران در ۳ فریم متوالی ..... ۴۸

- شکل ۲-۱۰: نحوه ی کدگذاری با استفاده از درخت هافمن ..... ۵۰
- شکل ۲-۱۱: روش کار الگوریتم LZ ..... ۵۲
- شکل ۲-۱۲: نحوه ی کار یک کدگذار و کدگشا ..... ۵۳
- ..... ۵۴
- شکل ۲-۱۳: دیاگرام فرمت فشرده سازی H.261 ..... ۵۴
- شکل ۲-۱۴: دیاگرام فرمت فشرده سازی H.263 ..... ۵۵
- شکل ۲-۱۵: دیاگرام فرمت فشرده سازی MPEG-1 ..... ۵۸
- شکل ۲-۱۶: دیاگرام فرمت فشرده سازی MPEG-2 ..... ۵۹
- شکل ۲-۱۷: دیاگرام فرمت فشرده سازی MPEG-4 ..... ۶۰
- شکل ۲-۱۸: نحوه ی توسعه ی استانداردها در ۴ دههی اخیر ..... ۶۲
- شکل ۳-۱: از راست به چپ، تصویر مولفه قرمز تصویر مولفه سبز و تصویر مولفه آبی از یک تصویر در مدل RGB ..... ۶۷
- شکل ۳-۲: از راست به چپ، تصویر مولفه روشنایی Y و تصویر مولفه رنگی cb و تصویر مولفه رنگی cr در مدل رنگی ..... ۶۷
- شکل ۳-۳: بردار مقادیر میانگین روشنایی ۷۰۰۰ فریم متوالی ..... ۷۱
- شکل ۳-۵: معیار حدآستانه بر روی بردار تفاضلی میانگین روشنایی فریمها ..... ۷۳
- شکل ۳-۶: دیاگرام مراحل پیش پردازشی برای فشردهسازی تصویر بر اساس روش پیشنهادی ..... ۷۴
- شکل ۳-۷: دیاگرام کدک شماره یک ..... ۷۷
- شکل ۳-۸: تصویر سمت راست گرید بندی با طول پنجرهی برابر و تصویر سمت چپ گرید بندی با طول پنجرهی متفاوت ..... ۷۹
- شکل ۳-۹: دیاگرام گریدبندی سکانسهای پرتحرک و ایجاد بلوک و زیربلوکهای هر گرید ..... ۸۲
- شکل ۳-۱۰: نحوه ی محاسبهی ماتریس تغییرات یک فریم ..... ۸۵



- شکل ۳-۱۱: سمت راست نحوه‌ی تمییم مقادیر بلوکهای حاشیه‌های و سمت چپ ماتریس تابع تخریب پس از محاسبه ..... ۸۶
- شکل ۴-۱: چهار فریم متفاوت از یک فایل ویدیویی پایگاه داده ..... ۹۲
- شکل ۴-۲: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۱ ..... ۹۵
- شکل ۴-۳: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۲ ..... ۹۶
- شکل ۴-۴: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۳ ..... ۹۶
- شکل ۴-۵: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۴ ..... ۹۶
- شکل ۴-۶: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۵ ..... ۹۷
- شکل ۴-۷: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۶ ..... ۹۷
- شکل ۴-۸: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۷ ..... ۹۷
- شکل ۴-۹: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۸ ..... ۹۸
- شکل ۴-۱۰: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۹ ..... ۹۸
- شکل ۴-۱۱: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۱۰ ..... ۹۸
- شکل ۴-۱۲: سکانس بندی ویدیوی شماره ۱ ..... ۱۰۰
- شکل ۴-۱۳: سکانس بندی ویدیوی شماره ۳ ..... ۱۰۱
- شکل ۴-۱۴: سکانس بندی ویدیوی شماره ۶ ..... ۱۰۲
- شکل ۴-۱۵: مقایسه حجم ویدیوها با فرمت های گوناگون ..... **Error! Bookmark not defined.**
- شکل ۴-۱۵: مقایسه حجم نهایی ویدیو با فرمت های گوناگون ..... ۱۱۷

## فهرست جداول

- جدول ۱-۱: درجه وضوح استانداردهای متفاوت تصاویر ویدیویی ..... ۴
- جدول ۱-۲: نرخ فریم و نرخ ارسال استانداردهای تصویر ..... ۵
- جدول ۱-۴: جزئیات تصاویر ویدیویی پایگاه داده ..... ۹۲
- جدول ۲-۴: مشخصات ویدوی های پایگاه داده ها ..... ۹۴
- جدول ۲-۴: میزان تغییرات هر سکانس ویدیوها ..... ۱۰۴
- جدول ۳-۴: حجم تصاویر ویدیویی پس از اعمال کدک ۱ بر روی سکانس ها ..... ۱۰۶
- جدول ۴-۴: تغییرات روشنایی در گریدهای سکانس ۲ از ویدیویی شماره ۱ ( یک سکانس پر تحرک) ۱۰۹
- جدول ۵-۴: تغییرات روشنایی در گریدهای سکانس ۱ از ویدیویی شماره ۳ ( یک سکانس کم تحرک) ۱۱۰
- جدول ۶-۴: مقادیر تابع تخریب برای سکانس ۲ از ویدیوی شماره ۱ ..... ۱۱۱
- جدول ۷-۴: تاثیر تابع تخریب بر حجم گریدها ..... ۱۱۲
- جدول ۸-۴: تاثیر متد فشرده سازی بر حجم سکانس ها ..... ۱۱۴
- جدول ۹-۴: مقایسه نرخ فشرده سازی فرمت های گوناگون ..... ۱۱۵

# فصل اول:

مقدمه

## فصل اول

### ۱-۱ مقدمه

آموزش از راه دور روشی برای تعلیم افراد با استفاده از تکنولوژی و بسترهای ارتباطی مخابراتی می‌باشد. در این روش بین فرد تدریس کننده و فرد دانش پژوه زمان و مکان تاثیر بسزایی ندارد. در بعضی مواقع بصورت همزمان<sup>۱</sup> و گاهی مواقع بصورت با تاخیر<sup>۲</sup> کلاسها برگزار می‌شود. در بعضی افراد فواصل چند کیلومتری بوده و برای بعضی کاربردهای دیگر فواصل چند هزار کیلومتری متصور است.

همانطور که گفته شد این روش تدریس و انتقال مفاهیم با توجه به عدم وابستگی به زمان و مکان و عدم تاثیر ظرفیت کلاسها بر کارایی و همچنین هزینه‌ی بسیار پایین در دهه‌های اخیر بسیار مورد توجه قرار گرفته است. اکثر دانشگاه‌ها واحدهای درسی را بدین گونه ارائه داده و در انتها نیز مدارک معتبری را به افراد ارائه می‌کنند. این سیستم نوین آموزشی در کنار مزایای بی‌شمار خود که توانسته است آموزش را در هر مکان جغرافیایی با هزینه‌ای بسیار کم برای دانش پژوهان فراهم نماید، نیازمندیهای اولیه‌ای را دارا می‌باشد. یکی از اصلی‌ترین نیازمندیهای آن خطوط ارتباطی با پهنای باند مناسب می‌باشد که با توجه به انتقال ویدیو یکی از اصلی‌ترین موانع پیش رو می‌باشد [1]. حجم تصاویر ویدیویی بسیار بالا بوده و متأسفانه به دلایل مختلف پهنای باند کافی برای ارسال این حجم در بسیاری از مناطق جغرافیایی فراهم نمی‌باشد. از این روی برای تکمیل و بهبود رسانی به این سیستم آموزشی افزایش پهنای باند یکی از راه حل‌ها محسوب می‌گردد. اما افزایش پهنای باند با در نظر گرفتن قیمت و مشکلات دیگر در بعضی مناطق جغرافیایی منطقی به نظر نمی‌رسد. در اینجا ایده‌ی فشرده‌سازی فایل‌های ویدیویی مطرح می‌شود. بجای

---

<sup>۱</sup> On-line

<sup>۲</sup> Off-line

آنکه اطلاعات را در مسیرهایی با پهنای باند بالا انتقال داده شود ، پس از فشرده‌سازی اطلاعات و کم کردن حجم آنان می‌توان آنان را در مسیرهایی با پهنای باند معمولی نیز با همان سرعت انتقال داد. در اینجا است که فشرده‌سازی نقش خود را در آموزش از راه دور به نمایش می‌گذارد. در حال حاضر در صنعت و پروژه‌های علمی رقابت بر سر متدهای جدید فشرده‌سازی می‌باشد که با در نظر گرفتن کیفیتی مطلوب بتوانند فایل های ویدیو را به فایلهایی با حجم کمتر تبدیل نمایند. برای این عمل باید چند مسئله را در نظر داشت:

✓ زمان فشرده‌سازی و بازکردن فایل‌های فشرده (برای انتقال بصورت برخط)

✓ میزان کیفیت کاهش یافته ی ویدیو پس از فشرده‌سازی

✓ میزان نرخ فشرده‌سازی تصاویر ویدیویی.

در آموزش از راه دور، آموزش به دو صورت انجام می پذیرد:

➤ نوع اول بصورت برخط<sup>۱</sup> می باشد که استاد و دانش پژوهان در یک زمان ولی در مکان های متفاوتی حضور دارند.

➤ نوع دوم بصورت برون خط<sup>۲</sup> است که ویدیوی برنامه ی کلاس درس از قبل ذخیره شده است و در زمان های آتی در اختیار دانش پژوهان قرار می گیرد.

ویدیو کنفرانسها و وبینارها و تماس‌های ویدیویی در دسته‌ی اول قرار می‌گیرند. در دسته اول از آموزش نیاز است با توجه به برخط بودن آموزش عمل فشرده‌سازی و بازکردن فایل‌های فشرده در زمان بسیار کوتاهی صورت پذیرد تا در روند آموزش تاخیری ایجاد نگردد.

---

<sup>۱</sup> On-line

<sup>۲</sup> Off-line

## ۱-۲ آشنایی با مفهوم ویدیو

عبارت ویدیو به تعدادی تصویر که بصورت متوالی به نمایش درمی‌آیند گفته می‌شود. هرکدام از این تصاویر را فریم<sup>۱</sup> می‌نامند. تمامی فریم‌ها در یک فایل ویدیویی دارای سایز مشخص و ثابتی می‌باشند. هر فریم دارای تعدادی پیکسل<sup>۲</sup> می‌باشد. پیکسل ریزترین و پایه‌ای ترین جزء یک ویدیو می‌باشد. هر پیکسل مقداری را در خود ذخیره کرده است که بیانگر رنگ تصویر در آن نقطه می‌باشد. هر چه چگالی پیکسل‌ها بیشتر باشد و یا به عبارت دیگر در یک پنجره با طول و عرض ثابت تعداد پیکسل‌های موجود بیشتر باشد می‌توان گفت که تصویر از وضوح بالاتری برخوردار می‌باشد. به تعداد پیکسل‌های یک فریم که بصورت تعداد پیکسل‌های عمودی و تعداد پیکسل‌های افقی بیان می‌شوند وضوح تصویر<sup>۳</sup> گفته می‌شود.

جدول ۱-۱: درجه وضوح استانداردهای متفاوت تصاویر ویدیویی

Acronym	Aspect ratio	Width (px)	Height (px)
VGA	4:3	640	480
SVGA	4:3	800	600
WSVGA	~17:10	1024	600
XGA	4:3	1024	768
XGA+	4:3	1152	864
WXGA	16:9	1280	720
WXGA	5:3	1280	768
WXGA	16:10	1280	800
SXGA- (UVGA)	4:3	1280	960
SXGA	5:4	1280	1024
HD	~16:9	1360	768
HD	~16:9	1366	768

فایل‌های ویدیویی بسته به کاربردشان دارای درجه وضوح‌های متفاوتی می‌باشند و این درجه وضوح تصاویر برای یکسان سازی در انتقال و استفاده در دستگاه‌های متفاوت از استانداردهای مشخصی پیروی

<sup>۱</sup>Frame

<sup>۲</sup>Pixel

<sup>۳</sup>Resolution

می‌کنند. در جدول شماره ۱-۱ تعدادی از استانداردهای شناخته شده به همراه درجه وضوح استفاده شده نشان داده شده است.

### ۱-۳ ضرورت فشرده‌سازی

تصاویر ویدیویی در حالت اولیه دارای حجم بسیار بالایی می‌باشند. هر تصویر ویدیویی از تعدادی فریم تشکیل شده است که بنا به کیفیت ویدیو برای هر ثانیه تعداد مشخصی فریم در نظر گرفته می‌شود. هر فریم دارای تعداد ثابت و مشخصی پیکسل می‌باشد. هر تصویر بنا به نوع کیفیت ویدیو می‌تواند تعداد ثابتی بیت برای ذخیره سازی آن پیکسل در نظر بگیرد. تصاویر ویدیویی به علت مقدار زیاد اطلاعات دارای حجم بسیار زیادی می‌باشند.

در جدول شماره ۱-۲ حجم چند نوع از تصاویر مرسوم ویدیویی در حالت غیر فشرده به همراه وضوح و نوع کیفیت آنان و میزان پهنای باند مورد نیاز جهت انتقال نمایش داده شده است [2].

جدول ۱-۲: نرخ فریم و نرخ ارسال استانداردهای تصویر

FORMAT	RESOLUTION	RAW BIT-RATE	REMARK
QCIF	176 ppl × 144 lfp × 25 fps	10.1 Mbps	digital video
CIF	352 ppl × 288 lfp × 25 fps	40.5 Mbps	digital video
SIF	360 ppl × 288 lfp × 25 fps	41.4 Mbps	digital video
SDTV	720 ppl × 576 lfp × 25 fps	165.8 Mbps	digital TV
HDTV	1920 ppl × 1080 lfp × 25 fps	829.4 Mbps	digital TV

انتخاب نوع استاندارد تصویر به کاربرد تصویر ویدیویی و همچنین بسترهای انتقال این تصاویر وابسته می‌باشد. می‌توان با انتخاب مناسب نوع استاندارد مناسب برای یک ویدیو از بهترین کیفیت و همچنین بهینه ترین استفاده از پهنای باند بهره برد.

همان گونه که مشاهده می‌شود انتقال این حجم داده حتی در پرسرعت ترین شبکه‌ها نیز امکان پذیر نمی‌باشد. پس نیازمند آن هستیم که با فشرده سازی تصاویر ویدیویی و کاهش حجم آنان، تصاویر فشرده سازی شده را منتقل نماییم. در غیر اینصورت امکان ارسال تصاویر ویدیویی از طریق شبکه‌های اینترنتی امکان پذیر نمی‌باشد.

پس می‌توان نتیجه‌گیری نمود که گسترش انتقالات و ارتباطات ویدیویی بر بستر شبکه‌های اینترنتی مرهون متدهای فشرده‌سازی می‌باشد. و همچنین می‌توان با بهبود متدها و روشهای فشرده سازی به گسترش بهتر و با کیفیت مطلوبتر تصاویر ویدیویی در شبکه اینترنتی کمک شایانی نمود. در حال حاضر رقابت اصلی بر روی کیفیت بالاتر ویدیوها بر روی شبکه‌های با پهنای باند کمتر می‌باشد. ارائه متدهای نوین فشرده سازی در حوزه ی ویدیو باعث شده است که تصاویر ویدیویی به تمامی حوزه‌ها راه یابد و به تمامی ابزارها قابلیت‌های جدیدی را ببخشد.

حال با توجه به حضور گسترده تصاویر ویدیویی و نیاز به جابجایی آن و مزایای گسترش ویدیو در فواصل بسیار زیاد، دیده می‌شود که متدهای فشرده سازی باعث می‌شوند در زمان، هزینه، حجم و ... در انتقال ویدیوها بهبود حاصل شده و با انجام عملیات فشرده سازی حجم یک فایل ویدیویی را تا میزان زیادی کاهش دهند.

در حال حاضر روش‌های فشرده سازی زیادی وجود دارند که معمولا اکثر آنها توسط سازمانهای استانداردسازی و یا تجاری در حال گسترش می‌باشند. رقابت این موسسات در میزان فشرده سازی فایل های ویدیویی و کیفیت تصاویر بعد از عملیات فشرده سازی می‌باشد. نرم‌افزارها و دستگاه‌های پخش



تصاویر ویدیویی به سرعت از فرمت های جدید ویدیویی پشتیبانی کرده و فایل های ویدیویی به سمت فرمت های با حجم کمتر و کیفیت بالاتری در حال حرکت می باشند.

## ۴-۱ تعریف فشرده سازی

فشرده سازی چیست؟ فشرده سازی به مجموعه عملیات هایی گفته می شود که با تغییر نحوه ی چینش، تغییر مقادیر پیکسل ها، پیش بینی مقادیر پیکسلها، استفاده از حوزه ی فرکانسی در توالی فریم ها، جابجایی مقادیر بیت ها برای رنگ ها، خلاصه سازی در ذخیره مقادیر بیت ها و غیره سعی در کاهش حجم یک فایل ویدیویی دارد.

فشرده سازی بسته به نوع تکنیک اجرایی بر روی ویدیو و محتوای بافت پیکسل ها و نحوه تغییرات فریمهای دارای نرخ های متفاوت فشرده سازی می باشد.

در فشرده سازی روش هایی وجود دارند که فارغ از نحوه ی بافت تصاویر به فشرده سازی ویدیو می پردازند. این روشها در بعضی ویدیوها نرخ بهتری در مقایسه با ویدیوهای دیگر دارند. حال اگر ما بخواهیم به نرخ بالاتری برای فشرده سازی دست پیدا کنیم بهتر است اطلاعاتی راجع به ویدیو از قبل داشته باشیم. در این صورت می توانیم با شناخت از نحوه ی چینش پیکسلها در یک فریم و رفتار یک پیکسل در فریمهای هم جوارش از روش مناسب تری برای فشرده سازی استفاده نماییم تا به نرخ بالاتری در کاهش حجم تصاویر ویدیویی دست پیدا نماییم.

## ۱-۵ تحرک در تصاویر ویدیویی

از یک دیدگاه برای دستیابی به هدف فوق می‌توان تصاویر ویدیویی را به دو دسته‌ی تصاویر پرتحرک<sup>۱</sup> و تصاویر کم‌تحرک<sup>۲</sup> طبقه‌بندی کرد. تصاویر پرتحرک مانند تصاویر مسابقات ورزشی، تصاویر با جابجایی زیاد دوربین و یا تغییرات زیاد در صحنه‌ها را شامل می‌شود که در آن یک پیکسل مشخص در فریم‌های مجاور دارای مقادیر متفاوتی می‌باشد. تصاویر کم‌تحرک مانند ویدیو کنفرانس‌ها، ویدیو چت‌ها، تصاویر ارسالی در آموزش از راه دور و غیره را در بر می‌گیرد. در تصاویر کم‌تحرک تغییر یک پیکسل در فریم‌های متوالی کمتر بوده است. در صورتیکه تصاویر کم‌تحرک را با روش‌های رایج در فشرده‌سازی ویدیو، فشرده‌سازی نماییم نرخ فشرده‌سازی قابل توجه نمی‌باشد. در صورتیکه با روش‌های فشرده‌سازی تعریف شده برای این دسته از تصاویر ویدیویی، فشرده‌سازی نماییم به نرخ فشرده‌سازی بالاتری می‌توان دست پیدا کرد. این میزان نرخ بالای فشرده‌سازی تصاویر ویدیویی می‌تواند کمک شایانی در انتقال این تصاویر می‌نماید. حال اگر تصاویر پرتحرک را با روش‌های فشرده‌سازی مربوط به تصاویر ویدیویی کم‌تحرک فشرده‌سازی شود دیده می‌شود که جدا از نرخ فشرده‌سازی که کاهش می‌یابد کیفیت تصاویر ویدیویی بعد از فشرده‌سازی نیز کاهش پیدا می‌نماید.

## ۱-۶ کیفیت در فشرده‌سازی

در فشرده‌سازی تصاویر ویدیویی دو دیدگاه اصلی وجود دارد. دیدگاه اول سعی در عدم تغییر مقادیر پیکسل‌های ویدیوی اولیه را دارد. این روش در انتهای عمل فشرده‌سازی هیچ تغییری در مقادیر پیکسل‌ها به‌وجود نمی‌آورد. معمولاً این روش برای فشرده‌سازی داده‌هایی که عدم تغییر در آنها پس از فشرده‌سازی

---

<sup>۱</sup>Fast Motion

<sup>۲</sup>Low Motion

بسیار مهم می باشند استفاده می شود. روش دوم روشی است که در آن ما بنا به موقعیت فریم و پیکسل‌ها تغییرات جزئی را در مقادیر پیکسل‌ها اعمال می‌نماییم تا بتوان به نرخ بالاتری در فشردن سازی دست پیدا کرد. این روش که با از دست دادن و تغییر مقادیر اولیه پیکسل‌های تصاویر ویدیویی همراه است بیشتر در حوزه ی فشردن سازی ویدیو به کار می‌رود و از نرخ بالاتری نسبت به روش قبل برخوردار می‌باشد. در روش‌هایی که با از دست دادن مقادیر اولیه داده در طی فشردن سازی همراه می‌باشند از یک معیار ارزیابی کیفیت استفاده می‌شود تا این میزان تغییر به محتوای تصاویر ویدیویی آسیب نرساند. معمولا این میزان تغییر تا حدی می‌باشد که تشخیص آن برای چشم انسان با توجه به مکانیزم بینایی دشوار باشد.

### ۱-۶-۱ اتلاف در فشردن سازی

فشردن سازی یک تصویر (یک تصویر دیجیتال و یا یک فایل ویدیویی) با فشردن سازی یک فایل داده‌ای تفاوت بسیار زیادی دارد. می توان این دو نوع فشردن سازی را به دو دسته‌ی فشردن سازی با اتلاف<sup>۱</sup> و فشردن سازی بدون اتلاف<sup>۲</sup> طبقه بندی نمود [3]. فشردن سازی یک فایل داده‌ای در دسته‌ی فشردن سازی بدون اتلاف قرار می‌گیرد و فشردن سازی تصاویر ویدیویی در دسته‌ی فشردن سازی با اتلاف قرار می‌گیرد. در فشردن سازی بدون اتلاف هیچ تغییری در مقادیر داده‌ها داده نمی‌شود بلکه با استفاده از حذف افزونگی<sup>۳</sup> در داده‌ها، باز چینش نحوه شماره گذاری داده‌ها (الگوریتم هافمن و غیره) و ... سعی در کاهش حجم داده دارد. این روش نرخ بسیار بالایی برای فشردن سازی تصاویر ویدیویی را دارا نمی‌باشد. در روش فشردن-سازی همراه با اتلاف، با تغییر مقادیر بعضی پیکسل‌ها، حذف بعضی از مقادیر پیکسل‌ها و جایگزینی آنان پس از خروج فایل از حالت فشردن با مقادیر پیش بینی شده از پیکسل‌های مجاور در یک فریم و یا در

---

<sup>۱</sup>Lossy compression

<sup>۲</sup>Lossless compression

<sup>۳</sup>Redundancy

توالی چندین فریم، سعی در کاهش حجم فایل ویدیویی دارد. در این روش بسته به میزان تخریب تصویر میزان فشرده‌سازی افزایش پیدا می‌کند. در روش فشرده سازی همراه با اتلاف سعی می‌شود میزان تخریب تا حدی باشد که چشم انسان متوجه تغییر و کاهش کیفیت ویدیو نشود. باید توجه داشت در بعضی از انواع ویدیوها تخریب جزئیات از اهمیت بالایی برخوردار نمی باشد. برای نمونه در یک تماس تلفنی تصویری می‌توان از یک سری از جزئیات تصویر چشم پوشی نمود و در ازای پهنای باند صرفه جویی شده صدا را با کیفیت بهتر و تصویر را بدون قطعی انتقال داد. برای بدست آوردن میزان تخریب که یک روش فشرده‌سازی بر روی یک ویدیو اعمال می‌کند، معیارهای اندازه گیری متفاوتی وجود دارد که در قسمت آتی توضیح داده خواهد شد.

می‌توان بسادگی بیان داشت که تصاویری که با روشهای بدون اتلاف فشرده می‌شوند را می‌توان دقیقاً به حالت اول بازگردانید اما در تصاویری که با روش‌های با اتلاف فشرده می‌شوند در صورت بازگرداندن تصاویر، دارای مقداری تفاوت (معمولاً کاهش کیفیت) نسبت به حالت اولیه می‌باشند. هر چه میزان فشرده سازی بیشتر باشد تغییر نسبت به ویدیو اولیه بیشتر خواهد بود. در شکل شماره ۱-۱ یک تصویر با سه نرخ متفاوت فشرده سازی با روش فشرده سازی به همراه اتلاف نمایش داده شده است.



شکل ۱-۲: تصویر a) فشرده سازی با ۸۴ درصد کمتر از اطلاعات تصویر اصلی، تصویر b) فشرده سازی با ۹۲ درصد کمتر از اطلاعات تصویر اصلی، تصویر c) فشرده سازی با ۹۸ درصد کمتر از اطلاعات تصویر اصلی

## ۱-۶-۲ معیارهای ارزیابی کیفیت

برای ارزیابی کیفیت یک ویدیو معیارهای زیادی وجود دارد. این معیارها را می‌توان به دو دسته تقسیم نمود. دسته‌ی اول معیارهای همراه با ناظر<sup>۱</sup> می‌باشد. در این نوع معیارها از تعدادی انسان درخواست می‌شود که نظرشان را نسبت به کیفیت یک ویدیو بیان نمایند. میانگین نظرات به عنوان معیار ارزیابی کیفیت ویدیو در نظر گرفته می‌شود. معیار دیگر را می‌توان معیار ارزیابی کیفیت ویدیو بدون ناظر<sup>۲</sup> دانست. در این معیار انسان هیچگونه دخالتی در ارزیابی کیفیت ویدیو ندارد، بلکه در این روش از ویدیوی اولیه استفاده شده و به مقایسه فریم به فریم با ویدیوی فشرده شده پرداخته می‌شود و میزان تغییرات این دو ویدیو فریم به فریم محاسبه می‌گردد. می‌توان روشهای دیگر بدون ناظر را که به محاسبه‌ی بافت تصویر پرداخته و میزان مات‌شدگی<sup>۳</sup> تصویر و یا زبرشدگی<sup>۴</sup> تصویر را بدست می‌آورند بیان نمود.

باید توجه داشت در بیشتر آزمایشات و محاسبات از روش بدون ناظر مقایسه‌ی فریم به فریم ویدیوی اصلی با ویدیوی فشرده شده استفاده می‌گردد. در ذیل چند نوع محاسبه برای روش ارزیابی بدون ناظر بیان شده است.

### ۱-۶-۲-۱ معیار میانگین مربع خطا MSE

معیار ارزیابی کیفیت ویدیوی میانگین مربع خطا<sup>۵</sup> یک معیار بدون ناظر می‌باشد. در این معیار هر فریم تصویر ویدیوی فشرده‌سازی شده با فریم متناظرش در ویدیوی اصلی مقایسه می‌گردد. در هر فریم برای تمامی پیکسل‌ها تفاوت مقادیر دو تصویر بدست می‌آید.

---

<sup>۱</sup>Supervised

<sup>۲</sup>Unsupervised

<sup>۳</sup>Blurring

<sup>۴</sup>Sharpness

<sup>۵</sup>Mean square error

$$\text{MSE} = \frac{1}{M_1 N_1} \sum_{(m,n) \in W} (b[m,n,k] - b[m,n,k-1])^2 \quad (1-1)$$

این تفاوت در صورتیکه صفر باشد بدین معنا است که هیچ تغییری در تصویر در طی فشرده‌سازی صورت نگرفته است. در صورتیکه مقدار این رابطه غیر از صفر باشد بدین معنا است که تصویر دچار تغییر (تخریب) شده است. هر چه این مقدار از صفر دورتر باشد یعنی میزان تغییر (تخریب) بیشتری صورت گرفته است. برای محاسبه‌ی این معیار ارزیابی کافی است تغییرات فریم را هر کدام به توان ۲ رسانیده و پس از مجموع تمامی مربعات تغییرات میانگین آنان را بدست آوریم [3].

هرچه عدد میانگین به صفر نزدیکتر باشد بیانگر آن است که تصویر پس از فشرده‌سازی کیفیت خود را کمتر از دست داده است. در صورتیکه این معیار برای یک تصویر ویدیویی برابر صفر باشد بدین معنا است که روش فشرده‌سازی اعمال شده، یک روش بدون اتلاف بوده است.

### ۱-۶-۲-۲ معیار میانگین قدر مطلق خطا MAE

معیار ارزیابی قدر مطلق خطا<sup>۱</sup> نیز یک معیار بدون ناظر می‌باشد. این معیار بسیار مشابه با معیار میانگین مربعات خطا می‌باشد. تنها تفاوت این دو معیار در این است که در معیار میانگین قدر مطلق خطا زمانیکه تفاوت مقادیر دو پیکسل متناظر در دو فریم متناظر را بدست می‌آورد به جای آنکه این مقدار تفاوت را به توان ۲ برساند از آن قدر مطلق گرفته می‌شود.

$$\text{MAE} = \frac{1}{M_1 N_1} \sum_{(m,n) \in W} |b[m,n,k] - b[m,n,k-1]| \quad (2-1)$$

---

<sup>۱</sup> Mean absolute error

می‌توان بیان داشت معیار میانگین مربعات خطا یک معیار درجه ۲ می‌باشد و میزان تفاوت پیکسل‌ها را به توان ۲ می‌رساند و تاثیر تفاوت‌های بزرگتر را بسیار بیشتر نشان می‌دهد. معیار میانگین قدر مطلق خطا یک معیار خطی بوده و میزان تفاوت مقادیر را به همان صورت نشان می‌دهد [3].

می‌توان مهم‌ترین مزیت معیار میانگین قدر مطلق خطا را در حجم کمتر محاسبات دانست. به علت آنکه در یک تصویر ویدیویی فریم‌های زیادی موجود است و هر فریم از تعداد بسیار زیادی پیکسل تشکیل شده است و برای بدست آوردن معیار مناسبی برای ارزیابی کیفیت باید تک تک پیکسل‌ها مورد محاسبه قرار گیرند.

حجم محاسبات بسیار بالا خواهد بود و معیار میانگین قدر مطلق خطا بعلا محاسبات کمتر نسبت به معیار میانگین مربعات خطا از پیچیدگی زمانی کمتری برخوردار بوده و مناسب‌تر به نظر می‌رسد.

## ۱-۷ ساختار بینایی انسان

مبنای دید انسان از محیط پرامون با تابش نور به اجسام و بازتابش آن به چشم انسان بوجود می‌آید. نور- مرئی قسمتی از طیف انرژی الکترو مغناطیسی می‌باشد. محدوده‌ی این طیف بیانگر نوع رنگ می‌باشد و چشم ناظر بسته به مکان نور در طیف انرژی رنگ خاصی را مشاهده می‌نماید.

مولفه‌هایی که با استفاده از آن می‌توان این انرژی الکترومغناطیسی را تعبیر کرد در ذیل بر مبنای مجمع جهانی استانداردسازی روشنایی<sup>۱</sup> CIE بیان شده است.

۱- روشنایی<sup>۱</sup> میزان نور دریافتی توسط چشم انسان که باعث تشخیص درخشنده تر بودن یک شی از شی دیگر می‌شود.

---

<sup>۱</sup><http://www.hike.te.chiba-u.ac.jp/ikeda/CIE/home.html>

۲- ته رنگ<sup>۲</sup> احساس انسان از شباهت بین دو ناحیه، توسط این پارامتر تعیین می‌گردد. در اصل ته- رنگ بیانگر رنگ غالبی است که چشم بیننده دریافت می‌نماید.

۳- پرننگی<sup>۳</sup> میزان کمتر یا بیشتر به چشم آمدن دو ناحیه با ته رنگ یکسان را نشان می‌دهد.

۴- درخشندگی<sup>۴</sup> میزان احساس درخشندگی یک ناحیه توسط چشم انسان نسبت به رنگ سفید مرجع.

۵- رنگینی<sup>۵</sup> میزان پرننگی یک ناحیه نسبت به درخشندگی رنگ سفید مرجع.

۶- اشباع<sup>۶</sup> میزان پرننگی یک ناحیه نسبت به درخشندگی آن ناحیه یا به عبارتی دیگر خلوص نسبی ته رنگ آن.

## ۸-۱ مدل‌های رنگی

### ۱-۸-۱ مدل رنگی RGB

یکی از اولین استانداردهای مورد استفاده در حوزه تصاویر دیجیتالی را که توسط CIE ارائه شده است را می‌توان استاندارد RGB<sup>۷</sup> بیان نمود. این استاندارد بر پایه ی سه رنگ اصلی قرمز، سبز و آبی پایه نهاده شده است. بر اساس این استاندارد هر رنگ را می‌توان با استفاده از این سه رنگ اصلی تشکیل داد. مطابق این استاندارد برای این سه رنگ اصلی تعریف ذیل را ارائه می‌شود.

- رنگ قرمز با طول طیف رنگ ۷۰۰ نانومتر

---

<sup>۲</sup> brightness

<sup>۳</sup> Hue

<sup>۴</sup> Colorfulness

<sup>۵</sup> Lightness

<sup>۶</sup> Chroma

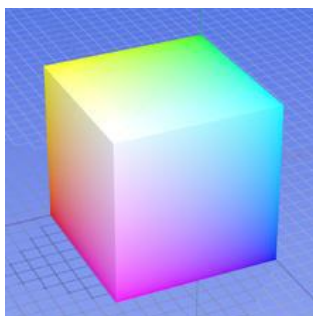
<sup>۷</sup> Saturation

<sup>۸</sup> Red Green Blue



- رنگ سبز با طول طیف رنگ ۵۴۶,۱ نانومتر
- رنگ آبی با طول طیف رنگ ۴۳۵,۸ نانومتر

در شکل شماره ۳-۱ فضای رنگ RGB در مکعب رنگها به نمایش گذاشته شده است.



شکل ۳-۱: پراکندگی رنگها در استاندارد RGB

در تصویر فوق بردار X که در راستای چپ افق امتداد داده شده است بیانگر رنگ قرمز و بردار Y که در راستای راست افق امتداد داده شده است رنگ آبی و بردار Z ها که عمود بر افق می باشد رنگ سبز را نشان می دهد. مرکز مختصات این تصویر که هر سه بردار دارای مقدار صفر می باشند را رنگ سیاه و دورترین نقطه تا مرکز مختصات که راس مکعب می باشد را رنگ سفید تشکیل داده است. مابقی رنگها از ترکیب مقادیر مختلف این سه بردار در داخل مکعب رنگی تشکیل می شود و بسته به میزان نمونه برداری از بردارها تعداد مقادیر مختلف تشکیل شده بیشتر و در نتیجه مکعب توانایی تشکیل رنگ های بیشتری را دارا خواهد بود. این مدل رنگ برای ایجاد تصویر در تلویزیون و مانیتورها (CRT, LCD, Plasma) بیشتر بکارگرفته می شود. در این استاندارد به سه رنگ قرمز، سبز و آبی رنگهای ابتدایی<sup>۱</sup> و به رنگهای دیگر که از این سه رنگ مشتق می گردند رنگهای ثانویه<sup>۲</sup> می گویند.

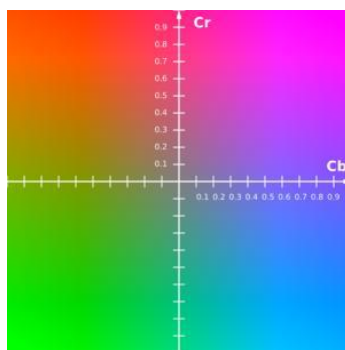
<sup>۱</sup>primary

<sup>۲</sup>secondary

تعریف فوق یک تعریف فیزیکی بر حسب طول طیف رنگ‌ها می‌باشد. در ادامه استاندارد دیگری ارائه می‌گردد که با ساختار بینایی انسان شباهت بیشتری را دارا می‌باشند.

### ۱-۸-۲ مدل رنگی Ycbcr

مدل RGB بهترین راه برای ارائه یک تصویر رنگی دیجیتال نمی‌باشد. در مدل RGB هر سه مولفه قرمز، سبز و آبی به میزان یکسانی از اهمیت برخوردار بوده و باید با یک دقت یکسان هر سه مولفه را ذخیره نمود. در بعضی از کاربردها ما نیازمند تغییر اندازه و مقادیر هر مولفه می‌شویم و برای این منظور از استاندارد Ycbcr استفاده می‌نماییم. در این استاندارد  $Y$  مولفه ی درخشندگی<sup>۱</sup> و از دو مولفه ی متفاوت رنگ ( $cb$ ,  $cr$  و یا  $U$ ,  $V$ ) استفاده می‌شود. مدل پراکندگی رنگها بر اساس دو محور  $cr$ ,  $cb$  در شکل شماره ۱-۲ نشان داده شده است.



شکل ۱-۴: مدل پراکندگی رنگها در Ycbcr

مدل رنگی RGB به راحتی می‌تواند به مدل رنگی Ycbcr یا YUV و بالعکس نگاشت گردد. در استاندارد ITU-R BT 601 [۴] تبدیلات این مدل به مدل RGB بیان شده است. این تبدیلات بر پایه‌ی روابط شماره ۱-۳ صورت می‌پذیرد.

---

<sup>۱</sup>luminance

$$\begin{bmatrix} Y_{601} \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.391 & -0.813 \\ 1.164 & 2.018 & 0 \end{bmatrix} \begin{bmatrix} Y_{601} \\ Cb \\ Cr \end{bmatrix} \quad (3-1)$$

همچنین سازمان ITU-R BT برای دستگاههای دیجیتالی با کیفیت بالاتر استاندارد ۷۰۹ را برای تبدیلات دو مدل رنگی RGB و Ycbcr بیان می‌دارد که در رابطه‌ی ۴-۱ نشان داده شده است.

$$\begin{bmatrix} Y_{709} \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.183 & 0.614 & 0.062 \\ -0.101 & -0.338 & 0.439 \\ 0.439 & -0.399 & -0.040 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.793 \\ 1.164 & -0.213 & -0.534 \\ 1.164 & 2.115 & 0 \end{bmatrix} \begin{bmatrix} Y_{709} \\ Cb \\ Cr \end{bmatrix} \quad (4-1)$$

### ۱-۲-۸-۱ فرمت های متفاوت نمونه برداری Ycbcr

عبارت اختصاری N:M:K بیانگر نرخ نمونه برداری متفاوت در سه مولفه ی Ycbcr می باشد. هر کدام از مولفه ها در حالت عادی ۸ بیت را اشغال می‌نمایند (در ویدیو های حرفه‌ای ۱۰ بیت). برای نمایش یک ویدیو با فرمت Ycbcr در ابتدا نیاز است تا نوع فرمت Ycbcr را شناسایی شود تا مقادیر جا افتاده از cb و cr را درون یابی شود. در ذیل انواع فرمت‌های استاندارد Ycbcr توضیح داده شده است.

➤ **نوع فرمت ۴:۴:۴** : نحوه قرار گیری مقادیر برای هر نمونه در Ycbcr برای نوع فرمت ۴:۴:۴ در

تصویر شماره ۵-۱ نشان داده شده است. در این نوع فرمت هر پیکسل از سه مقدار Y,cb,cr

تشکیل شده است که هر نمونه ۲۴ بیت (۳۰ بیت در ویدیوهای حرفه‌ای) فضا اشغال می‌نماید.



شکل ۵-۱: نوع فرمت ۴:۴:۴

➤ **نوع فرمت ۴:۲:۲:** در این نوع فرمت برای هر دو مقدار Y فقط یک مقدار cr, cb موجود است. در این نوع فرمت برای هر پیکسل ۱۶ بیت (و برای ویدیوهای حرفه‌ای ۲۰ بیت) فضا مورد نیاز می‌باشد. این نوع فرمت در شکل شماره ۶-۱ توضیح داده شده است.



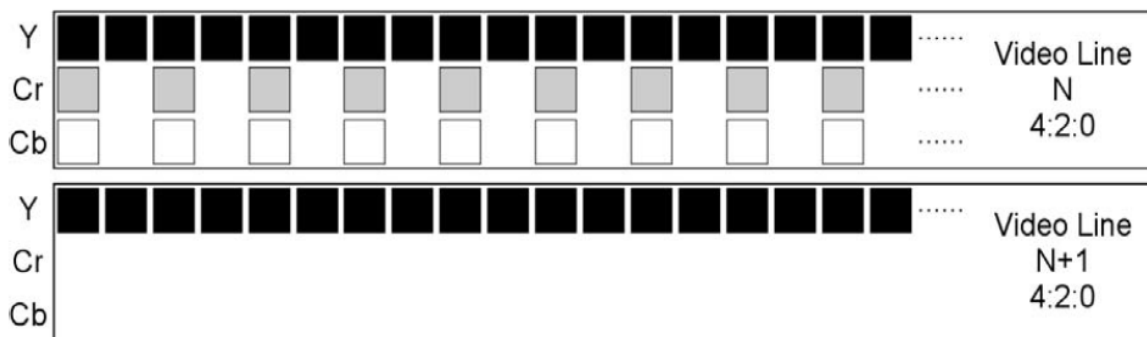
شکل ۶-۱: نوع فرمت ۴:۲:۲

➤ **نوع فرمت ۴:۱:۱:** در این نوع فرمت که با نام YUV12 نیز شناخته می‌شود و بیشتر برای کاربردهای فشرده سازی مورد استفاده قرار می‌گیرد برای هر ۴ مقدار Y فقط یک مقدار از cr, cb وجود دارد. هر پیکسل ۱۲ بیت فضا اشغال می‌نماید ( این نوع فرمت در ویدیوهای حرفه‌ای مورد استفاده قرار نمی‌گیرد). این نوع فرمت در شکل شماره ۷-۱ توضیح داده شده است.



شکل ۷-۱: نوع فرمت ۴:۱:۱

➤ **نوع فرمت ۴:۲:۰:** این نوع فرمت در فشرده سازی تصاویر ویدیویی بسیار استفاده می‌گردد. این نوع فرمت به غیر از حذف برخی از مقادیر cr, cb که در راستای افقی صورت می‌گیرد برای cr, cb در راستای عمودی نیز حذف را صورت می‌دهد. با این عمل نرخ فشرده سازی را ۲ به ۱ می‌نماید. این فرمت در تصویر شماره ۸-۱ توضیح داده شده است.



شکل ۸-۱: نوع فرمت ۴:۲:۰

## ۹-۱ ساختار پایان نامه

در این پایان نامه سعی شده است مطالب بصورت زنجیروار به یکدیگر پیوند داده شود. با توجه به حوزه تصاویر ویدیویی استفاده شده در حوزه‌ی آموزش از راه دور، که در دسته تصاویر ویدیویی کم‌تحرک قرار می‌گیرند در این پایان نامه تاکید بر روی روشهای این حوزه از تصاویر ویدیویی صورت گرفته است. در فصل اول، مقدمه به توضیح مختصری در مورد ضرورت فشرده سازی و تعریف اولیه ای از آن می‌پردازد. در این فصل همچنین با کیفیت فشرده سازی و معیارهای ارزیابی کیفیت ویدیو آشنا می‌شویم. برای شروع هر قدمی در مبحث فشرده‌سازی باید با سیستم بینایی انسان شناختی داشته باشیم. انواع مدل‌های رنگی و استانداردهای موجود در این زمینه نیز در این فصل بیان شده است.

در فصل دوم روش‌ها و استانداردهای فشرده سازی تصاویر ویدیویی توضیح داده شده است. در انتهای هر قسمت به مقالات و یا تولیدات هر سازمان و یا گروهی که از روش مذکور استفاده کرده اند اشاره می‌شود. در این فصل روش‌هایی که در فشرده‌سازی ویدیو باید از آنان استفاده نمود گفته شده است. تمامی استانداردها تا به امروز که در حوزه‌ی تصاویر ویدیویی استفاده شده در آموزش از راه دور مفید می‌باشند بیان شده است.

در فصل سوم روش پیشنهادی برای افزایش سرعت در انتقال تصاویر ویدیویی آموزش از راه دور تشریح خواهد شد. در ادامه پایگاه داده معرفی می شود و مراحل اجرای روش توضیح داده خواهد شد ، هر روش مورد استفاده، بررسی و آنالیز می شود. سپس در فصل چهارم نتایج و آزمایشات را گزارش می شود که به مقایسه ی نتایج بدست آمده با متدهای پیشین و همچنین آنالیز دلایل هر نتیجه پرداخته خواهد شد. در فصل پنجم به ارزیابی فعالیت های انجام شده، کارهای صورت پذیرفته و بیان نحوه ی ادامه ی مسیر در فعالیت های آتی پرداخته خواهد شد.

## فصل دوم:

مروری بر فعالیت‌های انجام شده (روش‌ها و استانداردهای فشرده‌سازی)

## ۲-۱ مقدمه

فشرده‌سازی ویدیو از سالهای بسیار گذشته در جهت کاهش حجم ویدیو و افزایش کیفیت ویدیو با ثابت نگهداشتن حجم بسیار مورد توجه قرار گرفته است. در دهه‌های اخیر با توجه به گسترش ارتباطات و قابلیت انتقال ویدیو در تماس‌های تلفنی، ویدیوکنفرانس‌ها، کلاسهای آموزش مجازی، تلویزیون‌های اینترنتی و کابلی، برنامه‌های چند رسانه‌ای تحت وب و ... نگرش خاصی به این حوزه شده است. به علت نیاز جامعه به افزایش سرعت در انتقال ویدیو در بستر شبکه‌های اینترنتی، این حوزه مورد توجه خاص صنعت قرار گرفته‌است. سازمان‌ها و کمپانی‌های مطرح در زمینه نرم افزارهای کامپیوتری و یا سخت افزارهای الکترونیکی به این حوزه نفوذ کرده و با ارائه برنامه‌ها و دستگاه‌های جدید خود به رقابت با یکدیگر در این عرصه می‌پردازند. مهمترین مزیت حضور کمپانی‌های تجاری به این حوزه، رشد و پیشرفت بسیار بالا در این حیطه می‌باشد. اما مهمترین ایراد آن را می‌توان در مخفی نگهداشتن روش‌ها و متدهای مخصوص به هر سازمان دانست. در صورتیکه تحقیقات آکادمیک برای استفاده‌ی همگان در دسترس عموم قرار می‌گیرد، تحقیقات تجاری بصورت مخفیانه نگهداشته می‌شود و می‌تواند عامل موفقیت یک سازمان در رقابت برای بدست آوردن جامعه متقاضی باشد. یکی از مهمترین دلایل کمبود مقالات معتبر در این حوزه در مقایسه با حوزه‌های دیگر را می‌توان در این مسئله دانست. در طول دهه‌ی اخیر با افزایش کاربرد فشرده‌سازی ویدیو بر بستر شبکه‌های اینترنتی تعداد مقالات این حوزه کاهش چشمگیری یافته است.

در این حوزه سازمان‌های استانداردسازی برای یکپارچگی متدها و روش‌های انتقال ویدیو سعی در ارائه‌ی استانداردهای فشرده‌سازی ویدیو می‌باشد. هر کدام از این استانداردها برای نوع خاصی از تصویر و با قابلیت مشخصی برای فشرده‌سازی تصویر استفاده می‌شوند. اکثر این استانداردها توسط سازمان‌ها و کمپانی‌های معتبر ارائه شده و در سازمان‌های استاندارد سازی به ثبت رسیده است. در ذیل در ابتدا با



روش های متنوعی که در فشرده سازی استفاده می شود آشنا شده و سپس به مروری بر استانداردهای موجود پرداخته می شود.

## ۲-۲ روش های فشرده سازی

در هر ویدیو همبستگی<sup>۱</sup> زیادی مابین مقادیر پیکسل های موجود در یک فریم و همچنین مقادیر یک پیکسل در توالی فریم های گوناگون وجود دارد. این همبستگی ها به ما در شناخت میزان افزونگی<sup>۲</sup> ها کمک کرده تا با حذف این افزونگی های درون فریمی<sup>۳</sup> و بیرون فریمی<sup>۴</sup> حجم یک فایل ویدیویی را کاهش دهیم. هرگاه افزونگی های یک فریم که یک تصویر است حذف گردد این روش را فشرده سازی درون فریمی و هرگاه افزونگی های توالی چند فریم برای یک پیکسل و یا یک ناحیه از تصویر حذف گردد، آن را فشرده سازی برون فریمی می نامند.

در فشرده سازی از سه تکنیک ذیل برای کاهش حجم یک فایل ویدیویی استفاده می شود.

- کاهش افزونگی مکانی : حذف افزونگی ها در یک تصویر با توجه به همبستگی بالای پیکسل های یک تصویر را فشرده سازی بر پایه ی کاهش افزونگی مکانی گویند. این روش یک روش درون فریمی محسوب می گردد.
- کاهش افزونگی زمانی : حذف افزونگی ها در توالی چندین فریم با توجه به همبستگی بالای یک پیکسل و یا یک ناحیه از تصویر در فریم های گوناگون را فشرده سازی بر پایه ی کاهش افزونگی زمانی گویند. این روش یک روش برون فریمی محسوب می گردد.

---

<sup>۱</sup>Correlation

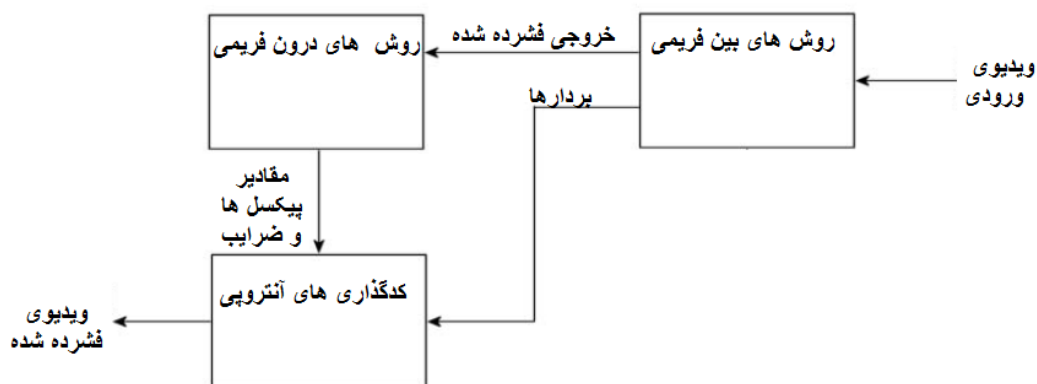
<sup>۲</sup>Redundancy

<sup>۳</sup>Intra-frame

<sup>۴</sup>Extra-frame

- کدگذاری آنتروپی: گروهی از روش های بدون اتلاف فشرده سازی برای کاهش حجم ویدیو بر مبنای تغییر طول داده ها و یا تغییراتی در راستای بهبود نرخ فشرده سازی در روشهای دیگر فشرده سازی را فشرده سازی بر پایه ی کدگذاری آنتروپی<sup>۱</sup> می نامند.

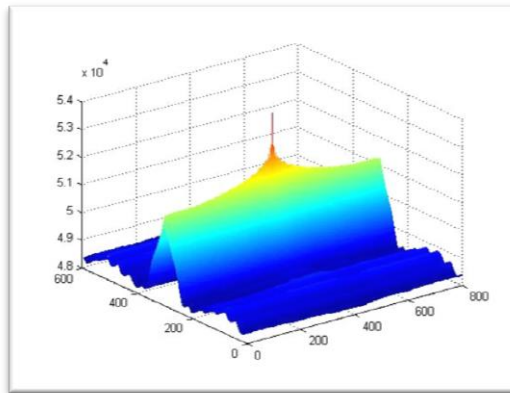
در تصویر شماره ۱-۲ نمای کلی از فشرده سازی تصویر با استفاده از سه روش بالا نمایش داده شده است.



شکل ۱-۲: نمای کلی فشرده سازی یک فایل ویدیویی

## ۱-۲-۲ افزودنی مکانی

در یک فریم ویدیو، پیکسل ها دارای همبستگی بسیار زیادی با یکدیگر می باشند. در شکل شماره ۲-۲ همبستگی دوبعدی پیکسل های یک فریم از تصویر و فریم بعدی تصویر نمایش داده شده است. تصویر شماره ۲-۲ نشانگر آن است که ما بین پیکسل های این فریم و فریم بعدی ارتباط و همبستگی بسیار بالایی برقرار است. حال برای کاهش حجم یک تصویر کافی است این افزودنی ها را از تصویر حذف شود.



شکل ۲-۲: میزان همبستگی دو فریم مجاور

تکنیک‌های ذیل به عنوان مهم ترین تکنیک های برای حذف افزونگی درون فریمی بکار می‌رود.

➤ **تکنیک پیش بینی کننده :** در این تکنیک مقدار بعضی از پیکسل ها را با استفاده از مقادیر

پیکسل های مجاور آن بدست می آوریم. این روش با توجه به محدودی بودن مقدار پیش بینی یک روش فشرده سازی همراه با اتلاف محسوب می گردد. با توجه به میزان نرخ فشرده سازی، تعداد پیکسل های پیش بینی شونده می توانند افزایش پیدا کنند. این روش با نام مدل

کدگذاری پالس تفاضلی<sup>۱</sup> شناخته می شود [5]

در تصویر شماره ۲-۳ همان گونه که نشان داده شده است برای بدست آوردن مقادیر پیکسل  $X$  از مقادیر پیکسل های مجاور استفاده شده است. تاثیر این روش وابسته به دقت تابع پیش بینی کننده می‌باشد. دقت تابع پیش بینی کننده به میزان همبستگی مقادیر پیکسل‌ها و همچنین تعداد پیکسل‌های معلوم به پیکسل های مجهول وابسته می باشد. هر چه تفاوت میان مقادیر پیکسل‌های اولیه با پیکسل های پیش بینی شده کمتر باشد کیفیت ویدیو بیشتر حفظ شده و تصاویر دچار تخریب کمتری می شود. میزان نرخ فشرده سازی و کیفیت تابع پیش بینی کننده با یکدیگر رابطه‌ی معکوس داشته و بسته به کاربرد و میزان

<sup>۱</sup>Differential Pulse Code Modulation

کیفیت مورد انتظار درصد تعداد پیکسل‌های پیش بینی شونده تغییر می‌نماید.

.	.	.	.	.	.	.
.	A	B	C	.	.	.
.	D	X	E	.	.	.
.	G	H	I	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

شکل ۲-۳: نحوه پیش‌بینی مقادیر پیکسلها

➤ **تکنیک بر پایه تبدیلات:** در این روش یک فریم از ویدیو به حوزه ی دیگری تبدیل پیدا می کند که در آن همبستگی کمتر و فشرده‌سازی راحت‌تر صورت می‌پذیرد. در ذیل دو دسته از این تبدیلات توضیح داده شده است.

- پایه بلوک : یک تصویر کامل و یا تصویری بعد از اعمال چند مرحله فشرده سازی به بلوک‌های مربعی تقسیم می‌گردد. این بلوک‌ها به علت حجم کمتر دارای محاسبات پردازشی کمتری بوده و براحتی می توان بر روی آنان محاسبات را انجام داد و از روشهای جبران حرکت<sup>۱</sup> نیز استفاده نمود.

ایراد اصلی این روش ایجاد ناهمگونی<sup>۲</sup> هایی در مرزهای این بلوک‌ها می باشد. زیرا هر بلوک با استفاده از داده‌های مقادیر داخلی اش فشرده شده و مرزها مطابق با این مقادیر کدگذاری شده اند و از مقادیر بلوک همجوار استفاده‌ای نشده است. این مسئله برای بلوک هم جوار نیز صادق می باشد. پس می‌توان در نظر داشت که دو مرز کنار هم دارای مقادیر مختلفی نسبت به یکدیگر خواهند بود. این تفاوت

<sup>۱</sup> Motion compensation

<sup>۲</sup> artifact

مقدار در پیکسل های هم جوار در مرزها باعث ناهمگونی شده و کیفیت تصویر را دچار افت می کند. می توان از <sup>1</sup>SVD و تبدیل کسینوسی گسسته<sup>2</sup> به عنوان مثالهایی از این روش نام برد.

- بر پایه تصویر: این روش بر روی کل تصویر و یا یک برشی<sup>3</sup> از تصویر عملیات فشرده سازی را اعمال می نماید. تبدیلات تصویر مانند تبدیل موجک گسسته کیفیت بهتری را در مقایسه با تبدیل کسینوسی گسسته را در فشرده سازی دارا می باشد. باید توجه داشت که تبدیل موجک گسسته نیاز به حافظه ی بیشتری می باشد و با روش های جبران حرکت به خوبی عمل نمی کند.

➤ **تکنیک کوانتیزاسیون<sup>4</sup>:** کوانتیزاسیون یک روشی است که در آن یک سیگنال با دامنه ی  $\alpha$  را به یک سیگنال با دامنه ی کمتر از  $\alpha$  نگاشت می کنند. در این حالت کوانتیزاسیون قابلیت دارد یک سیگنال را با تعداد بیت کمتری نسبت به حالت اولیه ذخیره سازی نماید. در کوانتیزاسیون یک مقدار به نزدیکترین مقدار کوانتیزه همگرا می گردد. این تکنیک یک تکنیک فشرده سازی به همراه اتلاف می باشد. بنابر این می توان نتیجه گیری نمود که داده هایی که با این روش فشرده می شوند را نمی توان بدون خطا دوباره به حالت اولیه بازگردانید. به مقدار تفاوت داده های کوانتیزه شده با حالت اولیه خطای کوانتیزاسیون می گویند.

➤ **تکنیک مرتب سازی دوباره:** به مجموعه تکنیکهایی که با دیدی متفاوت، مقادیر یک تصویر را پیمایش می کنند به قسمی که بتوان با چینش دوباره مقادیر به نرخ فشرده سازی بالاتری

---

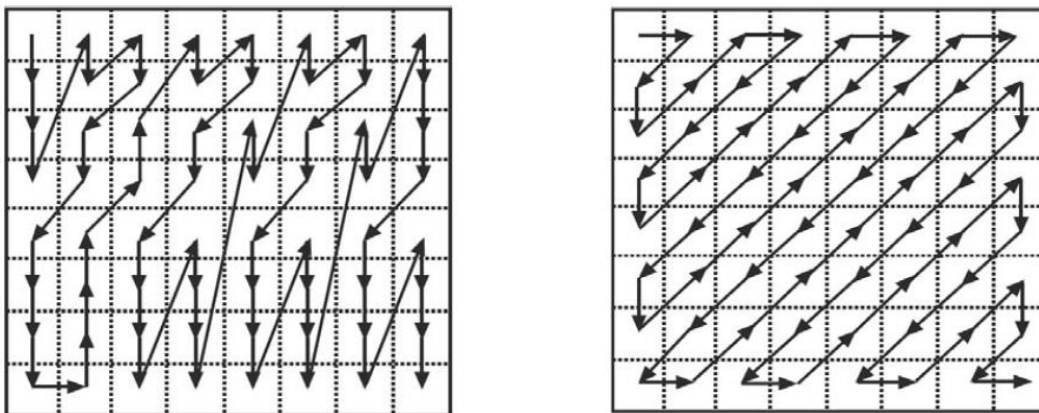
<sup>1</sup>Singular value decomposition

<sup>2</sup>Discrete cosine transform

<sup>3</sup>slice

<sup>4</sup>Quantization

دست پیدا کرد را تکنیکهای مرتب سازی دوباره می گویند. به عنوان مثال می توان بیان داشت که در DCT مهمترین ضرایب معمولا در نزدیکی گوشه ی سمت چپ بالای تصویر قرار دارند. بعد از وزن دهی، ضرایب با مقدار بسیار کم به صفر تغییر مقدار می دهند. در صورتیکه ضرایب غیر صفر در ابتدا ارسال شوند و سپس تعدادی صفر ارسال شود می توان به نرخ بهتری در انتقالات دست پیدا کرد. برای این عمل باید ضرایب به نحوی پیمایش گردند که احتمال این اتفاق بیشتر شود. در شکل شماره ۲-۴ دو تصویر از روش های گوناگون پیمایش ضرایب DCT نشان داده شده است که بیشترین احتمال دسترسی به حالت بهینه را دارا می باشد. در این حالات در ابتدا اکثر مقادیر غیر صفر و در انتها اکثر مقادیر صفر پیمایش می گردند.



شکل ۲-۴: نحوه ی پیمایش ضرایب DCT برای بهینه کردن فشردگی سازی

## ۲-۲-۲ افزودگی زمانی

برای فشردگی سازی یک ویدیو با استفاده از حذف افزودگی های ما بین فریمها روشهای بسیار زیادی وجود دارد. این روشها دارای پیچیدگیهای زمانی، مکانی متفاوت می باشند. این روشها بر پایه وجود همبستگی ما بین دو فریم متوالی طراحی شده اند. در این روشها با حذف این افزودگیها حجم ویدیو کاهش پیدا می کند. در تصویر شماره ۲-۵ نمونه ای از دو فریم متوالی و تصویر تفاضلی این دو فریم

نمایش داده شده است که می توان به خوبی میزان همبستگی دو فریم متوالی را در تصویر تفاضلی مشاهده نمود.



شکل ۲-۵: مقدار تفاضلی دو فریم متوالی

## ۲-۲-۱ کدگذاری تفاضلی

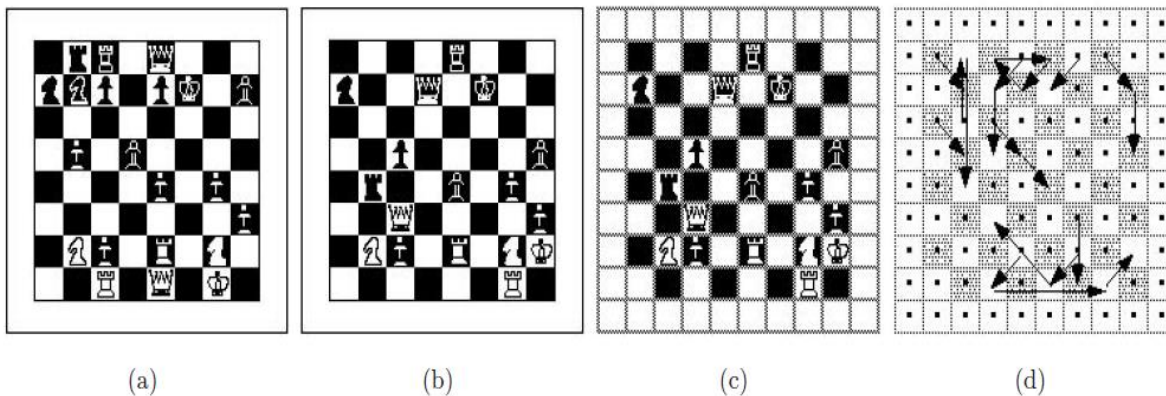
در روش کدگذاری تفاضلی، تفاضل مقدار دو فریم متوالی محاسبه و ذخیره می گردد. به عبارت دیگر مقادیر پیکسل ها در هر فریم با تفاوت مقادیر پیکسل ها با فریم قبلی جایگزین می گردد. در صورتیکه مقادیر پیکسل های دو فریم با یکدیگر شباهت زیادی داشته باشند آنگاه مقادیر فریم تفاضلی دارای تعداد بسیار زیادی مقدار صفر و یا نزدیک به صفر خواهد بود. میدانیم هر چه فراوانی یک مقدار در یک تصویر بیشتر باشد می توان آن تصویر را با نرخ بالاتری فشرده نمود. برای کد گشایی ویدیو هایی که با روش تفاضلی فشرده می شوند کافی است مقدار فریم تفاضلی را به مقدار پیکسل در فریم های قبلی اضافه نماییم تا مقادیر اصلی فریم بدست آید. با استفاده از این روش مقدار زیادی از افزونگی موجود در فریم های متوالی از بین رفته و حجم ویدیو کاهش می یابد.

در تصاویری که تغییرات در فریم های متوالی بسیار کم است می توان فقط پیکسل هایی را که تغییر کرده اند را کدگذاری نمود و برای ما بقی پیکسل ها که کد نمی شوند همان مقدار قبلی را در نظر گرفت. در این روش برای بهبود میزان فشرده سازی با صرف نظر از تغییرات بسیار جزئی پیکسل ها می توان فقط تغییرات بیشتر از یک حد آستانه ای مشخص را در نظر گرفت و مقادیر کمتر از حد آستانه را نادیده گرفته و

آن پیکسل‌ها را بدون تغییر در نظر داشت. با استفاده از این روش میزان فشرده‌سازی یک فایل ویدیویی به میزان زیادی افزایش پیدا می‌کند. باید توجه داشت هر چه این حدآستانه بزرگتر شود میزان مقادیر پیکسل‌هایی که تغییر می‌کنند بیشتر و با مقدار اولیه تفاوت بیشتری داشته و کیفیت ویدیو کاهش بیشتری خواهد داشت.

## ۲-۲-۲ تخمین حرکت و جبران آن بر پایه ی بلوک بندی

یک صفحه‌ی شطرنج را به عنوان یک فریم در نظر بگیرید. این صفحه به بلوک‌هایی بدون همپوشانی تقسیم شده است. مهره‌ها در این بلوک‌ها جابجا می‌شوند و فریم جدیدی را تشکیل می‌دهند. برای ذخیره سازی اتفاقات این صفحه شطرنج کافی است به جای ذخیره هر فریم تنها مسیرهای حرکت هر بلوک را ذخیره نماییم. این بدان معنا است که به جای ذخیره هر فریم کافی است بدانیم که محتویات بلوک به شماره ی  $\alpha_1$  به مکان بلوک  $\alpha_2$  انتقال پیدا کرده است. بدین صورت می‌توان به میزان بسیار بالایی ویدیو را فشرده نمود. در این روش بردار حرکت را به جای مقادیر پیکسل‌ها ذخیره می‌شود که بسیار کم حجم‌تر می‌باشد.



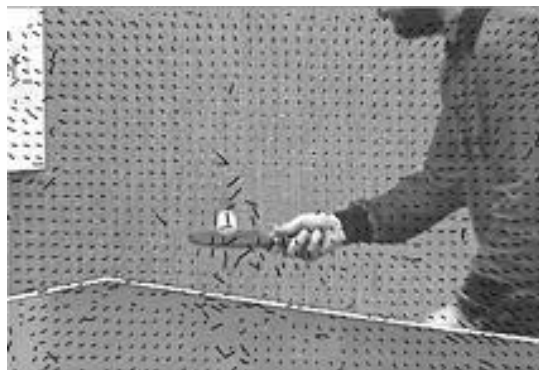
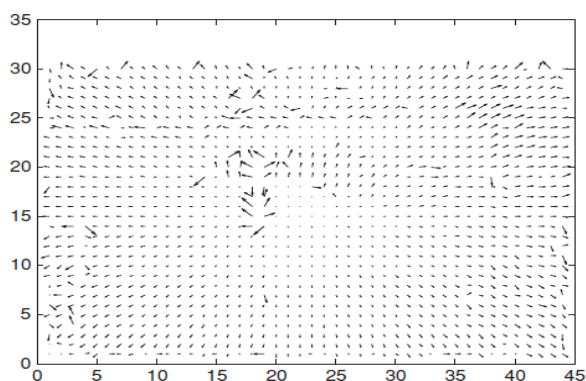
شکل ۲-۶: تصویر a, b دو صحنه ی متوالی شطرنج می‌باشند تصویر c نگاهت صفحه بر روی بلوک‌ها می‌باشد و تصویر d

بردار حرکت بلوک‌های شطرنج را نمایش میدهد



برای این منظور نیاز است که بتوان در فریم‌های متوالی به تخمین حرکت هر بلوک پرداخت و مشابه‌ترین بلوک را در فریم بعدی جستجو کرد. معیارهای متفاوتی را برای ارزیابی شباهت دو بلوک موجود می‌باشد [6] و روش‌های جستجوی دو بلوک در دو فریم نیز یکی از مهمترین زمینه‌های این نوع روش فشرده‌سازی می‌باشد زیرا این جستجو نیاز به محاسبات بالا و مقدار حافظه‌ی بالایی می‌باشد. هر چه الگوریتم جستجو بهینه‌تر باشد و از حجم محاسبات کمتری برخوردار باشد می‌تواند در موفقیت بیشتر این روش تاثیر بیشتری بگذارد [7].

برای فشرده‌سازی بدین روش از تکنیک تخمین حرکت<sup>۱</sup> استفاده می‌شود. در این روش حرکت هر بلوک توسط بردار حرکت آن مشخص می‌شود که شامل میزان جابجایی در راستای عمودی و میزان جابجایی در راستای افقی می‌باشد. در صورتیکه بلوک مطابق با بلوک مورد نظر همان بلوک اولیه باشد بردار حرکت که میزان جابجایی بلوک را در فریم آتی نمایش می‌دهد دارای مقدار صفر خواهد بود. این تغییرات در شکل شماره ۲-۷ نشان داده شده است.

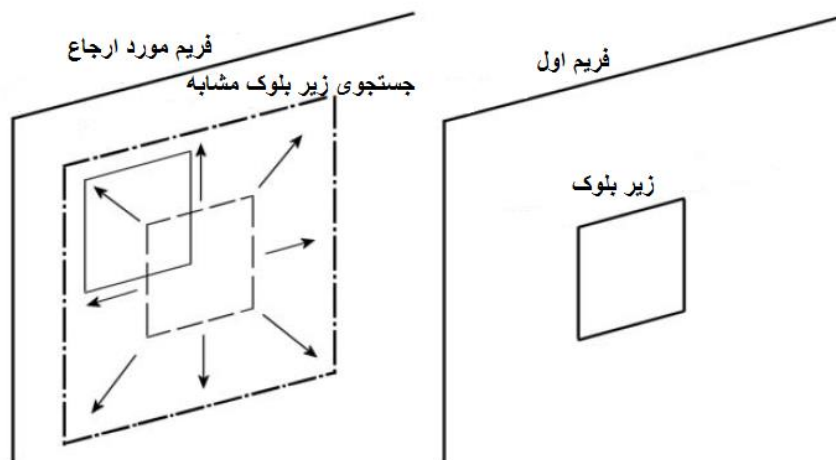


شکل ۲-۷: تصویر بردارهای حرکت در یک فریم که میزان جابجایی هر بلوک را نسبت به فریم بعدی نشان می‌دهد

یک روش بسیار معمول برای تطابق بلوک‌ها، محاسبه‌ی تفاضل مقادیر بلوک مورد نظر با بلوک کاندید

می‌باشد. در صورتیکه قدر مطلق تفاضلی SAD<sup>۱</sup> دو بلوک از مابقی بلوک‌ها کمتر بود می‌توان حدس زد که بلوک مشابه را پیدا کرده ایم. بلوک مشابه بلوکی است که بیشترین تطابق را با بلوک مورد نظر داشته باشد. باید در نظر داشت که این روش از لحاظ زمانی و مکانی بسیار پر هزینه خواهد بود زیرا در این روش کل بلوک‌های فریم جدید باید پیمایش شوند که بسیار زمانبر خواهد بود.

یکی از روش‌های بهینه استفاده از این روش، تکنیک محدود کردن فضای جستجو می‌باشد. در این روش جستجو بر اساس احتمال حضور بلوک در یک منطقه از فریم صورت می‌پذیرد. پس می‌توان نتیجه گرفت در این حالت کلیه بلوک‌های یک فریم جستجو نمی‌شوند بلکه بلوک‌هایی که احتمال تشابه بیشتری داشته باشند مورد جستجو قرار می‌گیرند. برای نیل به این هدف از متدهای پیش بینی رفتارهای بردار حرکت استفاده می‌شود. بدین گونه که رفتارهای بردار حرکت در فریم‌های گذشته محاسبه شده و برای فریم حاضر مکان جدیدی برای بلوک حدس زده می‌شود. در تصویر شماره ۲-۸ زیر فضای جستجوی یک بلوک در فریم بعدی آن به نمایش گذاشته شده است.



شکل ۲-۸: فضای جستجوی یک فریم برای پیدا کردن بلوک مشابه

<sup>۱</sup>Sum of Absolute Differences

روش بهبود یافته‌ی این تکنیک نیز با استفاده از تصحیح متد پیش بینی کننده صورت می‌پذیرد. پس از پیش بینی مکان جدید و مشاهده‌ی بردار حرکت، تابع پیش بینی کننده تصحیح شده و برای فریم‌های بعدی با توجه به تصحیح‌های صورت گرفته در فریم‌های گذشته پیش بینی رفتارهای بردار حرکت صورت می‌پذیرد. این روش که با استفاده از بازخورد محاسبات قبلی، محاسبات جدیدی را انجام می‌دهد را روش تخمین حرکت و جبران<sup>۱</sup> می‌نامند.

### ۳-۲-۲-۲ تخمین حرکت و جبران دو طرفه

برای تخمین بهتر و بهینه‌تر بردار حرکت روشهای بسیاری بیان شده است [8] یکی از روش‌هایی که باعث افزایش کارایی تابع پیش بینی کننده می‌شود را می‌توان تخمین حرکت و جبران دوطرفه<sup>۲</sup> معرفی نمود. در این روش برای بدست آوردن بردار حرکت یک بلوک در یک ویدیو از یک فریم پیشین و یک فریم آتی فریم مورد نظر استفاده می‌شود. سپس با دانستن مکان بلوک در فریم قبلی و بعدی، مکان فریم در بلوک فعلی را بدست می‌آورند.

یکی از مشکلاتی که در روش تخمین بردار حرکت موجود است، مساله‌ی مقادیر پیکسل‌هایی است که بعد از حرکت یک بلوک در پشت زمینه ظاهر می‌شوند. در روش‌های معمولی این مسئله بعلت عدم آگاهی از مقادیر آن پیکسل‌ها به مشکل دچار می‌شوند و ویدیو را دچار افت کیفیت می‌کنند. شی حرکت می‌کند اما جایگزین مکان قبلی شی مقداری وجود ندارد.

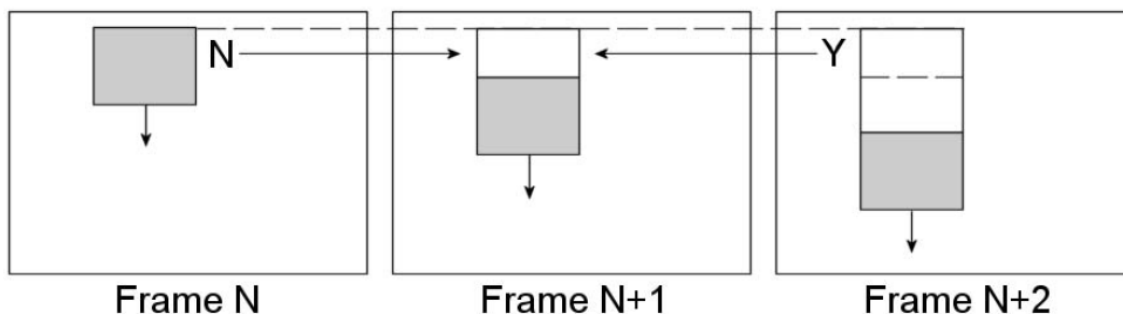
یکی دیگر از مزایای روش تخمین حرکت در دسترس بودن مقادیر بعدی یک بلوک پس از حرکت می‌باشد. زیرا در روش تخمین بردار حرکت دو طرفه مقادیر فریم قبلی و بعدی در دسترس است و مقادیری که قرار است جایگزین بلوک شوند در دسترس می‌باشند. در روش تخمین بردار حرکت دوطرفه این

---

<sup>۱</sup> Motion Estimation and Compensation

<sup>۲</sup> Bidirectional Motion Compensation

مشکل تا حد امکان کاسته شده و کیفیت ویدیو کمتر دچار افت می‌گردد. در زمانیکه پشت صحنه در وضعیت آشکار شدن می‌باشد مقادیر جدید جایگزین مقادیر بلوک سابق می‌گردند. در تصویر شماره ۹-۲ نحوه ی کار روش تخمین حرکت و جبران دو طرفه نشان داده شده است.



شکل ۹-۲: روش تخمین حرکت و جبران در ۳ فریم متوالی

## ۳-۲-۲ کدگذاری آنتروپی

کدگذاری آنتروپی<sup>۱</sup> یک روش برگشت پذیر فشرده سازی ویدیو می‌باشد. این بدین معنا می‌باشد که در طی این روش هیچ مقدار پیکسلی تغییر نمی‌یابد و هیچ داده‌ای از دست نرفته و کیفیت تصویر ویدیویی ثابت باقی می‌ماند. در این نوع کدگذاری، مقادیر پیکسل‌ها به نحوی پیمایش می‌شوند که مناسب‌ترین فضا برای ذخیره سازی داده‌ها ایجاد گردد [9,10] و بیشترین نرخ فشرده سازی برای ویدیو حاصل شود.

یکی از روش‌هایی که آن را می‌توان زیرمجموعه‌ای از کدگذاری آنتروپی دانست، روش دفترچه کد<sup>۲</sup> می‌باشد. در این روش از یک دفترچه کد برای ذخیره سازی مقادیر استفاده می‌شود. در نظر بگیرید یک الگو از مقادیر پیکسل‌ها در تصویر بیشتر تکرار گردد و مابقی الگوها با تعداد کمتری در تصاویر مشاهده شوند. در این صورت می‌توان تمامی الگوها را در یک دفترچه کد ذخیره سازی نمود.

<sup>۱</sup>Entropy coding

<sup>۲</sup>Codebook

در این روش برای الگوهایی که تکرار بیشتری دارند از کدی با طول کمتر استفاده شده و برای الگوهایی که دارای تکرار کمتری می باشند از کدی با طول بیشتر استفاده می شود. در این صورت می توان ادعا داشت که در فضای ذخیره سازی صرفه جویی مناسبی صورت گرفته است. و در انتها نیز می توان با مراجعه به دفترچه کد مقدار هر الگو را با توجه به مقدار ذخیره سازی شده بدست آورد و جایگزین نمود.

باید توجه داشت که دفترچه کد نیز به عنوان یک سربرار باید همراه با فایل ویدیویی منتقل گردد. اما از آنجائیکه میزان فشرده سازی با این روش برای تصاویر ویدیویی که الگوی تکراری در آنان از فراوانی بالایی برخوردار است آنقدر زیاد می باشد که سربرار دفترچه کد مقدار قابل ملاحظه ای نمی باشد.

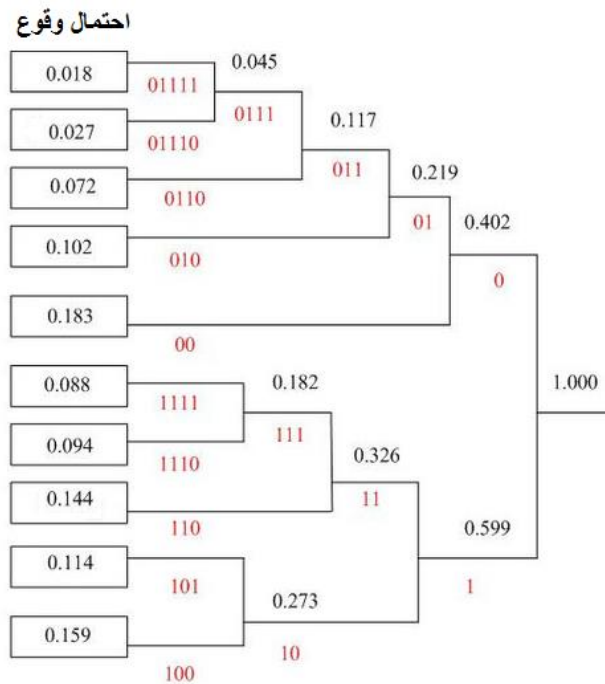
## ۲-۳-۲-۱ کدگذاری هافمن

تکنیک هافمن<sup>۱</sup> [11] یکی از روش های کلاسیک فشرده سازی برای ذخیره سازی مقادیر با طول رشته های متغییر می باشد. در این روش برای مقادیری که دارای بیشترین تکرار می باشند کدی با کمترین طول رشته در نظر گرفته می شود و برای مقادیری که کمترین تکرار را دارا می باشند از کدی با بیشترین طول رشته استفاده می شود.

در این روش کدها به نحوی انتخاب می شوند که در زمانیکه پشت سر یکدیگر قرار میگیرند قابل تمایز باشند. برای این عمل از درخت دودویی که در تصویر شماره ۱۰-۲ نمایش داده شده است، استفاده می - شود.

---

<sup>۲</sup>Huffman coding



شکل ۲-۱۰: نحوه ی کدگذاری با استفاده از درخت هافمن

کدها متمایز بوده و هنگامیکه کد های یک رشته با دفترچه کد هافمن تطبیق داده می شوند شناخته شده و مقدار اصلی جایگزین می شود و مقدار کد از رشته کدشده حذف می شود و کد بعدی در دفترچه جستجو می گردد. این عمل تا زمانیکه تمامی کدها از رشته کد شده حذف شوند ادامه پیدا می کند. همان گونه که توضیح داده شد بعلت متغییر بودن طول کدهای هافمن می توان کدهای با طول کمتر را برای مقادیری که دارای فراوانی بیشتری می باشند استفاده نمود و کدهای با طول بیشتر را برای متغییر-هایی که امکان مشاهده ی آنان بسیار کمتر است استفاده نمود. این کدگذاری برای داده هایی که فراوانی آنان نرمال می باشد، بهینه نمی باشد و بیشتر برای داده هایی استفاده می شود که انحراف معیار پراکندگی آنان بسیار کم می باشد. در این صورت می توان ادعا نمود این روش از نرخ فشرده سازی بالایی استفاده می کند. باید در نظر داشت که دفترچه کد هافمن نیز باید همراه با داده [1] [2]ها ذخیره گردد که حجم

این دفترچه در مقایسه با صرفه جویی که در حجم یک تصویر ویدیویی بدست آمده است قابل اغماض می باشد.

## ۲-۲-۳-۲ کدگذاری LZ

یکی دیگر از روش های کدگذاری آنتروپی استفاده از متد<sup>1</sup> LZ می باشد [12]. روش LZ بر مبنای تکرار یک الگو یا یک مقدار پشت سر یکدیگر می باشد. در این روش هنگامیکه چندین پیکسل دارای یک مقدار ثابت باشند می توان مقدار پیکسل ها و تعداد آنان را به جای ذخیره ی تک تک مقادیر آنان ذخیره نمود. در نظر بگیرید در یک تصویر ویدیویی مقدار یک پیکسل در تعداد مشخصی فریم هیچگونه تغییری نمی کند. می توان مقدار این پیکسل و تعداد فریم هایی را که این پیکسل بدون تغییر باقی خواهد ماند را ذخیره کرد. در این روش کافی است دو مقدار (مقدار پیکسل و تعداد تکرار آن) را ذخیره کرد و از ذخیره ی تعداد بالایی مقدار یکسان (افزونگی) جلوگیری نمود که باعث کاهش قابل ملاحظه ای در فشردن سازی خواهد شد. می توان ادعا نمود که این روش فشردن سازی یک روش بدون اتلاف و برگشت پذیر می باشد.

می توان در کدگذاری LZ برای افزایش کارایی تغییرات کوچک و ناچیز را در مقدار یک پیکسل در فریم های متوالی نادیده گرفت و تمامی مقادیری که از یک حد آستانه ای تجاوز نکنند را یک مقدار در نظر گرفت و سپس با روشی که در بالا توضیح داده شده است، فشردن سازی نمود. پر واضح است که این روش یک روش فشردن سازی با اتلاف بوده و مقادیر برگشت پذیر نمی باشند و ویدیو پس از فشردن سازی دچار افت کیفیت می شود. هر چه نرخ فشردن سازی را با بالا بردن حد آستانه بالا بریم افت کیفیت ویدیو بیشتر می شود. برای این منظور باید یک حد وسط برای افزایش نرخ فشردن سازی و کاهش کیفیت در نظر گرفت.

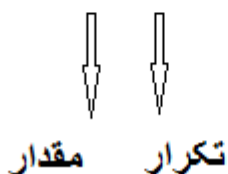
---

<sup>1</sup>Lempel-Ziv Coding

در روش LZ می توان تصویر را بگونه‌ای پیمایش نمود که پیکسل های با مقدار برابر و یا مشابه نیز پشت سر یکدیگر قرار گیرند. هدف آن است که نحوه‌ی پیمایش بگونه ای باشد که احتمال وقوع تکرار در پیکسل‌هایی که با مقادیر برابر و یا مشابه پشت سر یکدیگر افزایش پیدا کنند. هرچه این احتمال افزایش پیدا کند نرخ فشرده سازی ویدیو با این روش نیز افزایش پیدا می کند. بهینه ترین نوع از این روش تکنیک Lempel-Ziv-Welch Variation است که برای آشنایی بیشتر با آن می‌توان به [13]مراجعه نمود در تصویر شماره ۱۱-۲ روش کار روش فشرده سازی بر پایه LZ توضیح داده شده است.

**250-250-250-250-250-144-144-144-144-144-  
199-199-199-199-199-199-199-250-250-  
250-250-250-250-250-...**

**مقدار کدشده: 250(5)144(5)199(8)250(7)**



شکل ۱۱-۲: روش کار الگوریتم LZ

## ۳-۲ مدل های کدک ویدیو

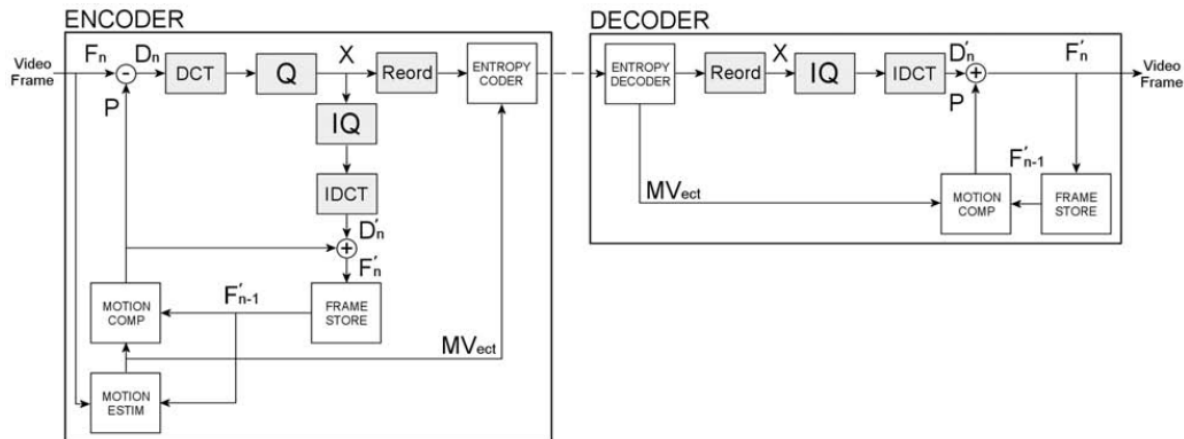
اکثر کدهای فشرده سازی استاندارد ویدیو دارای مدل ساختاری مشابهی می‌باشند. این مدل ساختاری شامل تخمین حرکت و جبران، تبدیلات و کدگذاری آنتروپی می باشند. یک کدک<sup>۱</sup> شامل یک کدگذار<sup>۲</sup> و

<sup>۱</sup>CODEC

<sup>۲</sup>Code



یک کدگشا<sup>۱</sup> می باشد که می تواند یک فایل ویدیویی را فشرده کرده انتقال دهند و پس از انتقال دوباره به حالت اول بازگردانند. در تصویر شماره ۱۲-۲ ذیل مدل کلی کدگذاری و کدگشایی ویدیو به نمایش گذاشته شده است.



شکل ۲-۱۲: نحوه ی کار یک کدگذار و کدگشا

## ۴-۲ سازمان ها و استانداردها

استانداردها ابزار مورد نیاز برای تعامل شرکت های گوناگونی را که بر روی سیستم های مشابهی کار می کنند فراهم می سازد. استانداردها باعث رشد و شکوفایی ابزارها و تکنولوژی ها می شوند. در دنیای تجاری، رقابت ها را استانداردها قانون مند و مسیر تکنولوژی را مشخص می کنند. سازمان ITU-T<sup>۲</sup> را می توان یکی از دو سازمان مطرح در زمینه ی استانداردسازی ویدیو در نظر گرفت. سازمان دیگری که در این زمینه فعالیت میکند سازمان ISO/IEC JTC1<sup>۳</sup> می باشد. سازمان ITU-T معمولا با استانداردهایی به فرم

<sup>۱</sup> Decode

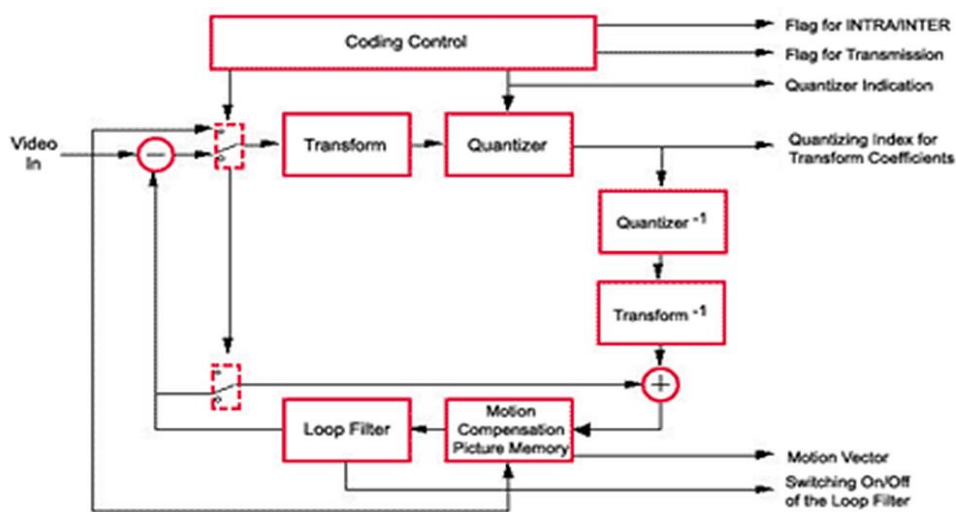
<sup>۲</sup> International Telecommunication Union Telecommunication

<sup>۳</sup> International Standardization Organization and the International Electrotechnical Commission, Joint Technical Committee number 1

H.26X شناخته شده است.(مانند H.26, H.262, H.263, H.26L و ...) در صورتیکه سازمان ISO/IEC را معمولاً با استانداردهایی به فرمت MPEG-X بیشتر شناخته می‌شوند.(مانند MPEG-1, MPEG-2, MPEG-4 و ...)

## ۲-۴-۱ سازمان ITU-T

سازمان ITU-T استانداردهایی را برای ارتباطات از راه دور، تلفن های تصویری و ویدیو کنفرانس ها ارائه می‌دهد. استاندارد H.261 که برای ارتباطات با پهنای باند  $64 \text{ kbps}$  طراحی شده بود بر پایه ترکیبی از روش فشرده سازی جبران حرکت بر پایه ی بلوک بندی و تبدیل کسینوسی گسسته می‌باشد. استاندارد H.261 از سال ۱۹۸۴ تا سال ۱۹۹۰ مورد استفاده قرار می گرفت. در شکل شماره ۱۳-۲ روش کاری این استاندارد برای انتقال ویدیویی تحت وب نشان داده شده است.

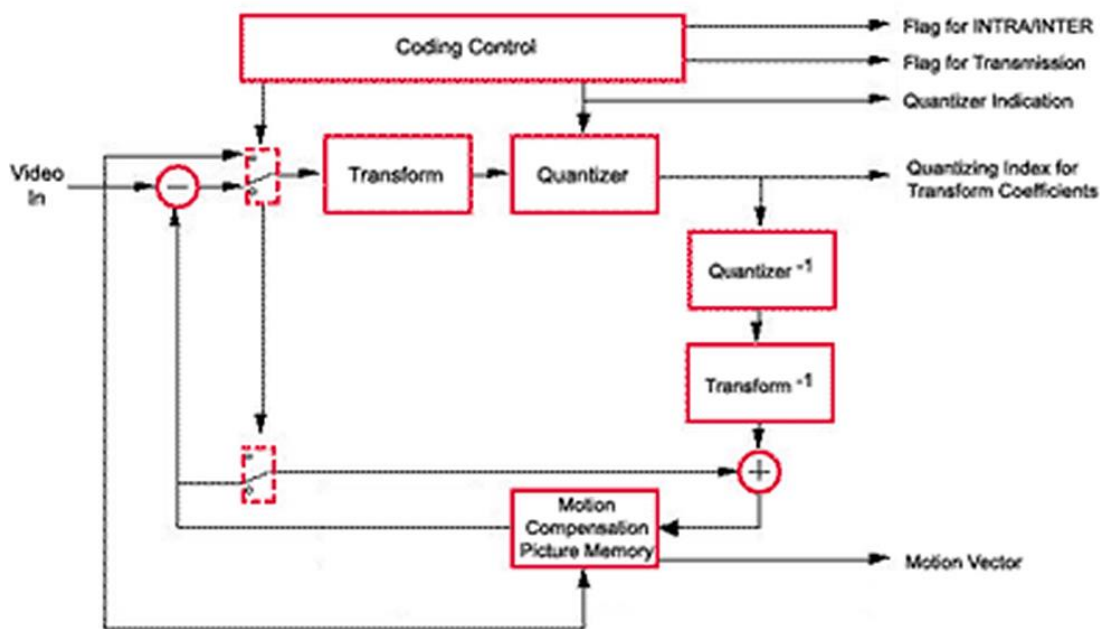


شکل ۲-۱۳: دیاگرام فرمت فشرده سازی H.261

نیاز به نرخ ارسال پائین (کمتر از 64 kbps) ویدیوهای ارتباطی، ITU-T را بر آن داشت تا استاندارد H.263 را ارائه دهد. این استاندارد که در سال ۱۹۹۳ معرفی شد قابلیت پشتیبانی از استاندارد H.261 را

دارا بود و همچنین با متدهای جدیدی که در آن بکار رفته بود (تغییر در نحوه‌ی کدگذاری ضرایب تبدیل و بردارهای حرکت) توانست تصاویر بهتری را با نرخ ارسال کمتر انتقال دهد. یکی از مهمترین مزایای H.263 نسبت به استاندارد H.261 را می‌توان پیچیدگی زمانی کمتر دانست.

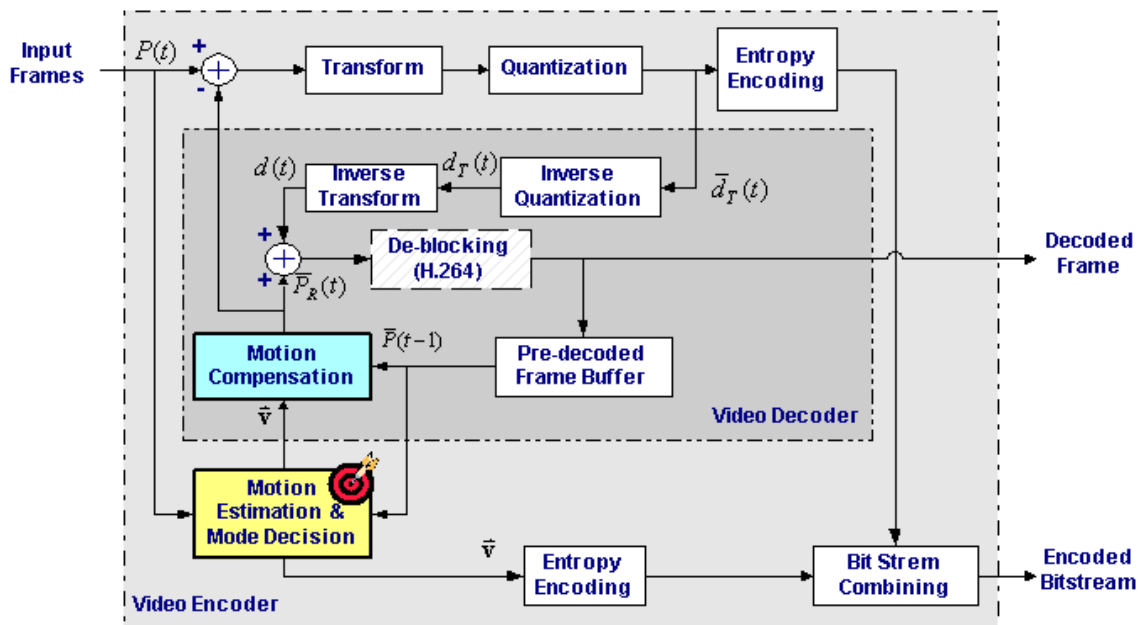
H.263 توانست در انتقال تصاویر ویدیویی یکی از بهترین استانداردها لقب بگیرد و دیگر استانداردهایی مانند 3GPP, H.323, H.310, H.320, H.324 بر پایه‌ی آن و به عنوان نسخه دوم H.263 معرفی شوند. استانداردهای نسخه دوم H.263 بیشتر بر بهبود تصاویر تمرکز داشتند. نسخه سوم H.263 در سال ۲۰۰۱ معرفی شد که برای کاربردهای نیاز به تاخیر کمتر در ارسال مناسب‌تر می‌بود. در تصویر شماره ۱۴-۲ نحوه‌ی کار استاندارد H.263 تحت وب توضیح داده شده است.



شکل ۱۴-۲: دیاگرام فرمت فشرده سازی H.263

## H.264 کدک ۱-۱-۴-۲

این کدک دارای نرخ فشرده سازی و کیفیت بالاتری در میان کدک‌های موجود می‌باشد. از لحاظ نحوه‌ی عملکرد مشابه با کدک H.263 می‌باشد. در این کدک پس از تبدیل تصویر به مایکروبلوک‌ها از جبران حرکت استفاده می‌شود که دارای بهبودهایی نسبت به کدک‌های قبلی می‌باشد. مهمترین تفاوت این کدک با کدک‌های قبلی را می‌توان در وجود فیلتر دیبلاکینگ دانست. [18] این کدک دارای انتخاب‌های زیادی برای میزان فشرده‌سازی توسط کاربر می‌باشد و می‌توان با استفاده از این، قدرت‌های انتخابی میزان فشرده‌سازی را جهت انتقال با پهنای باند کمتری اعمال نمود. از دیگر کاربردهای این کدک می‌توان در تلویزیون‌های کابلی و ماهواره‌ای دیجیتال و همچنین تصاویر ویدیویی کنفرانس از راه دور نام برد. [19] در شکل زیر نمایی از دیاگرام این کدک نمایش داده شده است.



شکل شماره ۲-۱۵ دیاگرام کدک H.263

## ۲-۴-۲ سازمان ISO MPEG

سازمان MPEG<sup>۱</sup> در سال ۱۹۸۸ به عنوان یک سازمان استانداردسازی در جهت ارائه استانداردهای لازم جهت فشرده سازی تصاویر ویدئو و سیگنال‌های صوتی شروع به کار نمود. هدف اصلی این سازمان ارائه الگوریتم‌های کدگذاری ویدئو برای ذخیره سازی بهینه تر تصاویر چندرسانه ای می باشد. این سازمان با ارائه استاندارد JPEG<sup>۲</sup> برای بهبود فشرده سازی تصاویر و عکسبرداری دیجیتالی فعالیت خود را آغاز نمود. تمامی استانداردهای ارائه شده توسط این سازمان قابلیت پشتیبانی از استانداردهای قبلی را نیز دارا می‌باشند.

### MPEG-1 ۱-۲-۴-۲

استاندارد MPEG-1 در سال ۱۹۹۴ ارائه گردید.[۱۴] این استاندارد نیز همانند استاندارد H.261 برای تصاویری با وضوح کم و همچنین نرخ ارسال پایین ارائه شده بود. این استاندارد اغلب در کاربردهایی مانند ویدئو کنفرانس‌ها استفاده می‌شده است. تصاویر استفاده شده در این فرمت از استاندارد SIF<sup>۳</sup> که شامل تصاویر ویدیویی با وضوح ۳۵۲\*۲۴۰ پیکسلی برای NTSC و تصاویر ویدیویی با وضوح ۳۵۲\*۲۸۸ پیکسلی برای PAL پیروی می‌کنند، می‌باشند.

نرخ ارسال تصاویر ویدیویی با فرمت MPEG-1 حدود ۱,۵ mbps می باشد. این استاندارد فشرده سازی تصاویر ویدیویی و صوتی را برای استفاده در CDها ارائه شد که برای تصاویر ویدیویی فرمت VCD<sup>۴</sup> و برای

---

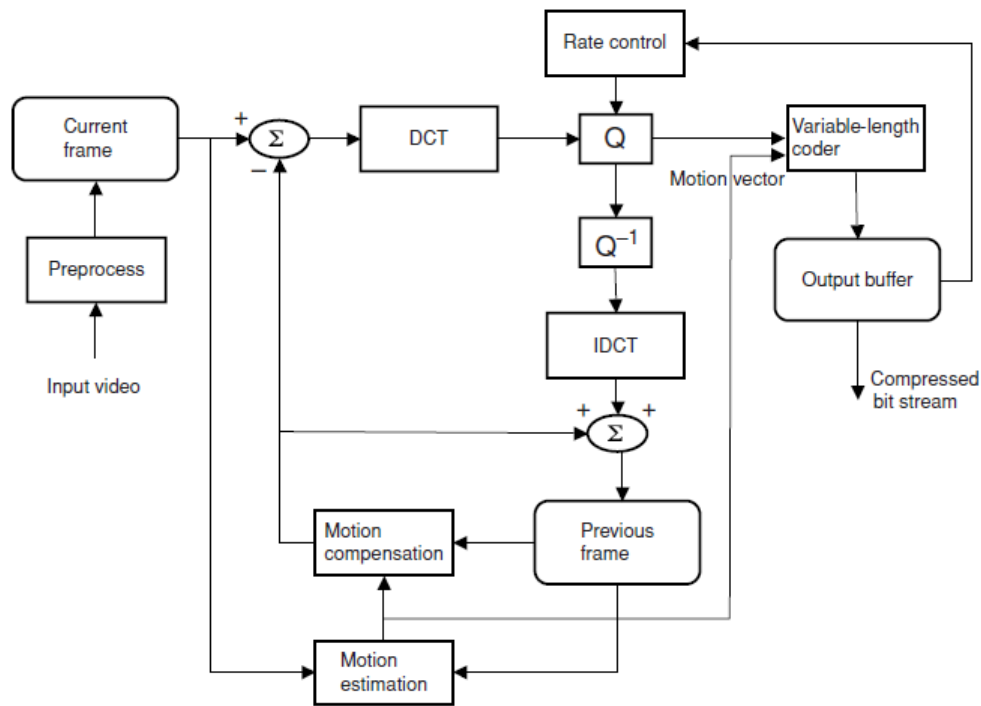
<sup>۱</sup>Moving Pictures Expert Group

<sup>۲</sup>Joint Photographic Experts Group

<sup>۳</sup>Standard Interchange Format

<sup>۴</sup>Video CD

فایل‌های صوتی فرمت MP3<sup>۱</sup> شناخته شده می باشد. در تصویر شماره ۱۶-۲ طرح ساده ای از نحوه‌ی کدگذاری در این فرمت نمایش داده شده است.



شکل ۱۶-۲: دیاگرام فرمت فشرده سازی MPEG-1

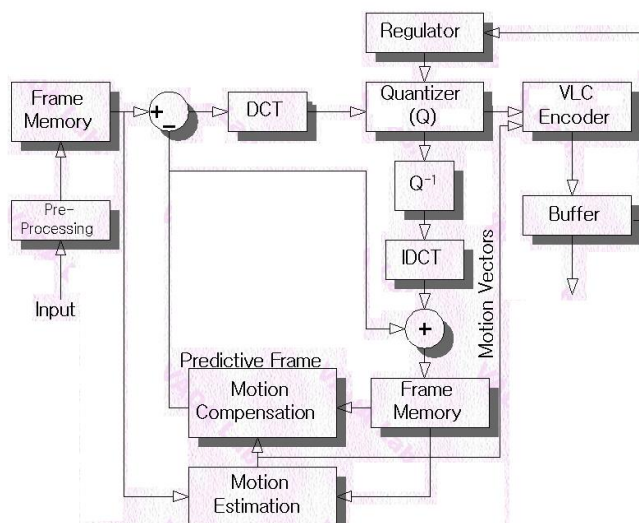
## MPEG-2 ۲-۲-۴-۲

در سال ۱۹۹۰ سازمان ISO MPEG کار بر روی استاندارد MPEG-2 را با هدف توسعه توانایی‌های استاندارد MPEG-1 آغاز نمود. [15] این استاندارد برای بهبود کیفیت تصاویر تلویزیونی در نرخ ارسال بالاتر از ۱,۵ Mbps ارائه شده است. می توان نتیجه گرفت که MPEG-1 و MPEG-2 از لحاظ کارایی با یکدیگر قابل مقایسه نمی باشند. زیرا MPEG-2 دارای پهنای باند اختصاص داده شده ی بیشتری نسبت به استاندارد MPEG-1 دارا می باشد و به همین علت از کیفیت بالاتری نسبت به استاندارد قبلی بهره می

<sup>۱</sup>MPEG-1 Layer-3

برد. یکی از ایرادات این استاندارد زمانی می باشد که با پهنای باند  $1.5 \text{ Mbps}$  تصاویر ویدیویی را انتقال می دهد. زیرا با این پهنای باند کیفیت بهتری را نسبت به استاندارد قبلی شامل نمی شود.

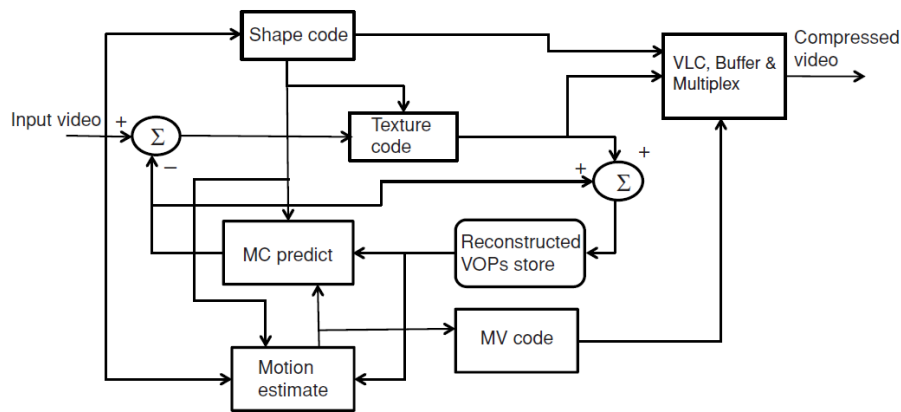
پهنای باند اختصاص داده شده برای این استاندارد  $3-10 \text{ Mbps}$  می باشد. این میزان پهنای باند برای ارسال تصاویری با وضوح  $720 \times 480$  پیکسلی NTSC و  $720 \times 576$  پیکسلی PAL در استاندارد MPEG-2 اختصاص داده شده است. استاندارد MPEG-2 قابلیت پشتیبانی از تصاویری با کیفیت HDTV را نیز به شرط اختصاص پهنای باند  $30-15 \text{ Mbps}$  را نیز دارا می باشد. این استاندارد برای تصاویر ویدیویی DVD بسیار مناسب بوده و به همین علت شبکه های تلویزیونی کابلی و همچنین شبکه های پخش ماهواره ای از این استاندارد برای پخش تصاویر ویدیویی خود استفاده کرده اند. استاندارد MPEG-3 را می توان مشابه استاندارد MPEG-2 دانست و تنها تفاوت این دو استاندارد پهنای باند  $20-40 \text{ Mbps}$  استاندارد MPEG-3 می باشد. ولی هیچگاه استاندارد MPEG-3 مورد استفاده قرار نگرفت. در شکل شماره ۲-۱۷ نحوه فشرده سازی توسط استاندارد MPEG-2 توضیح داده شده است.



شکل ۲-۱۷: دیاگرام فرمت فشرده سازی MPEG-2

## MPEG-4 ۳-۲-۴-۲

استاندارد MPEG-4 با ارائه یک الگوریتم پیشنهادی توانست به نرخ فشرده سازی بهتری نسبت به MPEG-2 دست پیدا کند. یکی از مزایای MPEG-4 نسبت به MPEG-2 در ارسال تصاویر در نرخ پایین بود. در هنگام ارسال تصاویر با نرخ پایین کیفیت تصاویر در استاندارد MPEG-2 همانگونه که در قسمت قبل توضیح داده شد دچار کاهش می گردید در صورتیکه در استاندارد MPEG-4 کیفیت تصاویر در نرخ پایین دچار کاهش چشمگیر نمی شود. نسخه اول این استاندارد در سال ۱۹۹۸ معرفی شد و نسخه دوم آن که دارای ویژگی‌های اضافه شده‌ای به نسخه ی اول می باشد در سال ۲۰۰۱ معرفی گردید. در استاندارد MPEG-4 جهت فشرده سازی از روشهای شناخت اشیاء، لایه بندی تصویر و فشرده سازی با درجات مختلف استفاده شده است. این استاندارد اولین گام در جهت فشرده سازی بر اساس محتوا می باشد. در تصویر شماره ۲-۱۸ روش کار این استاندارد توضیح داده شده است.



شکل ۲-۱۸: دیاگرام فرمت فشرده سازی MPEG-4



## MPEG-7 ۴-۲-۴-۲

استاندارد MPEG-7 با قصد جایگزینی استانداردهای قبلی ارائه نشد. این استاندارد برای تکمیل ویژگیهای دیگر استانداردهای MPEG ارائه شد. در این استاندارد اطلاعاتی راجع به محتوای تصویر ذخیره می‌گردد. با استفاده از این استاندارد می‌توان به جستجوی شی خاصی در تصویر پرداخت و به محتویات یک تصویر دسترسی پیدا کرد. در ذیل حوزه‌های کاربردی که با استفاده از این استاندارد گسترش یافته اند ارائه شده‌است.

- ✓ کتابخانه‌های دیجیتال (کاتالوگ‌های تصویری، دیکشنری‌های صوتی و ...)
- ✓ رسانه‌های جمعی (کانال‌های رادیویی، کانال‌های تلویزیونی و...)
- ✓ بهبود ابزارهای چند رسانه‌ای (روزنامه‌های الکترونیکی، برنامه‌های چند رسانه‌ای شخصی و...)

می‌توان ویژگی‌هایی زیر را برای خانواده‌ی استانداردهای MPEG دانست [16].

- ✓ امکان دسترسی تصادفی به هر فریم از تصویر
- ✓ حرکت به جلوی فریم‌های تصویر با سرعت بالا
- ✓ هماهنگی کامل صوت و تصویر
- ✓ مقاومت در برابر خطا و اشتباه
- ✓ قابل ویرایش بودن
- ✓ انعطاف پذیری فرمت‌ها

## JVT ۳-۴-۲

استانداردهای ارائه شده توسط ITU-T اکثراً برای انتقالات تصاویر ویدیویی برخط<sup>۱</sup> مانند ویدیو کنفرانسها و تماسهای تصویری بوده است. در صورتیکه استانداردهای ارائه شده توسط MPEG اکثراً بر پایه میزان حجم ذخیره سازی ویدیو بر روی حافظه‌های جانبی، انتقال برنامه‌های تلویزیونی و ... متمرکز شده است. هر دو سازمان بصورت جداگانه ای بر روی استانداردهای متفاوت تصاویر ویدیویی فعالیت می‌کنند. استاندارد مشترک این دو سازمان، استاندارد H.262/MPEG-2 می باشد که با همکاری یکدیگر به توسعه رسید. این استاندارد توسط گروهی مشترک از هر دو سازمان با نام JVT<sup>۲</sup> ارائه گردید. در استاندارد جدید H.26L که توسط هر دو سازمان بصورت مشترک ارائه گردید ارسال تصاویر با کیفیت مناسب و بدون تاخیر مدنظر هر دو سازمان بوده است. در تصویر شماره ۱۹-۲ نحوه توسعه استانداردها توسط این دو سازمان در طی سال های گذشته نمایش داده شده است.<sup>۳</sup>

The story of two groups - MPEG and VCEG						
Year	MPEG	Part	Layer/Profile/Type	Usage	VCEG	Variants
1984	Not formed			Practically not useful	H.120	
1988	Not formed			Videoconferencing	H.261	
1993	MPEG-1			VHS and Television Recording		
		Part 1	Systems			
		Part 2	Video	VCD	H.261	
		Part 3	Audio			
			Layer I			
			Layer II			
			Layer III	MP3		
1999	MPEG-2			Broadcast, Distribution, DVD		
		Part 1	Systems			
			Program Stream			
			Transport Stream			
		Part 2	Video		H.262	HDV, XDCAM
		Part 3	Audio			
			Layer I			
			Layer II			
			Layer III	MP3		
2004	MPEG-4			Broadcast, Internet, Blu-ray		
		Part 1	Systems			
		Part 2	Video		H.263	HDCAM SR
		Part 3	Audio			
		Part 10	Advanced Video Coding	MPEG-4 AVC	H.264	AVCHD, XAVC
		Part 14	MP4 Container	MP4		
2013	MPEG-H	Part 2	Video	HEVC	H.265	

شکل ۱۹-۲: نحوه توسعه استانداردها در ۴ دهه‌ی اخیر

<sup>۱</sup> On-line

<sup>۲</sup> Joint Video Team

<sup>۳</sup> Copyright saresh sudhakaran 2013

فصل سوم:

روش نشرده سازی پیشنهادی

### ۳-۱ مقدمه

در فصل‌های پیش با ماهیت تصاویر ویدیویی مورد استفاده در سیستم‌های آموزش از راه دور توضیح داده شد. همانگونه که بیان شد این تصاویر ویدیویی دارای تغییرات بسیار کمی در حوزه‌ی زمانی خود می‌باشند. بدین معنا که در طول فریم‌های متوالی تغییرات یک پیکسل مشخص معمولاً ثابت و یا بسیار جزئی می‌باشد. بدین تصاویر، تصاویر ویدیویی کم تحرک می‌گویند. همچنین در فصل‌های گذشته با روش‌های متفاوت فشرده سازی تصاویر ویدیویی، فرمت‌های متنوعی که در این حوزه سازمان‌های استاندارد سازی ارائه کرده‌اند و همچنین کاربرد متدهای فوق بیان گردید. حال در این فصل می‌خواهیم برای فشرده‌سازی تصاویر ویدیویی مورد استفاده در آموزش از راه دور که در حیطه‌ی تصاویر ویدیویی کم تحرک قرار می‌گیرند روش ارائه دهیم. در فصل آتی ادعا خواهد شد که این روش برای فشرده سازی تصاویر ویدیویی کم تحرک از روش‌های موجود از لحاظ حجم مورد نیاز برای ذخیره و همچنین کیفیت تصاویر در حجم برابر با روش‌ها و متدهای موجود رقابت می‌کند و بهینه تر می‌باشد.

ما در این روش پیشنهادی، برای بهینه تر استفاده کردن از پهنای باند شبکه‌های اینترنتی که مسیر انتقال تصاویر آموزش از راه دور می‌باشند سعی در کاهش حجم این تصاویر می‌باشد. باید توجه داشت که در روش فوق برخلاف روش‌های متداول فشرده‌سازی که همواره با استفاده از تخریب تصویر، سعی در بهینه تر نمودن میزان فشرده سازی یک فایل ویدیویی را دارند، ما با حذف افزونگی‌های موجود در یک فایل ویدیویی به جلوگیری از ارسال اطلاعات تکراری پرداخته می‌شود. افزونگی‌ها میزان قابل توجهی از اطلاعات ویدیویی را تشکیل می‌دهند و همانگونه که در فصل اول به تشریح بیان گردید، هرچه میزان همبستگی مابین فریم‌ها در یک تصویر ویدیویی بیشتر باشد میزان افزونگی‌ها در یک فایل ویدیویی بالاتر خواهد رفت.

از طرفی نیز واضح است که تصاویر ویدیویی هرچه دارای تحرک کمتری مابین فریم‌های متوالی خود داشته باشند، دارای همبستگی بیشتری ما بین فریم‌های خود می‌باشند. می‌توان نتیجه گرفت که با جلوگیری از ارسال افزونگی‌های تصاویر ویدیویی کم تحرک می‌توان فشرده سازی را تا میزان بسیار بالایی افزایش داد و از حجم تصاویر ویدیویی تا میزان زیادی کاهش داد. هر چه میزان حجم یک فایل ویدیویی کمتر باشد انتقال این تصاویر در مسیرهای شبکه‌های اینترنتی با سرعت بالاتری صورت می‌پذیرد. می‌توان از پهنای باند استفاده نشده در هنگام انتقال تصاویر فشرده شده با این روش، به ارسال تصاویر ویدیویی با کیفیت بالاتر و یا ارسال اطلاعات اضافی و جانبی به همراه ویدیو اقدام کرد.

## ۲-۳ پیش پردازش

در این روش پیشنهادی در مرحله‌ی اول پیش پردازش‌های مقدماتی را بر روی تصاویر ویدیویی اعمال می‌شود. پیش پردازش به کلیه مراحل گفته می‌شود که قبل از اعمال متد اصلی بر روی فایل ویدیویی صورت می‌پذیرد. هدف پیش پردازش آماده سازی فایل ویدیویی برای کسب نرخ بهینه در روش فشرده سازی می‌باشد. به عبارت دیگر در مرحله‌ی پیش پردازش فایل خام ویدیویی با اعمال پردازش‌های اولیه و مورد نیاز به فایلی مناسب و آماده برای استفاده در مرحله‌ی بعدی تبدیل می‌گردد. در این مرحله تغییر عمده‌ای بر روی تصاویر ویدیویی صورت نمی‌گیرد و تنها با تغییرات حوزه‌های تصاویر (تغییر حوزه مکانی به حوزه فرکانسی) و یا جابجایی مقادیر (بازچینی دوباره‌ی تصویر) و یا جداسازی یک فایل به قسمت‌های گوناگون، سعی در آماده سازی تصاویر ویدیویی برای اعمال متد اصلی فشرده‌سازی را دارند.

## ۱-۲-۳ تبدیل مدل رنگی RGB به مدل رنگی Ycber

همانگونه که در فصل مقدمه گفته شد استفاده از مدل رنگی Ycber نسبت به مدل رنگی RGB در فشرده‌سازی دارای مزایای گوناگونی می‌باشد. یکی از مزایای این مدل رنگی یکنواخت بودن تغییرات در مولفه‌های این مدل رنگی می‌باشد، بدین معنا که اگر یکی از مولفه‌های Ycber را تغییر داده شود، تصویر

دچار تخریب زیادی نمی‌شود. در صورتیکه با تغییر یکی از مولفه‌های مدل رنگی RGB کل تصویر دچار تخریب زیادی شده و مقادیر رنگ‌های پیکسل‌ها کاملاً دچار تغییر می‌گردد. یکی دیگر از مزایای مدل رنگی Ycbcr را می‌توان به استخراج جزئیات تغییرات تصویر فقط از مولفه‌ی روشنایی Y دانست. می‌توان میزان تغییرات تصویر را از میزان تغییرات روشنایی نور تصویر بدست آورد، در صورتیکه برای بدست آوردن میزان تغییرات در مدل رنگی RGB به هر سه مولفه‌ی رنگ نیاز می‌باشد.

در روش‌های فشرده‌سازی همانگونه که در فصل مقدمه ذکر شد، تصاویر را از مدل رنگی RGB به مدل رنگی Ycbcr تغییر می‌دهیم. برای تغییر مدل رنگی تصاویر ویدیویی از استاندارد ۶۰۱ ITU استفاده می‌کنیم. روابط این انتقال در ذیل نمایش داده شده است.

$$\begin{bmatrix} Y_{601} \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} \quad (۱-۳)$$

در این مرحله تصاویر ویدیویی را از مدل رنگی RGB به مدل رنگی Ycbcr تبدیل شد. حال با توجه به نوع محتوای تصویر که شامل تصاویر آموزشی کم تحرک می‌باشند، می‌توان از فرمت ۴:۲:۰ مدل رنگی Ycbcr استفاده کرد. این نوع فرمت برای تصاویر ویدیویی بکار برده می‌شود که دارای تحرک متوسط و کمی می‌باشند و به ازای تعداد مشخصی نمونه از مولفه‌ی روشنایی فقط یک نمونه از مولفه‌های رنگی cb,cr استفاده می‌شوند. این عمل باعث کاهش حجم داده‌های تصاویر ویدیویی می‌شوند و با توجه به محتوای تصاویر کم‌تحرک استفاده از فرمت ۴:۲:۰ در مدل رنگی Ycbcr تاثیری بر کیفیت تصاویر ویدیویی نداشته و کیفیت این تصاویر را تقریباً ثابت نگه می‌دارد. در این نوع فرمت پس از انتقال ویدیویی فشرده شده مقادیر حذف شده از مولفه‌های رنگی cb,cr درون یابی شده و نمونه‌های درون یابی شده به جای مقادیر حذف شده قرار می‌گیرند.

در شکل شماره ۳-۱ مولفه های رنگی مدل RGB برای یک فریم از تصویر نشان داده شده است و در شکل شماره ۳-۲ مولفه های روشنایی و رنگ مدل Ycbcr برای همان فریم به نمایش درآورده شده است.



شکل ۳-۱: از راست به چپ، تصویر مولفه قرمز تصویر مولفه سبز و تصویر مولفه آبی از یک تصویر در مدل RGB



شکل ۳-۲: از راست به چپ، تصویر مولفه روشنایی Y و تصویر مولفه رنگی cb و تصویر مولفه رنگی cr در مدل رنگی

### ۳-۲-۲-۲ سکانس بندی یک فایل ویدیویی

در این مرحله با استفاده از یک معیار مشخص سعی در محاسبه‌ی میزان تغییرات در فریم‌های مجاور را داریم. هدف آن است که یک فایل ویدیویی را که شامل تعداد N فریم است را به قسمت‌هایی که هرکدام دارای تعداد  $M, X, Z, \dots$  فریم می‌باشند تقسیم نماییم.  $M, X, Z, \dots$  لزوماً دارای تعداد برابری فریم نمی‌باشند و تعداد فریم‌های تشکیل دهنده‌ی هرکدام از این زیر فایل‌های ویدیویی به میزان تغییرات در فریم‌های متوالی وابسته می‌باشند. به هرکدام از این زیر فایل‌ها که دارای تغییرات مشخصی می‌باشند یک

سکانس می‌گوییم. سکانس‌ها افزای از کل تصویر ویدیویی می‌باشند و با توجه به تعریف افزای می‌توان گفت که هیچ فریمی در تصویر نمی‌باشد که جزو یک سکانس مشخص نباشد و همچنین هیچ فریمی نیز از تصویر وجود ندارد که در دو سکانس قرار گرفته باشد.

سکانس‌ها  $\sum$  = فایل ویدیویی

$$\sum \text{سکانس } a \cap \text{سکانس } b = \Phi$$

هر سکانس به نحوی انتخاب می‌گردند که شامل تعدادی فریم باشد که دارای بیشترین همبستگی نسبت به یکدیگر در مقایسه با فریم‌های قبل و بعد سکانس مورد نظر باشند. در تصاویر ویدیویی معمولاً یک سکانس یک محیط مشخص را نمایش می‌دهد. محیطی که در آن تغییرات زیادی صورت نگرفته و تعداد پیکسل‌هایی که در طول فریم تغییر میکنند بسیار کم می‌باشند. هرگاه محیط در فایل‌های ویدیویی تغییر نماید سکانس نیز تغییر پیدا می‌کند و سکانس به تغییرات مقادیر پیکسل‌ها وابستگی زیادی دارد.

در تصاویر ویدیوهای آموزش از راه دور تغییر هر صفحه از مطلب مورد ارائه، تغییر محیط نمایش، ایجاد صفحه‌ی جدید، پاک کردن بخشی زیادی از محتوای تصویر، ترسیم تصاویر و یا متون بیش از حد معمول، بزرگ نمایی و یا کوچک نمایی بیش از حد متوسط و ... را می‌توان دلایلی بیان نمود که در تصاویر ویدیویی باعث تغییر سکانس می‌گردند.

در فایل‌های ویدیویی که از جنس تصاویر پرتحرک و یا با تحرک متوسط می‌باشند نمی‌شود از روش سکانس بندی استفاده کرد زیرا طول هر سکانس بسیار کوتاه و در حد چند فریم خواهد شد زیرا اینگونه تصاویر ویدیویی دارای تغییرات زیادی در مقادیر پیکسل‌های خود در طی چندین فریم متوالی می‌باشند. و تعداد سکانس‌های زیاد با طول کم همانگونه که در ادامه توضیح داده خواهد شد نه تنها باعث افزایش نرخ



فشرده سازی نمی‌گردد که در تصاویر با میزان تحرک متوسط و بالا باعث افزایش چشمگیر حجم فایل‌های فشرده خواهند شد.

دلایل تقسیم بندی یک فایل ویدیویی کم تحرک به تعدادی سکانس را می‌توان به دلایل ذیل دانست:

➤ اعمال روش فشرده سازی بر روی یک سکانس باعث نیاز کمتر به حافظه‌ی داخلی در حین پردازش می‌گردد. به علت حجم بالای افزونگی‌ها در تصاویر ویدیویی کمبود حافظه‌ی داخلی یکی از مشکلات عمده‌ی پردازش‌های فشرده سازی می‌باشد که با تقسیم یک فایل به تعدادی سکانس و اعمال روش‌های فشرده سازی در هر مرحله بر روی یک سکانس می‌توان این مشکل را رفع نمود.

➤ با تقسیم یک فایل ویدیویی به تعدادی سکانس که دارای همبستگی بسیار بالایی در مابین فریم‌های خود می‌باشند می‌توان باعث افزایش افزونگی در یک سکانس شد. این افزایش افزونگی در یک سکانس باعث کاهش بیشتر حجم و افزایش نرخ فشرده سازی می‌گردد. به عبارت دیگر می‌توان بیان نمود که جداسازی فریم‌های مشابه از فریم‌های دیگر و اعمال روش‌های فشرده‌سازی بر روی هر کدام از این سکانس‌ها باعث بهبود روند فشرده سازی تصاویر ویدیویی می‌گردند.

➤ در صورت پردازش با سخت افزارهای با حافظه‌ی اصلی مناسب می‌توان تعدادی سکانس را بصورت موازی و با استفاده از روش‌های فشرده سازی پردازش نمود. به علت اینکه هر سکانس بصورت کاملاً جداگانه‌ای مورد پردازش قرار گرفته‌اند و هیچگونه نیازی به اطلاعات سکانس‌های قبلی و بعدی ندارند می‌توانند بصورت موازی مورد پردازش قرار بگیرند و پیچیدگی زمانی و مدت پردازش را تا حد بسیار زیادی بهبود دهند.

باید توجه داشت که تبدیل یک فایل ویدیویی به چندین سکانس خود به دلیل دارا بودن سربرار اطلاعات اضافی برای هر سکانس حجم فایلی نهایی را اضافه خواهد کرد. اما بدلیل اینکه یک فایل ویدیویی که دارای فریم‌هایی با همبستگی متوسط است را به چندین ویدیو با میزان همبستگی بسیار بالا تبدیل کرده‌ایم و نرخ فشرده سازی را تا حد بسیار زیادی بالا برده‌ایم می توان از این سربرار چشم پوشی کرد.

نحوه‌ی تقسیم بندی یک فایل ویدیویی را به سکانس‌های متفاوت را می‌توان به شرح ذیل توضیح داد:

همانگونه که در قسمت قبل توضیح داده شد، مقادیر پیکسل‌های هر فریم از مدل رنگی RGB به مدل رنگی Ycbcr را تغییر دادیم. حال برای محاسبه‌ی ابتدا و انتهای هر فریم از مشخصه‌ی روشنایی Y استفاده می‌کنیم.

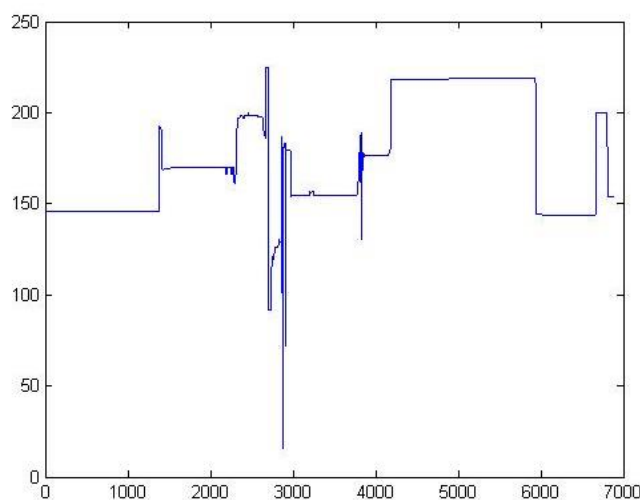
استفاده از این مشخصه همانگونه که در فصل اول توضیح داده شد بدین دلیل است که تغییرات روشنایی در هر پیکسل می‌تواند بیانگر بهتری از تغییرات در محیط باشد. در صورتیکه در مدل رنگی RGB تغییر مقادیر رنگ‌ها همیشه بیانگر تغییر در محیط نمی‌باشد و معیار خوبی برای ارزیابی تغییرات نمی‌باشد.

حال تمامی مقادیر پیکسل‌ها در مشخصه روشنایی را برای یک فریم محاسبه می‌کنیم. برای این کار کافی است برای فریم مشخصی، تمامی مقادیر ماتریس روشنایی (Y) را که هم ابعاد با پنجره تصویر می‌باشد را با یکدیگر جمع کرد. سپس مقدار حاصل شده را بر تعداد کل پیکسل‌های فریم تقسیم شود. در این صورت می‌توان میانگین مقدار روشنایی پیکسل‌های یک فریم را بدست آورد.

$$\text{Pixel\_mean} = 1/(M*N) \sum Y(x,y)$$

حال برای هر فریم میانگین مقدار روشنایی پیکسل‌ها را بدست آمده است. فریم‌هایی که در آنان سکانس تغییر می‌کند، فریم‌هایی می‌باشند که مقادیر میانگین روشنایی پیکسل‌ها نیز در آنان به میزان قابل ملاحظه‌ای تغییر کرده است. احتمال آنکه در یک تصویر ویدیویی مقادیر میانگین روشنایی پیکسل‌ها از

حد مشخصی بیشتر تغییر نکند ولی میزان تغییرات در یک فریم به حدی باشد که سکانس تغییر کرده است، به عبارت دیگر تغییرات در فریم زیاد باشد اما مقادیر روشنایی پیکسل‌ها به گونه ای تغییر کند که مقدار تغییرات را ثابت نگه دارد، به علت همبستگی موجود در مابین پیکسل‌های یک فریم از تصویر بسیار کم و تقریباً امکان پذیر نمی باشد. در شکل شماره ۳-۳ بردار مقادیر میانگین روشنایی ۷۰۰۰ فریم از یک تصویر ویدیویی آموزش از راه دور نشان داده شده است.



شکل ۳-۳: بردار مقادیر میانگین روشنایی ۷۰۰۰ فریم متوالی

می توان گفت در صورتیکه میانگین مقدار روشنایی پیکسل‌ها در دو فریم متوالی از حدآستانه‌ای<sup>۱</sup> بیشتر تغییر کند می توانیم ادعا نماییم که همبستگی میان آن دو فریم نسبت به همبستگی با فریم‌های قبلی کاهش پیدا کرده و در نتیجه فریم جدید متعلق به یک سکانس جدید می باشد.

برای محاسبه و پیدا کردن فریم‌های ابتدایی سکانس‌ها می توان از معیار همبستگی<sup>۲</sup> استفاده نمود و با محاسبه‌ی نسبت میزان همبستگی داخلی به میزان همبستگی خارجی و ایجاد یک حدآستانه سعی در جداسازی سکانس‌ها از یکدیگر داشت. به علت جم بالای محاسبات در این روش و نیاز به پردازش‌های

<sup>۱</sup>Threshold

<sup>۲</sup>Correlation Measure

زمان بر (به علت محاسبات معیار همبستگی) و حجم بالای اطلاعات موجود در یک ویدیو (به علت وجود درصد بالایی افزونگی) در این روش پیشنهادی از معیار ساده و کم هزینه‌ی معیار تفاضلی<sup>۱</sup> استفاده شده است.

دلیل استفاده از معیار تفاضلی بجای معیار همبستگی، محاسبات بسیار کمتر و پردازش سریعتر تصویر می‌باشد. باید توجه داشت که مسئله‌ی پیچیدگی زمانی یکی از مهمترین مباحث و نگرانی‌ها در حوزه‌ی فشرده‌سازی تصاویر ویدیویی می‌باشد.

برای استفاده از معیار تفاضلی کافی است میانگین مقادیر روشنایی هر فریم را از فریم بعدی‌اش کسر نماییم و در یک بردار جدید ذخیره نماییم. بردار موجود بیانگر میزان تغییرات میانگین مقادیر روشنایی هر فریم نسبت به فریم بعدی‌اش می‌باشد. حال کافی است مقادیری را که در این بردار بیش از حدآستانه می‌باشند و یا به عبارت دیگر فریم‌هایی که مقدار میانگین روشنایشان نسبت به فریم بعدی بیش از حدآستانه تغییر کرده است را بعنوان شماره‌ی فریم ابتدایی یک سکانس در نظر بگیریم و ویدیو را به سکانس‌هایی متفاوت که هر سکانس دارای بیشترین مقدار همبستگی مابین فریم‌های خودش و کمترین میزان همبستگی مابین فریم‌های سکانس بعدی و قبلی را دارا می‌باشد تقسیم بندی نمود.

### ۳-۲-۱ پیدا کردن حدآستانه برای سکانس بندی

برای بدست آوردن حدآستانه برای سکانس بندی یک فایل ویدیویی باید حتما از روش بدون ناظر<sup>۲</sup> استفاده نمود. برای این کار کافی است حدآستانه‌ای انتخاب گردد که اگر میزان تغییرات میانگین روشنایی فریم از این مقدار بیشتر بود، فریم به سکانس جدیدی تعلق پیدا نماید. می‌توان اعلام کرد که با توجه به

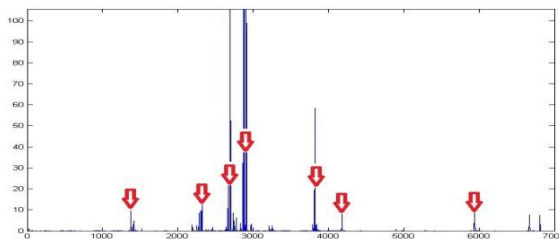
---

<sup>۱</sup>Differential Measure

<sup>۲</sup>Unsupervised

آزمایشات صورت گرفته بر روی ویدیوهای آموزش از راه دور ، حدآستانه‌ای را پیدا کرد که با استفاده از آن بتوان به خوبی و با دقت مطلوبی یک فایل ویدیویی را به تعدادی سکانس تقسیم کرد.

این حد آستانه با توجه به آزمایشات صورت گرفته بر روی پایگاه داده‌ای از تصاویر ویدیویی آموزش از راه دور مقدار ۱۰ بدست آمده است. این بدان معنا می‌باشد که در صورتیکه میزان تغییرات روشنایی دو فریم از عدد ۱۰ بیشتر بود باید فریم ثانویه به عنوان اولین فریم از سکانس جدید و فریم اولیه به عنوان آخرین فریم از سکانس قبلی در نظر گرفته شود. در تصویر شماره ۳-۵ معیار حدآستانه بر روی بردار تفاضلی میانگین روشنایی فریم‌ها اعمال شده و نقاط مشخص شده نقاط ابتدایی هر سکانس می‌باشند.



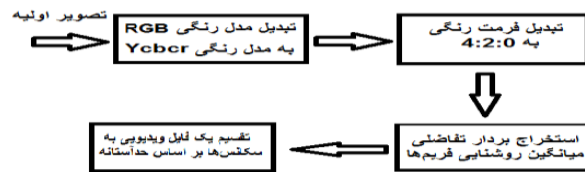
شکل ۳-۵: معیار حدآستانه بر روی بردار تفاضلی میانگین روشنایی فریم‌ها

می‌توان این حدآستانه را برای کلیه‌ی تصاویر ویدیویی کم تحرک آموزش از راه دور استفاده نمود. لازم به ذکر است که برای تصاویر ویدیویی کم‌تحرک در حوزه‌های دیگر مانند تماس‌های ویدیویی، ویدیو کنفرانس‌ها و ... نیز باید حدآستانه‌ای مجزا بدست آید و با استفاده از آن حدآستانه به سکانس بندی تصاویر ویدیویی اقدام نمود.

باید توجه داشت که حداقل طول هر سکانس باید ۲۱۰ باشد. در صورت عدم رعایت این حداقل طول تعدادی از سکانس‌ها دارای طول بسیار کوتاه بوده و این باعث سربار زیاد در روش فشرده سازی و کاهش نرخ فشرده‌سازی تصاویر ویدیویی می‌شوند.

### ۳-۲-۳ نتیجه گیری

حال می توان بیان کرد که تصاویر ویدیویی بر ای مراحل بعدی آماده است. در مرحله ی پیش پردازش تصاویر ویدیویی از مدل رنگی RGB به مدل رنگی Ycber تبدیل یافتند. سپس فرمت آنان با توجه به محتوای تصاویر آموزشی (کم تحرک) به ۴:۲:۰ تبدیل یافت. پس از تبدیل مدل رنگی در مرحله ی پیش پردازش یک فایل ویدیویی با معیار و حدآستانه ی مربوط به تصاویر ویدیویی کم تحرک آموزشی به سکانس هایی تقسیم گردید که سعی شده است هر سکانس بیشترین همبستگی را مابین فریم های خود و کمترین همبستگی را با فریم های سکانس های مجاور داشته باشد. حال تصاویر ویدیویی مدل رنگی Ycber به تعدادی سکانس با میزان تغییرات مشابه تقسیم گردیده است و فشرده سازی بر روی هر کدام از این سکانس ها اعمال می گردد. در شکل شماره ۳-۶ دیاگرام مراحل پیش پردازش نمایش داده شده است.



شکل ۳-۶: دیاگرام مراحل پیش پردازشی برای فشرده سازی تصویر بر اساس روش پیشنهادی

### ۳-۳ جایگزینی سکانس های بدون تغییر

در مرحله پیش پردازش تصاویر ویدیویی برای اعمال روش های فشرده سازی آماده شدند. حال در این مرحله فشرده سازی را بر روی تصاویر اعمال می شود. ورودی مرحله ی پیش پردازش، یک فایل ویدیویی با مدل رنگی RGB می باشد و خروجی مرحله ی پیش پردازش چندین سکانس با مدل رنگی Ycber می باشد. در مرحله ی پیش پردازش تنها فشرده سازی صورت گرفته تغییر فرمت مدل رنگی Ycber از ۴:۴:۴ به ۴:۲:۰ می باشد که با توجه به ماهیت کم تحرک بودن تصاویر ویدیویی آموزشی، براحتی می توان مقادیر حذف شده ی مولفه های رنگی cb,cr را پس از انتقال فایل فشرده، درون یابی نمود.

در مرحله‌ی جایگزینی سکانس‌های بدون تغییر، افزونگی‌های اصلی مابین فریمی را حذف می‌شود. این افزونگی‌ها شامل سکانس‌هایی می‌باشند که در طول مدت سکانس هیچ یک از فریم‌های سکانس تغییر نمی‌کنند. اینگونه سکانس‌ها در تصاویر ویدیویی دیگری به غیر از تصاویر ویدیویی آموزش از راه دور بسیار نادر می‌باشند و در صورت وجود تعداد فریم‌های این سکانس‌ها بسیار کم می‌باشند. باید در نظر داشت که سکانس‌های بدون هیچگونه تغییری در تصاویر ویدیویی آموزش از راه دور بسیار متداول بوده بصورت میانگین دارای تعداد فریم‌های نسبتاً بالایی می‌باشند. سکانس‌های بدون تغییر را می‌توان به لحظاتی که استاد در حال توضیح یک اسلاید و یا یک متن و یا صحبت بر روی یک تصویر و یا یک دیاگرام است نسبت داد. در این سکانس‌ها اطلاعات صوتی دارای تغییر زیادی می‌باشند اما اطلاعات تصویری هیچگونه تغییری را مابین فریم‌های خود ثبت نمی‌کنند.

در روش پیشنهادی برای حذف افزونگی سکانس‌هایی که در آنان فریم‌ها هیچگونه تغییری ندارند و بردار تفاضلی میانگین روشنایی فریم در طول سکانس در آنان صفر می‌باشد از کدک شماره ۱ استفاده می‌گردد.

### ۳-۳-۱ کدک شماره یک

کدک شماره یک روش فشرده‌سازی است که برای سکانس‌های بدون تغییر در روش پیشنهادی فشرده‌سازی ویدیوهای آموزش از راه دور استفاده می‌گردد. در این روش به جای ارسال اطلاعات پیکسل‌های هر فریم در طول سکانس فقط پیکسل‌های فریم اول ارسال شده و تعداد فریم‌های سکانس نیز به عنوان سربرار به فایل فشرده شده اعلام می‌گردد. بدین صورت در حین کدگشایی از فایل فشرده شده تصویر ارسال شده را به تعداد فریم‌های سکانس تکرار می‌نماید.

این روش نه تنها باعث افت کیفیت تصاویر ویدیویی نمی‌شود، بلکه باعث افزایش کیفیت تصاویر ویدیویی در مقایسه با دیگر روشها و فرمت‌های فشرده‌سازی استاندارد و شناخته شده می‌شود. در روشها و

استانداردهای فشرده‌سازی شناخته شده این افزونگی‌ها با توجه به مقادیر فریم‌های دیگر سکانس‌های قبلی و بعدی به کدک فشرده ساز تحویل داده می‌شوند و مقادیر این پیکسل‌ها با توجه به پیکسل‌های سکانس‌های مجاور فشرده شده و از مقادیر آنان تاثیر گرفته و دچار تغییر و تخریب می‌شوند. در صورتیکه مقادیر پیکسل‌های سکانس‌های بدون تغییر باید ثابت مانده و تغییری نکنند. در صورتیکه کدک شماره یک ارائه شده، هیچگونه پیکسلی تاثیری از پیکسل‌های سکانس‌های مجاور نمی‌گیرد و مقادیر اصلی پس از کدگشایی بازتولید می‌گردند. با این روش به جای ارسال اطلاعات یک فریم چندین بار (به طول سکانس) فقط یکبار اطلاعات ارسال گشته و باعث صرفه‌جویی در پهنای باند می‌گردد.

در کدک شماره یک در این مرحله فقط افزونگی‌های زمانی تصویر ویدیویی حذف می‌گردد. باید دانست که یک تصویر یا فریم دارای مقدار زیادی افزونگی مکانی ویا به عبارت دیگر افزونگی درون فریمی نیز می‌باشد که با حذف آنان می‌توان به نرخ بالاتری از فشرده‌سازی دست پیدا کرد. در این مرحله از کدک شماره یک می‌خواهیم افزونگی‌های درون فریمی را حذف نماییم. برای این کار کافی است از روش‌های متداول فشرده‌سازی تصویر استفاده شود.

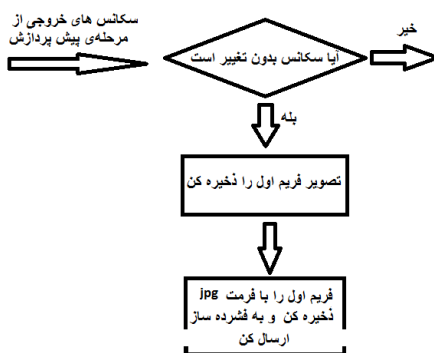
مطابق نتایج و محاسبات صورت گرفته بر روی تصاویر پایگاه داده‌ی تصاویر ویدیویی آموزش از راه دور می‌توان بیان داشت که بهترین فرمت فشرده‌سازی فریم اول یک سکانس فرمت jpg می‌باشد که دارای بهترین نرخ فشرده‌سازی تصویر نسبت به اتلاف مقادیر پیکسل می‌باشد که این مسئله به تفصیل در فصل نتایج و آزمایشات ارزیابی شده و دلیل انتخاب این فرمت فشرده‌سازی تصویر بیان شده است.

با استفاده از فرمت فشرده‌سازی تصویر jpg توانستیم افزونگی‌های مکانی یک سکانس را نیز در سکانس‌های بدون تغییر حذف نماییم. حذف افزونگی‌های مکانی و زمانی در سکانس‌های بدون تغییر باعث افزایش چشمگیر نرخ فشرده‌سازی تصاویر ویدیویی آموزش از راه دور می‌گردد. در کدک شماره یک به جای



انتقال یک سکنس که دارای تعدادی فریم بدون تغییر می‌باشند، یک عکس فشرده‌شده را به‌همراه تعداد فریم‌های سکنس منتقل می‌نماییم و سپس در قسمت کدگشا این تصویر را جایگزین تمامی فریم‌های سکنس شود.

باید توجه داشت که معیار انتخاب یک سکنس به عنوان سکنس بدون تغییر، دارا بودن حداکثر مقدار میانگین روشنایی فریم کمتر از ۰,۰۵ برای هر فریم می‌باشد. این مقدار ۰,۰۵ را می‌توان به عنوان میزان تغییراتی از تصویر که توسط چشم انسان قابل تشخیص نبوده و خطای دید محسوب می‌گردد برشمرد. مقدار این حدآستانه با آزمایش بر روی سکنس‌های بدون تغییر بدست آمده است. این حدآستانه مختص به تصاویر ویدیویی آموزش از راه دور بوده و مابقی تصاویر کم تحرک حوزه‌های دیگر دارای حدآستانه‌ی مشخص و مختص به خود می‌باشند. در شکل شماره ۳-۷ دیاگرام نحوه‌ی کارکرد کدک شماره یک نمایش داده شده است.



شکل ۳-۷: دیاگرام کدک شماره یک

### ۳-۴ گرید بندی سکانس‌ها

در بخش قبلی توانستیم سکانس‌هایی را که بدون تغییر می‌باشند را با ارسال تصویر فشرده فریم اول سکانس و تعداد فریم‌های موجود در سکانس فشرده نماییم. حال می‌خواهیم سکانس‌های دیگر فایل ویدیویی را که در آنان بردار تفاضلی میانگین روشنایی فریم‌ها دارای تغییرات می‌باشند را فشرده کرد. برای حذف بیشتر افزونگی‌های مکانی و زمانی موجود در یک سکانس با تغییر، در ابتدا باید اقداماتی را انجام دهیم. در یک سکانس با تغییر تمامی قسمت‌های سکانس دارای تغییرات نمی‌باشند بلکه بخش‌هایی از تصویر دچار تغییر شده و مابقی تصویر بدون تغییر در طول سکانس باقی می‌ماند. برای شناسایی مناطق پرتحرک از مناطق بی‌تحرک در یک سکانس همراه با تغییر باید سکانس را به بخش‌هایی تقسیم کرد. به این بخش بندی گریدبندی<sup>۱</sup> می‌گویند.

### ۳-۴-۱ گریدبندی با سائز پنجره‌های یکسان

در گرید بندی می‌توان از دو روش مرسوم استفاده نمود:

➤ روش اول عبارت است از بخش بندی پنجره‌ی فریم بصورت قسمت‌های مساوی و برابر. در این روش پیچیدگی زمانی بسیار کمی برای گریدبندی وجود دارد. طول پنجره‌ها از قبل مشخص می‌باشد.

---

<sup>۱</sup> Gridding

➤ روش دوم بنا به میزان تغییرات، گریدهای مجاور دارای تغییرات مشابه را یک گرید در نظر می‌گیریم. در نتیجه گریدهایی با سایز پنجره‌های بزرگ و پنجره‌هایی با سایز پنجره‌ی کوچک تشکیل می‌شود. این روش از لحاظ نرخ فشرده‌سازی بهینه‌تر می‌باشد.

مشکل اصلی روش گریدبندی با سایز پنجره‌های متفاوت در پیچیدگی بالای زمانی آن می‌باشد که باعث افزایش پردازش‌های فشرده‌سازی می‌گردد.

در شکل شماره ۳-۸ دو روش متفاوت گریدبندی در تصویر به نمایش گذاشته شده است.



شکل ۳-۸: تصویر سمت راست گرید بندی با طول پنجره‌ی برابر و تصویر سمت چپ گرید بندی با طول پنجره‌ی متفاوت

در مبحث فشرده‌سازی به علت وجود اطلاعات بسیار زیاد موجود در یک فایل ویدیویی و پردازش‌های بسیار زیاد موجود در یک فایل ویدیویی و پردازش‌های بسیار زیاد برای حذف افزونگی‌های زمانی و مکانی و روش‌های آنروپی سعی در کم کردن پردازش‌های روش‌های فشرده سازی دارد. به همین علت با گریدبندی پنجره‌هایی با سایز برابر از میزان کم افزایش نرخ فشرده‌سازی چشم پوشی کرده و مانع از افزایش پیچیدگی بیشتر زمانی می‌شود.

## ۲-۴-۳ محاسبه‌ی سایز پنجره‌های گریدبندی

در گریدبندی با سایز مساوی نیاز است که اندازه‌ی بلوک‌ها را مشخص نماییم. با مشخص کردن سایز بلوک‌ها می‌توانیم تعداد بلوک‌های هر سکانس را بدست آورد. هر چه سایز بلوک‌ها به سایز محدوده‌ی تغییرات در فریم نزدیک‌تر باشد و یا به عبارت دیگر محیط به تغییرات در فریم باشد، به فشردگی سازی با نرخ بالاتری دست پیدا کرد. هرچه سایز بلوک‌ها بزرگ‌تر از محدوده‌ی تغییرات در فریم و یا کوچک‌تر از محدوده‌ی تغییرات در فریم باشد فشردگی سازی با نرخ کمتری صورت می‌پذیرد. برای بدست آوردن سایز بهینه‌ی پنجره می‌توان از محاسبه‌ی میزان تغییرات در فریم و استفاده از الگوریتم‌های بهینه سازی استفاده نمود. به علت وجود داده‌های بسیار زیاد یک فایل ویدیویی و محاسبات پیچیده الگوریتم‌های بهینه‌سازی، در صورت محاسبه‌ی اندازه‌ی بلوک‌ها، پیچیدگی زمانی تا حد بسیار بالایی افزایش پیدا می‌کند و برای کاربردهای عمومی به هیچ وجه امکان پذیر نبوده و پیشنهاد نمی‌گردد. پیشنهاد می‌شود فقط در کاربردهایی که تصاویر ویدیویی به تعداد بسیار زیادی در بستر شبکه‌های اینترنتی استفاده می‌شوند از روش محاسبه‌ی سایز بلوک‌ها بکار گرفته شود.

در اینجا با محاسبه‌ی حیطه‌ی تغییرات در هر سکانس و آزمایش بر روی پایگاه داده‌های موجود تصاویر آموزش از راه دور با کیفیت AVI و کدک فشرده ساز H264، سایز هر بلوک را  $160 * 160$  در نظر گرفت. می‌توان ادعا نمود در صورت تبدیل یک سکانس به بلوک‌هایی با سایز  $160 * 160$ ، سربار هر بلوک برای سکانس به علت کوچک نبودن بیش از حد سایز بلوک مقدار زیادی نخواهد شد. این سایز بلوک همچنین برای کاربرد تصاویر آموزش از راه دور تضمین می‌دهد اکثر تغییرات در یک سکانس را احاطه نماید.

## ۳-۴-۳ محاسبه تغییرات در بلوک‌های یک سکانس

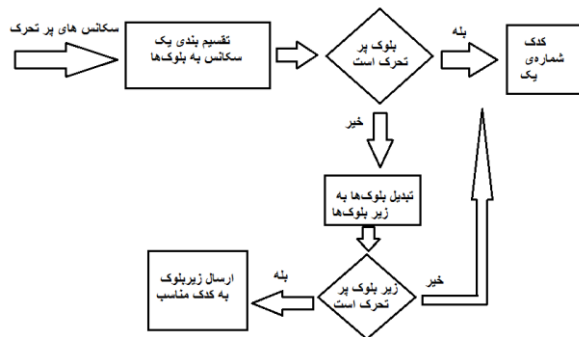
حال که یک سکانس را به بلوک‌هایی با سایز ثابت و اندازه‌ی  $160 * 160$  تبدیل شده است، باید به محاسبه‌ی تغییرات هر بلوک پرداخت. در یک سکانس همراه با تغییر، زمانیکه از روش گریدبندی استفاده

شود و یک سکانس را به بلوک‌هایی تقسیم کرد، بعضی از این بلوک‌ها بدون تغییر می‌باشد و بردار تفاضلی میانگین روشنایی در این بلوک‌ها برای تمامی فریم‌ها مقدار صفر می‌باشد. حال می‌توان این بلوک‌های دارای بردار تفاضلی میانگین روشنایی با مقدار صفر را به کدک شماره یک که در قسمت قبلی توضیح داده شد ارسال کرد و با استفاده از کدک شماره‌ی یک این بلوک‌ها نیز فشرده شده و نرخ فشرده سازی فایل ویدیویی را افزایش داده و افزونگی‌های موجود در سکانس‌های دارای تغییرات را نیز یک مرحله دیگر کاهش داد.

حال یک سکانس نیز تبدیل به تعدادی بلوک فشرده‌سازی شده و تعدادی بلوک نیز دارای بردار تفاضلی میانگین روشنایی غیر صفر می‌باشند. باید توجه داشت که در طول بلوک تمامی فریم‌ها دارای تغییر نمی‌باشند. بلکه بلوک‌های دارای مقدار بردار تفاضلی میانگین روشنایی غیر صفر در بعضی از فریم‌های خود دارای تغییر بوده و در بعضی دیگر از فریم‌های خود بدون تغییر می‌باشند. اگر بتوان توالی فریم‌هایی را که در یک بلوک بدون تغییر می‌باشند را از فریم‌های دارای تغییر جدا کرد، می‌توانیم ادعا نمود که تمامی افزونگی‌های موجود در یک تصویر ویدیویی را چه بصورت سکانس‌های بدون تغییر و یا بلوک‌های بدون تغییر یک سکانس و یا فریم‌های بدون تغییر یک بلوک را حذف شده است. برای رسیدن به این هدف کافی است بردار تفاضلی میانگین روشنایی یک بلوک را محاسبه کرده و در صورت پیدا کردن حداقل ۲۱۰ فریم (در قسمت قبل حد‌آستانه‌ی تعداد فریم‌های بدون تغییر را ۲۱۰ فریم محاسبه شده است) بدون تغییر در یک بلوک، آن را از بلوک جدا کرده و به عنوان یک زیربلوک به کدک شماره‌ی یک برای فشرده‌سازی ارسال کرد. با استفاده از این روش یک بلوک را نیز به چندین زیربلوک تقسیم بندی می‌کند که هرکدام از این زیربلوک‌ها دارای بردار تفاضلی میانگین روشنایی متفاوتی می‌باشند و بسته به مقدار این بردار به کدک‌های مناسب برای فشرده‌سازی ارسال می‌گردند.

### ۳-۴-۴ نتیجه گیری

در این قسمت با گریدبندی یک سکانس پرتحرک به تعدادی بلوک، در مرحله‌ی اول با حذف بلوک‌های دارای بردار تفاضلی میانگین روشنایی صفر و ارسال آنان به کدک شماره‌ی یک، افزونگی‌های مکانی و زمانی فریم‌های سکانس‌های پرتحرک را حذف کرد. سپس با محاسبه‌ی بردار تفاضلی میانگین روشنایی یک بلوک، آن را بر حسب مقادیر بردار تفاضلی میانگین روشنایی به زیر بلوک‌هایی تقسیم کرد. در این زیربلوک‌ها نیز با ارسال زیربلوک‌هایی با بردار تفاضلی میانگین روشنایی صفر به کدک شماره‌ی یک ادعا نمود که تمامی افزونگی‌های یک سکانس پرتحرک را حذف کرد. حال می‌توان با ارسال زیربلوک‌های یک سکانس پرتحرک بسته به میزان تحرک و مقدار بردار تفاضلی میانگین روشنایی به کدک‌های مناسب، این زیربلوک‌ها را نیز بسته به میزان تغییرات با کدک مناسب فشرده‌سازی نمود تا نرخ فشرده‌سازی تصویر ویدیویی را تا حد امکان افزایش داد. در شکل شماره‌ی ۳-۹ دیاگرام مراحل اجرا شده بر روی سکانس‌های دارای بردار تفاضلی میانگین روشنایی غیر صفر نمایش داده شده است.



شکل ۳-۹: دیاگرام گریدبندی سکانس‌های پرتحرک و ایجاد بلوک و زیربلوک‌های هر گرید

## ۳-۵ تابع تخریب

فرمت‌ها و استانداردهای فشرده‌سازی همانگونه که در فصول قبل بیان شد، بر دو نوع فرمت‌های فشرده-سازی تصاویر ویدیویی با اتلاف می‌باشند و یا فرمت‌های فشرده‌سازی تصاویر ویدیویی بدون اتلاف می‌باشند. در اکثر حوزه‌های فشرده‌سازی با توجه به ماهیت تصویر از فرمت‌های فشرده‌سازی با اتلاف استفاده می‌گردد. به علت تغییرات کم و محدود در تصاویر ویدیویی آموزش از راه دور و عدم نیاز به کیفیت بسیار بالا در این تصاویر از فرمت‌های با اتلاف برای فشرده‌سازی استفاده می‌گردد.

میزان اتلاف تصاویر و کیفیت تصاویر نیاز به یک حد تعادل<sup>۱</sup> دارد. هرچه میزان اتلاف داده‌های تصاویر بیشتر باشد، نرخ فشرده‌سازی نیز بیشتر می‌گردد ولی کیفیت تصاویر کاهش می‌یابد. هرچه میزان کیفیت تصاویر را افزایش دهیم، میزان اتلاف کاهش یافته و نرخ فشرده‌سازی کمتر می‌شود.

در روش پیشنهادی سعی ما بر این است که زیربلوک‌هایی را که مورد توجه کمتری می‌باشند را با میزان اتلاف بیشتر داده‌های تصویر فشرده کرده و زیربلوک‌هایی را که توسط بیننده بیشتر مورد توجه قرار می‌گیرند را با اتلاف کمتری فشرده نماییم. بنابر مشاهدات صورت گرفته در این زمینه، در تصاویر آموزش از راه دور هنگامی که در زیربلوکی از یک سکانس جزئیات بیشتر تغییر کرده و دارای بردار تفاضلی میانگین روشنایی با مقدار بیشتری می‌باشد، مطلب مهمی در آن زیر بلوک در حال بیان بوده و نیاز به توجه بیشتر بیننده را طلب می‌کند. در این هنگام توجه بیننده به زیر بلوک‌هایی که دارای بردار تفاضلی میانگین روشنایی کمتری می‌باشند، بسیار کمتر است.

به عنوان مثال در یک تصویر ویدیویی آموزش از راه دور زیربلوکی که تصویر ساعت سیستم عامل را نمایش می‌دهد دارای بردار تفاضلی میانگین روشنایی با مقدار بسیار کمتری نسبت به بردار تفاضلی

---

<sup>۱</sup>Trade Off

میانگین روشنایی زیربلوکی است که در حال نمایش نوشتارهای جدیدی بر روی متن می‌باشد. همچنین می‌توان ادعا نمود در هنگامی که متنی بر روی تصویر در حال رسم می‌باشد توجه بیننده به زیر بلوکی که ساعت سیستم عامل را نمایش می‌دهد بسیار کمتر می‌باشد. پس می‌توان نتیجه گرفت که می‌توان زیربلوکی را که دارای مقدار بردار تفاضلی میانگین روشنایی بیشتری می‌باشد را با نرخ فشرده سازی کمتری فشرده نمود و زیربلوکی را که دارای مقدار بردار تفاضلی میانگین روشنایی کمتری می‌باشد را با نرخ فشرده‌سازی بالاتری فشرده نمود و نرخ فشرده سازی تصویر ویدیویی آموزش از راه دور را افزایش داد. برای بدست آوردن مکان زیربلوک‌هایی که در زمان‌های مشخص نیاز به نرخ فشرده‌سازی کمتر و کیفیت بالاتری می‌باشند و زیربلوک‌هایی را که می‌توان با نرخ فشرده‌سازی بیشتر و کیفیت پایین‌تری فشرده نمود نیاز به یک معیار بدون ناظر می‌باشد. به همین جهت ما در این روش پیشنهادی تابع تخریبی را معرفی می‌نماییم که به کدک فشرده‌سازی اعلام نماید که هر زیربلوک را در هر بازه‌ی زمانی با چه نرخ می‌توان فشرده نمود.

این تابع تخریب باعث عدم تغییر کیفیت در زیربلوک‌های مهم شده و نرخ فشرده‌سازی را در زیربلوک‌هایی که از اهمیت پایین‌تری برخوردارند را افزایش می‌دهد. در نتیجه می‌توان ادعا نمود با استفاده از این تابع تخریب، هم کیفیت تصاویر ویدیویی را کاهش نیافته و هم باعث افزایش نرخ فشرده سازی تصاویر ویدیویی شده‌است.

### ۳-۵-۱ معرفی تابع تخریب


برای تعریف یک تابع تخریب نیاز به یک معیار برای تغییرات در زیربلوک‌ها می‌باشد. بهترین معیار برای محاسبه‌ی تغییرات در یک زیربلوک همانگونه که در مراحل قبلی بیان شد، استفاده از بردار تفاضلی میانگین روشنایی پیکسل‌ها می‌باشد. کافی است بتوانیم میزان تغییرات یک پیکسل را در طول چندین



فریم قبل و بعد محاسبه نمود. برای این کار کافی است از مجموع قدر مطلق تغییرات یک پیکسل در طول سکانس استفاده شود. در رابطه‌ی ذیل محاسبه‌ی معیار تغییرات یک پیکسل مشاهده می‌گردد.

$$\text{معیار تغییرات یک پیکسل} = \sum |B(x,y,z) - B(x,y,z+1)|$$

حال این مقادیر را برای تمامی پیکسل‌های موجود در هر زیر بلوک محاسبه و با یکدیگر جمع نموده و بر تعداد پیکسل‌های موجود در زیر بلوک تقسیم می‌شود. سپس مقدار کل هر زیر بلوک را در ماتریسی به ابعاد تعداد زیر بلوک‌های فریم قرار می‌گیرد. در شکل شماره‌ی ۳-۱۰ تصویر اصلی که به زیر بلوک‌هایی تقسیم شده است را به همراه ماتریس تغییرات روشنایی هر زیر بلوک نشان داده شده است.



منابع و مراجع مرتبط

کتابخانه دیجیتال دانشگاه گیلان  
مجموعه‌های نازی در دهکده پدیا

[http://en.wikipedia.org/wiki/Fuzzy\\_set](http://en.wikipedia.org/wiki/Fuzzy_set)

- کتاب «تئوری مجموعه‌های فازی و کاربردهای آن» نوشته‌ده مرتضی زاهدی

- Fuzzy sets and applications, By: Lotfi Asker Zadeh, Renald E. Yager

- Fuzzy set theory: foundations and applications, By: George J. Klir, Uta H. St. Ulrici, Be Yun

- Fuzzy set theory and its applications, By: Hans-Jürgen Zimmermann

0.1	0.1	0.1	0.1
0.12	0.12	0.12	0.12
14	14	14	14
31	31	31	31

شکل ۳-۱۰: نحوه‌ی محاسبه‌ی ماتریس تغییرات یک فریم

پس برای محاسبه‌ی مقدار تابع تخریب برای هر بلوک، کافی است مقدار هر بلوک را با مقادیر همسایه‌هایش جمع نماییم و به عنوان مقدار تابع تخریب برای آن بلوک در نظر گرفت. برای بلوک‌هایی که در حاشیه قرار داده شده‌اند کافی است مطابق شکل ۳-۱۱ بلوک‌های کناری را تعمیم داده و سپس به محاسبه‌ی مقدار تابع تخریب پرداخت.

0.1	0.1	0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1	0.1	0.1
0.12	0.12	0.12	0.12	0.12	0.12
14	14	14	14	14	14
31	31	31	31	31	31
31	31	31	31	31	31

0.96	0.96	0.96	0.96
42.66	42.66	42.66	42.66
135.36	135.36	135.36	135.36
228	228	228	228

شکل ۳-۱۱: سمت راست نحوه‌ی تمییم مقادیر بلوک‌های حاشیه‌ای و سمت چپ ماتریس تابع تخریب پس از محاسبه

حال برای هر بلوک مقدار تابع تخریب را بدست آمده است. کافی است با در نظر گرفتن میزان نرخ فشرده‌سازی تصویر ویدیویی به حد بندی مقادیر تابع تخریب پرداخت. این بدین معنا است که در صورتیکه مقدار تابع تخریب برای یک بلوک مقدار مشخص  $\alpha$  بود در آن هنگام میزان نرخ فشرده‌سازی کدک برای این بلوک به چه میزان خواهد بود. این معیار به میزان فشرده‌سازی تصویر ویدیویی وابسته بوده و می‌توان به عنوان یک انتخاب به اپراتوری که می‌خواهد با استفاده از این روش پیشنهادی به فشرده سازی تصویر ویدیویی آموزشی بپردازد، محول نمود. هرچه نرخ فشرده سازی بیشتر باشد می‌توان دامنه‌ی مقادیر تابع تخریب را به سمت حدودی پیش برد که دامنه‌ی مقادیر تابع تخریب کمتر بیشتر از دامنه‌ی مقادیر تابع تخریب بیشتر باشند.

مقادیر تابع تخریب در تمامی کدک‌ها مورد استفاده قرار گرفته و حتی در هنگامی که فریم اول یک گرید به‌جای یک زیر سکانس انتقال پیدا می‌کند (کدک شماره ۱) می‌توان با استفاده از این تابع تخریب، با افزایش نرخ فشرده‌سازی تصویر و کاهش کیفیت آن حجم فایل ویدیویی فشرده شده را کاهش داد.

## ۳-۶ کدک‌های زیربلوک‌های دارای تحرک

در قسمت قبل برای سکانس‌ها و زیربلوک‌های بدون تغییر از کدک شماره‌ی یک استفاده شده است. کدک شماره‌ی یک، تصویر فشرده شده (نرخ فشرده سازی تصویر توسط تابع تخریب اعلام می‌گردد) را به‌مراه تعداد فریم‌های سکانس و یا زیربلوک ارسال می‌گردند.

برای زیربلوک‌هایی که دارای بردار تغییرات میانگین روشنایی بیش از صفر می‌باشند، بسته به میزان بردار تغییرات میانگین روشنایی از کدک‌های متفاوتی استفاده خواهد شد که در ذیل به چند نمونه از این کدک‌ها اشاره خواهد شد. باید توجه داشت که ورودی هرکدام از این کدک‌ها ماتریس مشخصه‌ی روشنایی زیر بلوک و ماتریس‌های رنگ زیربلوک با فرمت ۴:۲:۰ می‌باشند.

## ۳-۶-۱ کدک شماره‌ی دو

کدک شماره‌ی دو، کدگذار و کدگشایی می‌باشد که برای زیر بلوک‌هایی استفاده می‌گردد که در آنان مقدار بردار تفاضلی میانگین روشنایی پیکسل‌های یک زیربلوک بیش از صفر باشد، اما این مقدار بسیار کوچک باشد. معمولاً زیربلوک‌هایی که به این کدک وارد می‌شوند، زیربلوک‌هایی هستند که در آنان در گوشه‌ای از بلوک تغییراتی جزئی ولی مهمی صورت گرفته است و باعث افزایش مقدار بردار تفاضلی میانگین روشنایی پیکسل‌های زیربلوک شده است. می‌توان به ترسیم نوشته با قلم نازک و یا تغییرات مهمی که در منطقه‌ی کوچکی از زیر سکانس قرار دارد و میانگین روشنایی تصویر را تا حد کمی تغییر داده که مناسب برای کدک شماره ۳ نمی‌باشد اما نمی‌توان آن را به کدک شماره ۱ انتقال داد، زیرا اطلاعات تغییرات پیکسل‌ها مهم بوده و نمی‌خواهیم حذف گردند.

در این کدک در مرحله‌ی اول از روش Run-length جهت حذف افزونگی‌های زمانی زیربلوک استفاده شده است. برای هر پیکسل در طول فریم‌های زیربلوک مقادیر را بدست می‌آورد. مقادیری را که بصورت

متوالی تکرار شده‌اند را همراه با تکرار آنان ذخیره کرده و مقادیری را که فقط یکبار در توالی فریم‌ها مشاهده شده‌اند را با مقدار تکرار ۱ ذخیره می‌نماید.

در این کدک با توجه به میزان تابع تخریب می‌توان مقادیر پیکسل‌های فریم‌های متوالی را دچار تغییر و تخریب نمود. بدین صورت که می‌توان برای ثبت مقدار یک پیکسل، اگر این مقدار با تفاوت  $\beta$  درجه روشنایی با پیکسل فریم قبلی بود آن را نیز با مقدار پیکسل قبلی ذخیره می‌نماییم. با این روش می‌توان فراوانی مقادیر متوالی پیکسل‌ها را بیشتر کرده و نرخ فشرده‌سازی را افزایش دهیم. در صورتیکه بخواهیم نرخ فشرده‌سازی را بالاتر برده و کیفیت را کاهش دهیم کافی است مقدار  $\beta$  را افزایش دهیم. در این صورت پیکسل‌های با تفاوت بیشتری را یکسان در نظر گرفته‌ایم.

### ۳-۶-۲ کدک شماره‌ی سه

در این قسمت می‌خواهیم تصاویر ویدیویی را که شامل تغییرات بیشتری نسبت به قسمت‌های دیگری از تصویر می‌باشند را فشرده نماییم. لازم به ذکر است که این تصاویر خود در حیطه‌ی تصاویر کم‌تحرك قرار می‌گیرند ولی در مقایسه با قسمت‌های دیگر تصاویر ویدیویی آموزش از راه دور دارای تغییرات بیشتری می‌باشند. همانگونه که در فصل پیش توضیح داده شد کدک H264 یکی از کدک‌های مناسب برای فشرده‌سازی تصاویر کم‌تحرك به شمار می‌آید. این کدک فشرده‌سازی، می‌تواند تصاویر را به خوبی فشرده کرده و به کیفیت تصویر ویدیویی کمترین آسیب را وارد نماید. توضیح بیشتر این کدک در فصل قبل بیان شده است. در روش پیشنهادی این کدک برای تصاویری به کار می‌رود که بیشترین تغییرات را در ویدیو داشته‌اند و با استفاده از این کدک گریدهای اسکانس‌های دارای تغییرات فشرده می‌گردند.

## فصل چهارم:

### نتایج و محاسبات روش پیشنهادی

## فصل چهارم

در این فصل می‌خواهیم روش پیشنهادی را که در فصل قبل به تفصیل بیان گردید را با اعمال بر روی چندین فایل ویدیویی آموزش از راه دور با روشهای گذشته و فرمت‌های فشرده‌سازی شناخته شده مقایسه کرده و مزایا و معایب هر روش را با استناد به مشاهدات و نتایج بدست آمده، بیان نماییم.

برای این منظور از فایل‌های ویدیویی آموزش از راه دور استفاده خواهیم نمود. متأسفانه به علت تجاری بودن این حوزه از تحقیقات و ورود شرکت‌های بزرگ نرم‌افزاری و سخت‌افزاری به این حیطه، پایگاه داده‌ی استاندارد در این زمینه موجود نبوده است. به همین منظور پایگاه داده‌ای برای تصاویر ویدیویی آموزش از راه دور جمع‌آوری شده است که در ذیل توضیح داده خواهد شد.

### ۴-۱ پایگاه داده تصاویر ویدیویی آموزش از راه دور

در اکثر مقالات و کارهای پژوهشی برای اعمال متدها و مقایسه‌ی آنان، از تصاویر ویدیویی استفاده می‌شود که بعنوان یک پایگاه داده استاندارد مطرح نمی‌باشند. بر روی یک تصویر ویدیویی متدهای مختلفی آزمایش شده و حجم فشرده‌سازی بر روی این تصاویر با در نظر گرفتن میزان اتلاف داده که با معیارهایی که در پیش توضیح داده شده است، مقایسه می‌گردد. و به ندرت مشاهده می‌شود که چندین مقاله بر روی یک پایگاه داده مشترک مقایسه متدهای گوناگون را اعمال نمایند. استفاده از پایگاه داده بیشتر در مقالاتی صورت پذیرفته است که نیاز به معتبرسازی بیشتری مانند پایان نامه‌ها و یا ... را دارا می‌باشند.

در حیطه فشرده‌سازی تصاویر ویدیویی، پایگاه داده‌ی منحصر بفردی موجود نمی‌باشد. بلکه در هر حیطه بسته به نوع و مشخصات ویدیو، پایگاه داده‌های متنوعی موجود است. تفاوت این پایگاه داده‌ها در چند مورد خلاصه می‌گردد. مهمترین عامل تمایز این پایگاه داده‌ها در نوع رفتاری اشیاء در تصاویر می‌باشد. بسته به نوع روش فشرده‌سازی که با چه روشی و بر روی چه تصاویری کاربرد دارد پایگاه

داده مربوطه انتخاب می‌گردد. تفاوت دیگر در پایگاه های داده نوع و مشخصات ویدیو می باشد که شامل سایز فریم، تعداد فریم‌ها در یک ثانیه، تعداد رنگ های موجود در تصاویر ، فرمت ذخیره سازی ویدیو و.. می باشد.

پایگاه داده‌های متنوعی در حوزه‌ی تصاویر ویدیویی موجود است اما پایگاه داده‌ی استاندارد که محتوی تصاویری با اشیاء تحرک کم باشند وجود ندارد. و با توجه به کارایی زیاد اینگونه تصاویر نیاز به وجود این پایگاه داده احساس می شود. و به همین خاطر پایگاه داده ای را که در ذیل توضیح خواهیم داد برای این منظور طراحی شده است.

این پایگاه داده [17] شامل ۹ ویدیو می‌باشد که از تصویر برداری بدون کاهش حجم از یکی از دروس دانشکده‌ی آموزش‌های از راه دور دانشگاه شاهرود برگرفته شده است. در این ویدیوها صحنه‌هایی ثابت، صحنه‌هایی با حرکت موس و یا یک نشانگر با ابعاد بزرگتر از موس، نوشتار بر روی متن و یا صفحه، ترسیم اشکال مختلف بر روی صفحه، تغییر سایز تصویر و بزرگنمایی بخش‌هایی از تصویر، تعویض صفحه‌های محتوی مطلب، تغییر در محتوای مطلب ارائه شده، نمایش نرم افزارها در حین ارائه درس، به اشتراک گذاری میزکار ویندوز و ... موجود می باشد. بعضی از این اتفاقات، تغییرات کمی را در تصاویر ویدیویی ایجاد کرده و بعضی دیگر تغییرات عمده‌ای را در تصاویر ویدیویی ایجاد می‌کنند. زمان اتفاق هر کدام از تغییرات معلوم نبوده و قابل پیش بینی نمی‌باشد. مکان تغییرات نیز در فریم از الگوی خاصی تبعیت نمی‌کند. در جدول ذیل چند تصویر از یک ویدیوی موجود در این پایگاه داده به نمایش گذاشته شده است.



شکل ۴-۱: چهار فریم متفاوت از یک فایل ویدیویی پایگاه داده

تصاویر ویدیویی این پایگاه داده شامل ۹ تصویر است که همگی دارای نرخ فریم ثابتی در یک ثانیه را دارا می باشند. این نرخ فریم برابر با ۲۴ فریم بر ثانیه می باشد. ابعاد تمامی تصاویر شامل ۱۲۰۰ \* ۸۰۰ می باشد و هر پیکسل تصویر سه بیت برای هر رنگ را شامل می شود. می توان گفت تعداد رنگ-های مختلف در این تصاویر شامل  $2^{24}$  و یا به عبارتی ۱۶ میلیون رنگ متمایز می باشد. مدت زمان ویدیوها و بقیه ی اطلاعات در جدول ذیل قرار داده شده است.

جدول ۴-۱: جزئیات تصاویر ویدیویی پایگاه داده

	bpp	Duration	Frame rate	Number of frames	height	width
Video_1	24	4002.7	23.976	80017	800	1280
Video_2	24	4313.1	23.976	86212	800	1280
Video_3	24	5891.3	23.976	117746	800	1280
Video_4	24	4708.9	23.976	94132	800	1280
Video_5	24	4487	23.976	89692	800	1280
Video_6	24	4711.1	23.976	94177	800	1280
Video_7	24	2043.4	23.976	40870	800	1280
Video_8	24	3829.8	23.976	76569	800	1280
Video_9	24	3875.8	23.976	77493	800	1280



## ۴-۲ محاسبه عملکرد روش پیشنهادی بر روی پایگاه داده

حال می‌خواهیم روش پیشنهادی را بر روی ویدیوهای آموزش از راه دور پایگاه داده اعمال کرده و با فرمت‌های شناخته شده‌ای که بر روی همین تصاویر اعمال شده‌اند به مقایسه بپردازیم.

فایل ویدیوی اولی که متد پیشنهادی را بر روی آن اعمال می‌نماییم، تکه‌ی اول از یکی از جلسات آموزش از راه دور می‌باشد. این تکه شامل ۱ دقیقه‌ی اول کلاس می‌باشد و نرخ فریم آن ۲۴ فریم بر ثانیه می‌باشد.

در مرحله‌ی اول این تصویر ویدیویی را از مدل رنگی RGB به مدل رنگی Ycbcr تبدیل می‌نماییم. برای این کار میزان پیچیدگی زمانی پردازش ما به علت ضرب ماتریسی از مرتبه‌ی  $X^2$  می‌باشد که  $X$  برابر با تعداد پیکسل‌های هر فریم می‌باشد. این پیچیدگی زمانی برای یک تبدیل تنها یک فریم از مدل رنگی RGB به مدل رنگی Ycbcr می‌باشد. در صورتیکه ما این عمل را برای تمامی فریم‌ها انجام می‌دهیم. پس با فرض اینکه تعداد فریم‌های فایل ویدیویی  $Z$  و هر فریم دارای  $x$  پیکسل در طول و  $y$  پیکسل در عرض باشد، پیچیدگی زمانی تبدیل مدل رنگی از مرتبه‌ی زمانی  $x^2y^2z$  می‌باشد.

تبدیل فرمت ۴:۴:۴ به فرمت ۴:۲:۰ به علت حذف تعدادی از مقادیر ماتریس‌های رنگی از پیچیدگی زمانی خطی  $x$  پیروی می‌نماید.

برای اعمال این روش فشرده سازی بر روی تصاویر ویدیویی پایگاه داده، بصورت انتخابی تعداد ۱۰ ویدیو با مدت زمان ۶۰ ثانیه و با احتساب ۲۴ فریم بر ثانیه، ۱۴۴۰ فریم تصویر انتخاب گردیده است. دلیل کوتاهی ویدیوها، حجم بالای داده‌ها می‌باشد که باعث پیچیدگی مکانی و زمانی محاسبات می‌باشد.

برای این منظور ۱۰ ویدیو که از مجموع تصاویر ویدیویی با تنوع پیچیدگی‌های تصویر از جمله: تغییر محتوای تصویر، کوچک و بزرگ شدن قسمتی از بافت تصویر، ترسیم اشکال بر قسمتی از تصویر و...

که در تصاویر آموزش از راه دور مرسوم می باشد انتخاب گردیده است. در ذیل با تجزیه و تحلیل محتوای ویدیوها و تغییرات آنان و اعمال روش های فشرده سازی و مقایسه ی حجم و کیفیت تصاویر فشرده شده با فرمت های دیگر فشرده سازی خواهیم پرداخت.

در ذیل جدول مشخصات این ۱۰ ویدیو به نمایش گذاشته شده است.

جدول ۴-۲: مشخصات ویدیو های پایگاه داده ها

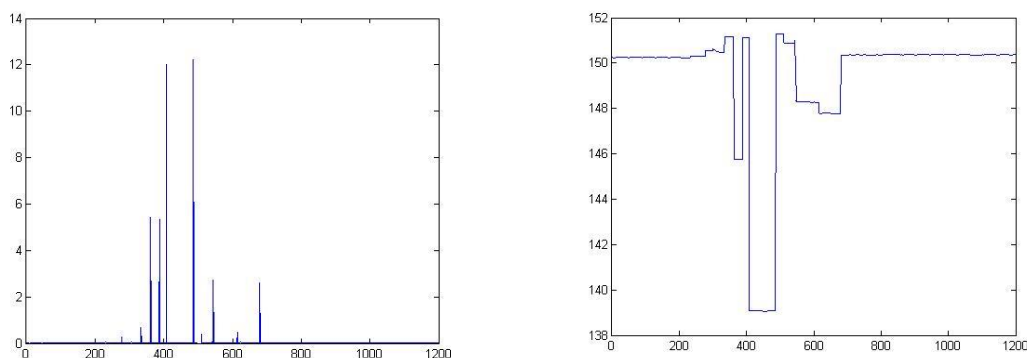
	<b>bpp</b>	<b>fps</b>	<b>duration</b>	<b>Size (bytes)</b>	<b>width</b>	<b>Height</b>
<b>Sample 1</b>	24	24	60	1900544	1280	800
<b>Sample 2</b>	24	24	60	1561709	1280	800
<b>Sample 3</b>	24	24	60	1620791	1280	800
<b>Sample 4</b>	24	24	60	1484875	1280	800
<b>Sample 5</b>	24	24	60	1388353	1280	800
<b>Sample 6</b>	24	24	60	1499278	1280	800
<b>Sample 7</b>	24	24	60	1546203	1280	800
<b>Sample 8</b>	24	24	60	1443939	1280	800
<b>Sample 9</b>	24	24	60	1576487	1280	800
<b>Sample10</b>	24	24	60	1538385	1280	800

در مرحله اول فشرده سازی فرمت رنگی تصاویر از RGB به Ycber تغییر پیدا می کند و نوع ۴:۲:۰ برای Ycber انتخاب می گردد. این نوع فرمت Ycber برای تصاویر ویدیویی کم تحرک متداول بوده و در فرمت های فشرده سازی شناخته شده برای این تصاویر بکار می رود. پیچیدگی زمانی این تبدیل شامل انتقال تمامی فریم های ویدیو توسط یک ماتریس دو بعدی می باشد که از پیچیدگی بالای برخوردار نمی باشد.

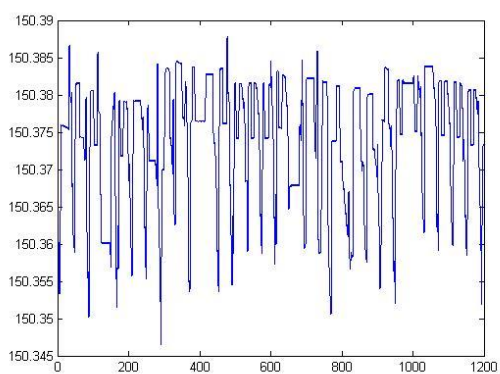
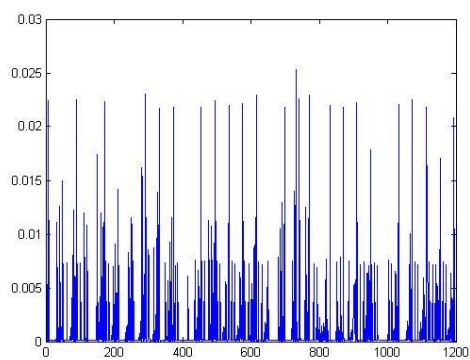
حال در این مرحله ماتریس سه بعدی روشنایی تمامی فریم های تصویر را در اختیار داریم. ماتریس cb و cr شامل مولفه های رنگی تصویر می باشند. این دو ماتریس با توجه به فرمت ۴:۲:۰ درون یابی شده و مقدار زیادی از اطلاعات موجود در آنان در مرحله کدگذاری حذف می شوند و سپس در مرحله کد گشایی با درون یابی دوباره بدست می آیند.

### ۳-۴ استخراج بردار میانگین روشنایی تصویر

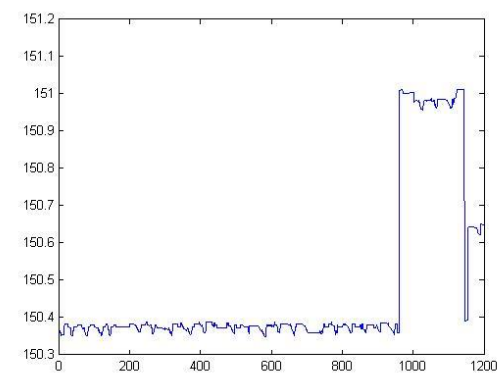
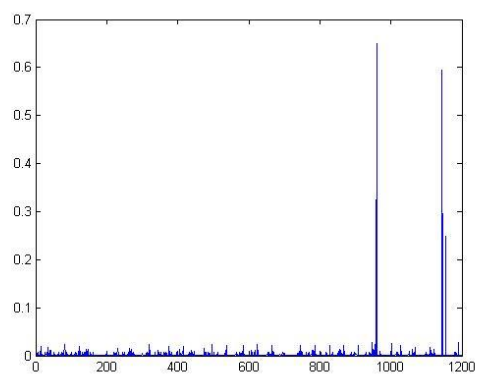
در این مرحله می خواهیم بردار تغییرات روشنایی نور را برای هر تصویر ویدیویی بدست آوریم تا بتوانیم تصویر ویدیویی را با توجه به میزان تغییرات روشنایی به سکانس هایی مجزا تبدیل نماییم. برای این منظور همانگونه که در فصل پیش بیان شد از میانگین قدرمطلق تغییرات روشنایی هر فریم استفاده می نماییم. پیچیدگی زمانی محاسبه ی میانگین قدرمطلق تغییرات روشنایی نسبت به روش های دیگر برای شناسایی تغییرات در هر فریم کمتر بوده و در تصاویر ویدیویی آموزش از راه دور (به علت عدم پیچیدگی تغییرات زیاد در اکثر فریم ها) مناسب می باشد. در تصاویر ذیل بردار میانگین روشنایی و بردار تغییرات میانگین روشنایی فریم های هر تصویر ویدیویی نمایش داده شده است.



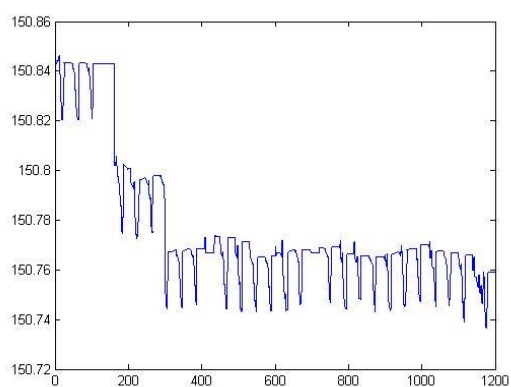
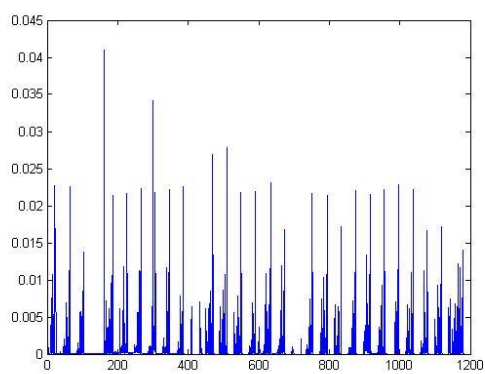
شکل ۳-۴: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۱



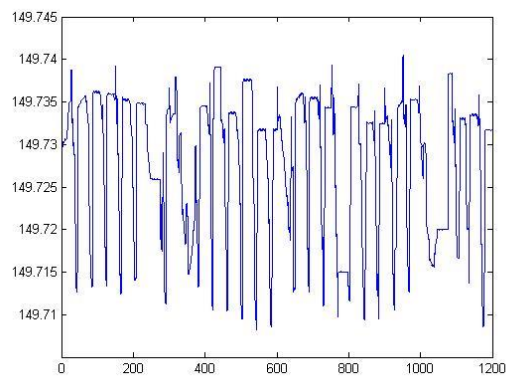
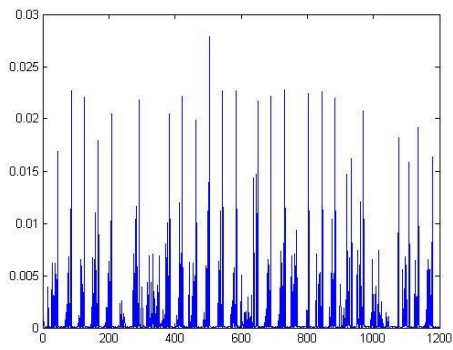
شکل ۳-۴: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۲



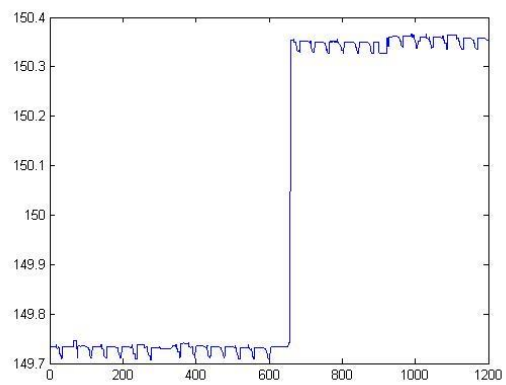
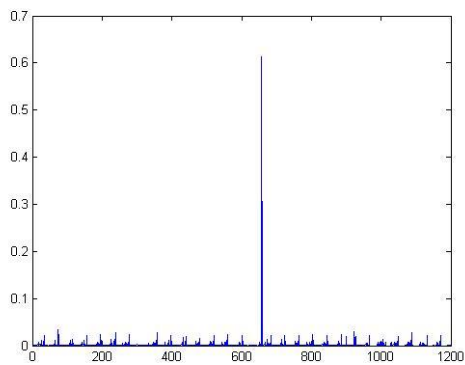
شکل ۴-۴: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۳



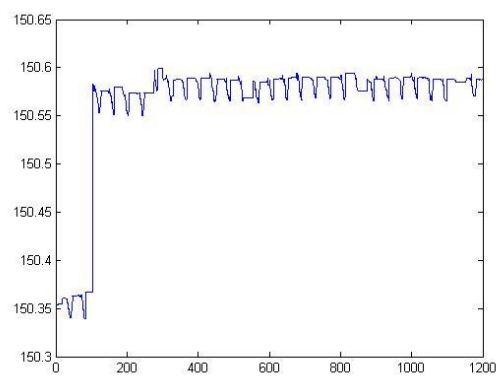
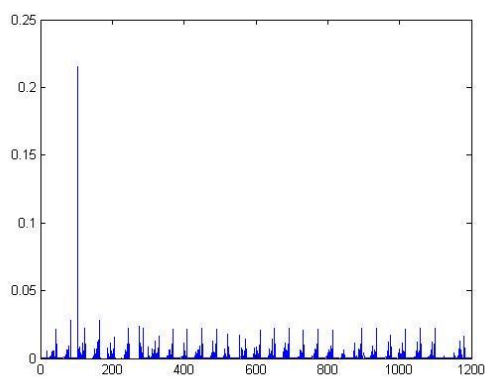
شکل ۵-۴: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۴



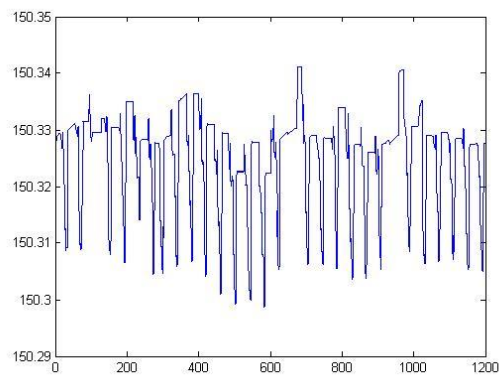
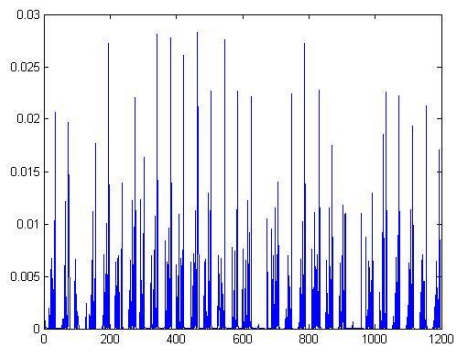
شکل ۴-۶: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۵



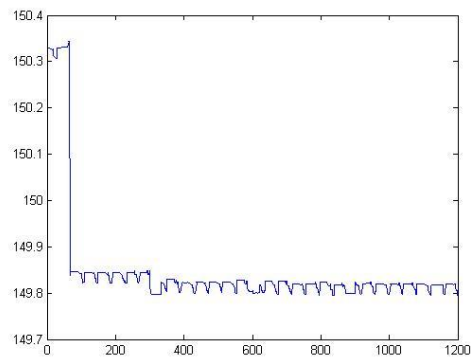
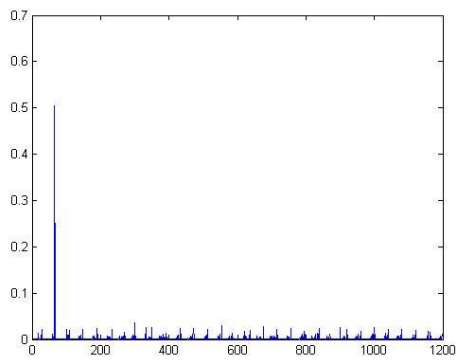
شکل ۴-۷: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۶



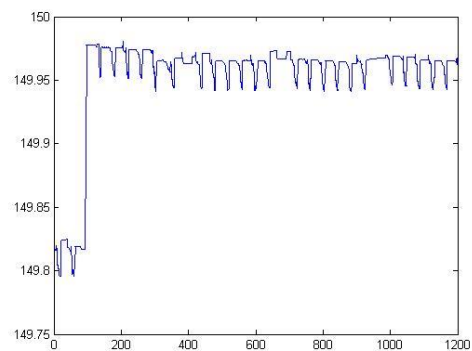
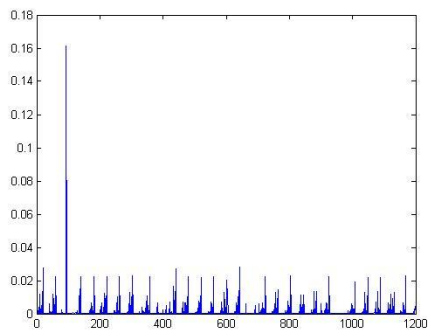
شکل ۴-۸: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۷



شکل ۴-۹: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۸



شکل ۴-۱۰: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۹



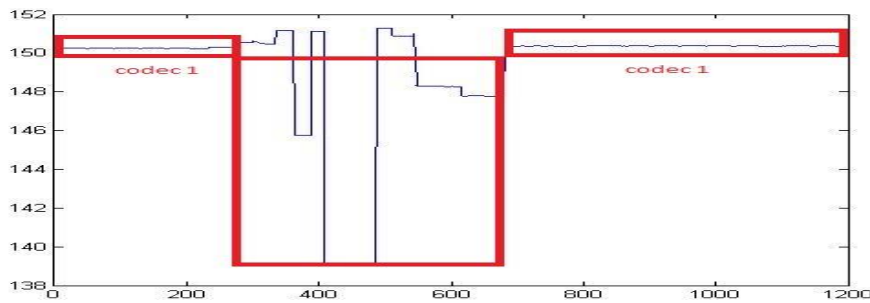
شکل ۴-۱۱: بردار میانگین روشنایی و میانگین تغییرات روشنایی ویدیوی شماره ۱۰

## ۴-۴ سکانس بندی تصاویر ویدیویی

در این مرحله می خواهیم به آنالیز جریان تغییرات در یک تصویر ویدیویی بپردازیم. در قسمت قبل در اشکال ۲-۴ تا ۱۱-۴ نحوه تغییر روشنایی در فریم های یک ویدیو مشاهده گردید. تغییرات تصاویر یک ویدیو با تغییرات میانگین روشنایی پیکسل های آن رابطه مستقیم دارد. می توان نتیجه گیری کرد هر چه تغییرات میانگین روشنایی های چند فریم متوالی بیشتر باشد در این فریم ها تغییرات بیشتری در تصویر رخ می دهد، و هر چه تغییرات میانگین روشنایی پیکسل های چند فریم کمتر باشد تغییرات اجزای تصویر در این فریم ها کمتر می باشد. در صورتیکه تغییرات میانگین روشنایی پیکسل های چند فریم برابر صفر و یا از حد آستانه ی مشخصی که بسیار نزدیک به صفر است تجاوز نکرد می توان نتیجه گرفت که این فریم ها بدون تغییر می باشند. همانگونه که در فصل پیش به تفصیل بیان شد، هرگاه تعداد مشخصی فریم دارای بردار تغییرات میانگین روشنایی پیکسل های برابر و یا حدود صفر باشند، می توان نتیجه گیری نمود که آن فریم ها بدون تغییر بوده و برای ذخیره سازی مقادیر این فریم ها می توان با استفاده از کدک شماره ۱ که در فصل پیش توضیح داده شده است، فقط یک تصویر از آن ذخیره نمود. اگر تغییرات بردار میانگین بیشتر از حد آستانه باشد برای گرید بندی به مراحل دیگر روش پیشنهادی ارجاع داده خواهد شد.

همانگونه که در بالا بیان شد انتخاب نوع کدک به محتوای تصویر وابسته می باشد. برای انتخاب کدک مناسب باید از محتوای تصویر مطلع باشیم، که این عمل بوسیله ی بردار میانگین تغییرات صورت می پذیرد. حال می خواهیم در این قسمت بردار میانگین روشنایی را برای سه ویدیو از پایگاه داده تجزیه و تحلیل نماییم.

#### ۴-۴-۱ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۱

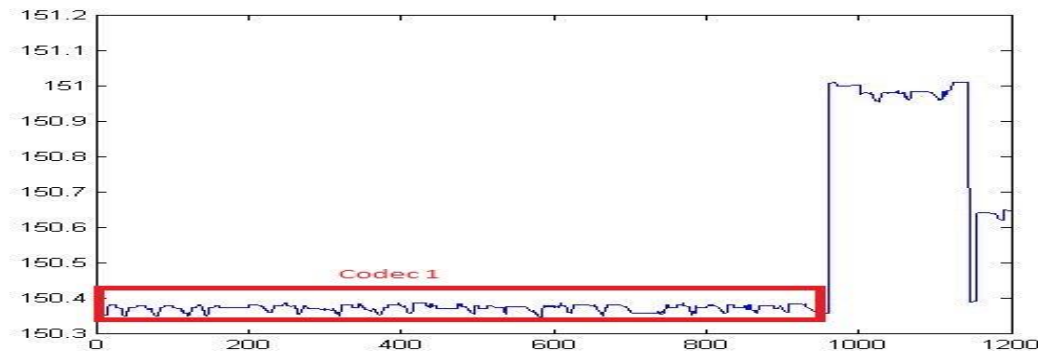


شکل ۴-۱۲: سکانس بندی ویدیوی شماره ۱

در ویدیوی شماره ۱ همانگونه که بردار میانگین روشنایی در تصویر بالا نشان می دهد، این ویدیو را می توان به سه سکانس تبدیل نمود. در سکانس اول تغییرات روشنایی بسیار کم می باشد و می توان آن را به کدک شماره ۱ جهت فشرده سازی ارسال نمود. در سکانس دوم تغییرات بعثت اجرای یک نرم افزار در تصویر ویدیویی و باز و بسته نمودن این نرم افزار، تصویر ویدیویی را شامل تغییرات نسبتاً زیادی در بردار میانگین روشنایی کرده است. در سکانس سوم استاد در حال تدریس صفحه جدیدی از پاورپوینت می باشد که تغییرات در این سکانس نیز به علت نزدیکی به صفر باعث شده است این سکانس نیز جهت فشرده سازی به کدک شماره ۱ ارجاع داده شود. این ارجاع باعث می شود بیش از ۶۳٪ از ویدیو با استفاده از کدک شماره ۱ به دو فریم تبدیل گردد و سکانس دوم نیز به مراحل بعدی فشرده سازی جهت کاهش افزونگی های زمانی و مکانی انتقال پیدا کند.



## ۴-۲-۴ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۳



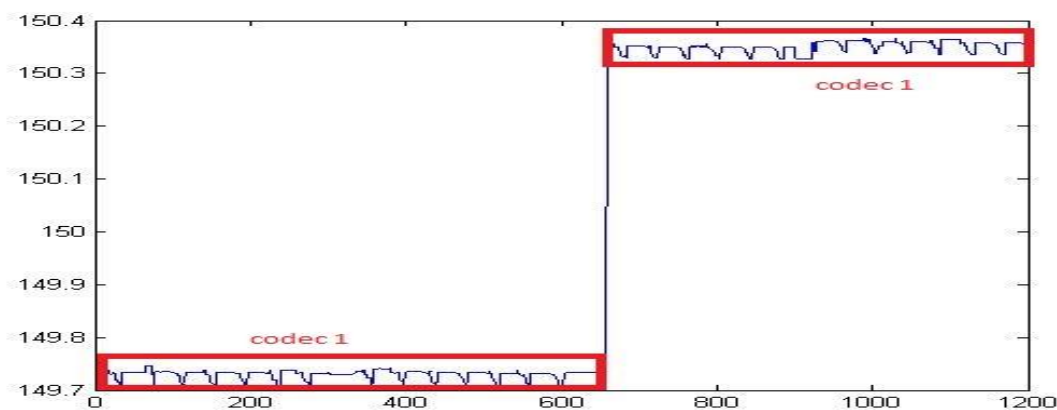
شکل ۴-۱۳: سکانس بندی ویدیوی شماره ۳

همانگونه که در بردار میانگین روشنایی ویدیوی شماره ۳ مشاهده می‌گردد تا فریم شماره ۹۴۰ تغییرات کمتر از ۰.۵٪ روشنایی برای هر پیکسل است. در صورت مشاهده ی تصویر ویدیویی تا زمان تصویر ثابت می باشد و استاد در حال حضور و غیاب کلاس درس می باشد. در این زمان تنها تغییر در تصویر چشمک زدن اشاره گر می باشد و این چشمک زدن باعث ۰.۵٪ تغییر میانگین روشنایی مقادیر پیکسل ها شده است. در صورتیکه این تعداد فریم را به کدک شماره ۱ انتقال دهیم ناظر بیرونی اشاره گر موس را بدون تغییر می بیند. باید توجه داشت این اشاره گر به هیچ عنوان به مفهوم و یا شی‌ای در تصویر اشاره نمی کند و حذف آن هیچگونه تاثیری بر محتوا و کیفیت ویدیو نخواهد داشت. با حذف این تغییر بسیار جزئی و به اهمیت می توان با استفاده از کدک شماره ۱ میزان بسیار زیادی از حجم داده‌ها را کاهش داد و به سرعت انتقال و استفاده بهینه از پهنای باند برای ارسال مابقی تصویر و یا بهبود کیفیت صوت استفاده نمود.

از فریم ۹۴۰ به بعد استاد صفحه تصویر را عوض می‌کند که این مسئله باعث تغییر سکانس تصویر ویدیویی می‌گردد. پس می‌توان با استفاده از روش فشرده سازی پیشنهادی این ویدیو را به دو سکانس تبدیل نمود. سکانس اول شامل فریم های ۱ تا ۹۴۰ و سکانس دوم شامل فریم های ۹۴۰ الی ۱۱۹۹. سکانس اول به علت تغییرات کمتر از ۰.۵٪ در میانگین روشنایی به کدک شماره ۱ انتقال پیدا کرده

و این سکانس از ویدیو را به یک تک فریم تبدیل می‌نماید و حجم آن را به میزان بسیار زیادی کاهش می‌دهد. سکانس دوم جهت ادامه عملیات فشرده سازی و گرید بندی به مراحل بعدی فشرده سازی ارجاع داده می‌شود. این کاهش حجم در سکانس اول هیچگونه کاهشی در کیفیت در محتوای آموزش از راه دور نداشته و باعث افزایش نرخ فشرده سازی می‌گردد.

#### ۴-۳-۴ آنالیز تغییرات بردار روشنایی در ویدیوی شماره ۶



شکل ۴-۱۴: سکانس بندی ویدیوی شماره ۶

همانگونه که در تصویر بالا مشاهده می‌گردد با استفاده از بردار میانگین روشنایی می‌توان تصویر ویدیویی را به دو سکانس تبدیل نمود. در این ویدیو به علت اینکه هر دو سکانس دارای بردار میانگین روشنایی کمتر از ۰.۵٪ روشنایی یک پیکسل می‌باشد می‌توان هر دو سکانس را به کدک شماره ۱ ارجاع کرد. در صورتیکه به تصویر ویدیویی مراجعه نماییم تغییر سکانس بعلاوه تغییر صفحه پاورپوینت می‌باشد و تغییرات بسیار کم به علت تغییر رنگ در shortcut مربوط به نرم افزار ضبط ویدیویی می‌باشد که در نوار ابزار ویندوز قرار داشته است و هیچگونه توجهی را به خود جلب نمی‌نماید. با ارسال این دو سکانس به کدک شماره ۱، دو تصویر از ابتدای هر فریم جایگزین این تصویر ویدیویی به تعداد ۱۱۹۹ فریم شده و باعث کاهش چشمگیر حجم ویدیو می‌گردد. این فشرده سازی هیچگونه کاهش

کیفیت در محتوای آموزشی از راه دور نداشته و باعث اختصاص پهنای باند بیشتری جهت ارسال مابقی ویدیو و یا افزایش کیفیت صوت می گردد.

#### ۴-۳-۴ نتیجه گیری

در این قسمت با تجزیه و تحلیل دلیل تغییرات میانگین روشنایی پیکسل های یک تصویر و نحوه ی سکانس بندی آنان آشنا شدیم. تصاویر به همراه بردار تغییرات میانگین روشنایی که صفر و یا نزدیک به صفر می باشد (حدآستانه در این پایگاه داده ۰.۵٪ در نظر گرفته شده است) به کدک شماره ۱ انتقال پیدا می کند. مقدار حدآستانه برای انتقال یک سکانس به کدک شماره ۱ بسته به میزان مورد نظر برای نرخ فشرده سازی و کیفیت تصویر فشرده شده قابل تغییر می باشد. در آزمایشات صورت گرفته بر روی پایگاه داده ۰.۵٪ برای حدآستانه در نظر گرفته شده است که می تواند بسته به نیاز تغییر پیدا نماید. تعداد حداقل ۲۱۰ فریم برای هر سکانس برای جلوگیری از ایجاد تعداد زیادی سکانس با تعداد فریم های کم در نظر گرفته شده است. دلیل این حداقل تعداد فریم برای جلوگیری از سربار ایجاد شده برای هر سکانس می باشد که در نرخ فشرده سازی تاثیر نامطلوبی خواهد داشت. هدف این است که تصاویر ویدیویی که با استفاده از کدک شماره ۱ میزان نرخ فشرده سازی آنان از سربار ایجاد شده مناسب تر ارزیابی می گردند به این کدک انتقال پیدا نمایند. تعداد ۲۱۰ فریم نیز با توجه به تصاویر این پایگاه داده به دست آمده است. در صورت زیاد کردن این تعداد فریم، تصاویری را که می توان با این کدک فشرده سازی نمود و نرخ فشرده سازی را افزایش داد، با اضافه شدن فریم های دارای تغییر به یک سکانس تا تغییرات بیشتر از حدآستانه تبدیل شده و با ارسال به کدک های دیگر باعث کاهش نرخ فشرده سازی می گردد.

## ۴-۵ ارزیابی کدک شماره ۱

در کدک شماره ۱ تنها تصاویر ویدیویی ارسال می‌گردند که هیچگونه تغییری در روشنایی پیکسل‌های موجود نداشته باشند. لازم به ذکر است چشمک زدن اشاره گر موس و یا تغییر در اعداد ساعت در نوار ابزار که تغییرات بی ارزش در فریم می‌باند قابل حذف می‌باشند. پس می‌توان سکانس‌هایی از تصاویر ویدیویی را که دارای میانگین تغییرات کمتر از ۰,۰۵ در پیکسل‌های روشنایی می‌باشند را به عنوان سکانس‌های بدون تغییر در نظر گرفت. از لحاظ محتوای ویدیویی این تصاویر شامل زمانهایی است که استاد در حال توضیح یک صفحه از مطلب درسی می‌باشد بدون آنکه با موس به مکان خاصی اشاره نماید و یا مطلبی را در صفحه اضافه و یا کم نماید.

در تصاویر ویدیویی آموزش از راه دور اینگونه سکانس‌ها از فراوانی مناسبی برخوردار می‌باشند. در مقایسه با تصاویر ویدیویی پر تحرک که اینگونه سکانس‌ها در آنان بسیار نادر می‌باشند در تصاویر ویدیویی آموزش از راه دور به دفعات اینگونه سکانس‌ها مشاهده می‌گردند.

لازم به ذکر است به علت جلوگیری از سربار تصاویر ویدیویی این سکانس‌ها حداقل شامل ۲۱۰ فریم می‌باشند که در آزمایشات متعدد بر روی پایگاه داده‌ها این مقدار حدآستانه برای طول سکانس‌ها حاصل گردیده است. در جدول ذیل میزان تغییرات در سکانس‌های هر ویدیو نمایش داده شده است.

جدول ۴-۲: میزان تغییرات هر سکانس ویدیوها

کدک شماره ۱	میانگین تغییرات روشنایی	تعداد فریم سکانس	شماره سکانس	شماره تصویر ویدیویی
قابل استفاده است	۰,۰۰۱۹	۲۵۰	سکانس ۱	ویدیو شماره ۱
قابل استفاده نیست	۰,۱۱۶۶	۳۴۷	سکانس ۲	
قابل استفاده است	۰,۰۰۲	۶۰۲	سکانس ۳	
قابل استفاده است	۰,۰۰۱۶	۱۱۹۹	سکانس ۱	ویدیو شماره ۲

ویدیو شماره ۳	سکانس ۱	۹۴۰	۰,۰۰۲۲	قابل استفاده است
	سکانس ۲	۲۵۹	۰,۰۰۸۳	قابل استفاده نیست
ویدیو شماره ۴	سکانس ۱	۱۱۹۹	۰,۰۰۱۳	قابل استفاده است
ویدیو شماره ۵	سکانس ۱	۱۱۹۹	۰,۰۰۱۲	قابل استفاده است
ویدیو شماره ۶	سکانس ۱	۷۷۸	۰,۰۰۲۱	قابل استفاده است
	سکانس ۲	۴۲۱	۰,۰۰۳۱	قابل استفاده است
ویدیو شماره ۷	سکانس ۱	۱۱۹۹	۰,۰۰۱۵	قابل استفاده است
ویدیو شماره ۸	سکانس ۱	۱۱۹۹	۰,۰۰۱۴	قابل استفاده است
ویدیو شماره ۹	سکانس ۱	۸۳	۰,۰۰۷۱	قابل استفاده نیست
	سکانس ۲	۱۱۱۶	۰,۰۰۱۶	قابل استفاده است
ویدیو شماره ۱۰	سکانس ۱	۸۰	۰,۰۰۱۷	قابل استفاده است
	سکانس ۲	۱۱۱۹	۰,۰۰۱۵	قابل استفاده است

در کدک شماره ۱ یک تصویر ویدیویی شامل حداقل ۲۱۰ فریم به یک فریم کاهش حجم داده و افزونگی زمانی موجود در آن از بین خواهد رفت. این حذف افزونگی باعث افزایش نرخ فشرده سازی تصاویر ویدیویی آموزش از راه دور می‌گردد. با این روش می‌توان اعلام نمود به علت عدم تغییر پیکسل‌های ارزشمند تصویر کاهش حجم تصویر ویدیویی بر کیفیت محتوای آموزشی هیچگونه تأثیری نخواهد داشت. می‌توان ادعا نمود هیچگونه اطلاعات مفید تصویر ویدیویی دچار تغییر نمی‌گردد و تغییرات ایجاد شده از دید کاربر مخفی بوده و افزایش نرخ فشرده سازی بدون کاهش کیفیت تصویر ویدیویی همراه خواهد بود.

از لحاظ میزان نرخ فشرده سازی و کاهش حجم تصاویر ویدیویی در جدول ذیل میزان کاهش هر سکانس که توسط کدک شماره ۱ فشرده شده است نمایش داده شده است. باید در نظر داشت که در قسمت قبل که یک ویدیو را به سکانس های متفاوتی بنا بر میزان تغییرات میانگین روشنایی تقسیم نموده ایم، میزان کمی سربار وجود داشته است که با توجه به میزان فشرده سازی توسط کدک شماره-۱ در این مرحله سربار بسیار ناچیز بوده و این کدک باعث کاهش قابل ملاحظه ای در حجم تصاویر ویدیویی گردیده است.

جدول ۴-۳: حجم تصاویر ویدیویی پس از اعمال کدک ۱ بر روی سکانس ها

شماره تصویر ویدیویی	شماره سکانس	حجم اولیه (مگا بایت)	حجم پس از اعمال کدک	درصد کاهش حجم
ویدیو شماره ۱	سکانس ۱	۰,۲۵	۰,۱۷۸	٪۵۳
	سکانس ۲	۱,۳۸	-	٪۰
	سکانس ۳	۲,۳۶	۰,۱۶	٪۹۳
ویدیو شماره ۲	سکانس ۱	۱,۵۲۶	۰,۱۵۱	٪۹۰
ویدیو شماره ۳	سکانس ۱	۱,۳۵	۰,۱۵۱	٪۸۸
	سکانس ۲	۲,۴۴	-	٪۰
ویدیو شماره ۴	سکانس ۱	۱,۴۵۱	۰,۱۴	٪۹۳
ویدیو شماره ۵	سکانس ۱	۱,۳۵۶	۰,۱۶۳	٪۸۸
ویدیو شماره ۶	سکانس ۱	۰,۹۲۷	۰,۱۶۳	٪۸۳
	سکانس ۲	۲,۹۶	۰,۱۷	٪۹۴

ویدیو شماره ۷	سکانس ۱	۱,۵۱	۰,۱۶۸	%۸۹
ویدیو شماره ۸	سکانس ۱	۱,۴۱۱	۰,۱۴۳	%۹۰
ویدیو شماره ۹	سکانس ۱	۰,۱۶۷	-	%۰
	سکانس ۲	۳,۶۲	۰,۱۸۳	%۹۵
ویدیو شماره ۱۰	سکانس ۱	۱,۵۰۳	۰,۱۷۶	%۸۸
	سکانس ۲	۳,۵۲	۰,۱۸۵	%۹۵
مجموع		۲۵,۳۱۲	۶,۱۱۸	%۷۶

#### ۴-۵-۱ نتیجه گیری

می توان با توجه به اطلاعات فوق نتیجه گیری نمود در صورتیکه قبل از اعمال کدک شماره ۱ یک تصویر ویدیویی، سکانس بندی نگردیده بود نمی توانستیم از کدک شماره ۱ به خوبی استفاده نماییم و مجبور بودیم برای فشرده سازی یک تصویر ویدیویی از کدک های دیگری که در مقایسه با کدک شماره ۱ از نرخ فشرده سازی بسیار کمتری برخوردار می باشند استفاده نماییم. در صورتیکه با سکانس بندی مناسب همانگونه که در نتایج بدست آمده مشاهده می گردد تعدادی از تصاویر ویدیو به چند سکانس که هر کدام دارای بردار میانگین تغییرات روشنایی نزدیک به صفر می باشند تبدیل شده اند. در صورتیکه اگر سکانس بندی صورت نمی پذیرفت این ویدیوها، تصاویر ویدیویی کم تحرک محسوب می شدند و قابلیت فشرده سازی با کدک شماره ۱ را که برای تصاویر ویدیویی بی تحرک می باشند، مناسب نبودند.

می توان نتیجه گیری کرد که بیشترین میزان فشرده سازی توسط کدک شماره ۱ صورت می پذیرد و در صورت سکانس بندی مناسب یک تصویر ویدیویی، می توان ادعا نمود که در تصاویر ویدیویی

آموزش از راه دور درصد قابل توجهی از سکانس ها توسط این کدک فشرده سازی خواهند شد. هر چه درصد سکانس های بی تحرک در یک تصویر ویدیویی بیشتر باشند نرخ فشرده سازی تصاویر ویدیویی آموزش از راه دور بصورت قابل ملاحظه ای افزایش پیدا خواهد کرد.

#### ۴-۶- گرید بندی سکانس ها

در این مرحله سکانس های ویدیویی که دارای تحرک بیش از حد آستانه ی کدک شماره ۱ می- باشند، گرید بندی می شوند. این گرید بندی برای تبدیل یک سکانس به تعدادی زیر سکانس می باشد که تعدادی از این زیرسکانس ها دارای بردار تغییرات میانگین روشنایی نزدیک به صفر می باشند که با ارسال دوباره به کدک شماره ۱ می توان آنان را فشرده سازی نمود. به عبارت دیگر این گرید بندی باعث می گردد که افزونگی های بخش هایی از تصویر که بی تحرک می باشند حذف گردد و تنها زیر سکانس هایی که دارای تحرک می باشند باقی بمانند. همانگونه که در فصل پیش بیان شد در تصاویر ویدیویی آموزش از راه دور این پایگاه داده سایز  $160 * 160$  پیکسل برای هر گرید مناسب می باشد. این سایز گرید بندی می تواند با توجه به سر بار ایجاد شده، نرخ فشرده سازی کدک شماره ۱ را بهینه نماید.

پس از گرید بندی سکانس های دارای تحرک که در این پایگاه داده دارای فریم هایی به سایز  $1280 * 800$  پیکسل می باشند، هر فریم تبدیل به  $8 * 5$  زیرسکانس می گردد که با توجه به به سر بار کم این تعداد زیر سکانس برای یک سکانس مناسب می باشد. در صورت افزایش تعداد زیرسکانس ها، نرخ فشرده سازی به علت ایجاد سر بار کاهش پیدا می کنند.

در این مرحله هر سکانس به ۴۰ زیر سکانس تبدیل می گردد. بردار تغییرات میانگین روشنایی را برای هر زیر سکانس بدست می آوریم. زیر سکانس هایی را که شامل بردار تغییرات میانگین روشنایی کمتر از حد آستانه ی کدک شماره ۱ می باشند را به کدک شماره ۱ ارسال می نماییم و زیرسکانس هایی را



که دارای بردار تغییرات میانگین روشنایی بالاتری می باشند و یا به عبارت دیگر زیرسکانس های با  
تحرك را به كدك شماره ۳ ارسال می نماییم.

می توان ادعا نمود پس از اعمال سکانس بندی و كدك شماره ۱ و گريد بندی و ارسال زیرسکانس های  
بی تحرك به كدك شماره ۱، توانسته ایم تمامی افزونگی های زمانی یک تصویر ویدیویی را از آن  
حذف نماییم. كدك شماره ۱ نیز افزونگی های مکانی زیر سکانس ها و سکانس های بی تحرك را از  
بین می برد و باعث افزایش نرخ فشرده سازی تا حد امکان در این مرحله می گردد. در جدول ذیل  
جدول میانگین تغییرات روشنایی گريدهای ۲ ویدیوی پایگاه داده نمایش داده شده است.

جدول ۴-۴: تغییرات روشنایی در گريدهای سکانس ۲ از ویدیوی شماره ۱ (یک سکانس پر تحرك)							
۰,۰۰۵	۰,۰۰۶	۰,۰۰۳	۰,۰۰۳	۰,۰۰۴	۰,۰۰۳	۰,۰۰۳	۰,۰۰۵
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۲۵	۰,۰۷۴	۰,۰۷۱	۰,۰۲۵	۰,۰۰۱	۰,۰۰۵
۰,۰۰۱	۰,۰۰۰۰	۰,۰۶۴	۰,۱۹۹	۰,۱۹۶	۰,۰۶۱	۰,۰۰۰۰	۰,۰۰۳
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۲	۰,۰۶۳	۰,۱۴۸	۰,۰۶۵	۰,۰۱۷
۰,۰۰۱	۰,۰۰۲	۰,۰۱۷	۰,۰۲۳	۰,۰۲۴	۰,۰۶۳	۰,۰۴۰	۰,۰۱۳

جدول ۴-۵: تغییرات روشنایی در گریدهای سکانس ۱ از ویدیویی شماره ۳ (یک سکانس کم تحرک)							
۰,۰۱۰	۰,۰۰۹	۰,۰۰۶	۰,۰۰۷	۰,۰۰۶	۰,۰۰۷	۰,۰۰۷	۰,۰۰۷
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۲	۰,۰۰۰۰	۰,۰۰۱	۰,۰۰۱	۰,۰۰۰۰	۰,۰۰۰۰
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۴	۰,۰۰۴	۰,۰۰۱	۰,۰۰۳	۰,۰۰۰۰	۰,۰۰۰۰
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۲	۰,۰۰۳	۰,۰۰۰۰	۰,۰۰۱	۰,۰۰۱
۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۱۴	۰,۰۰۲	۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۰۰	۰,۰۰۰۰

همانگونه که مشاهده می شود اکثر زیر سکانس های بدون تحرک بوده و می توان آنان را به کدک شماره ۱ جهت فشرده سازی انتقال داد. در تصاویر ویدیویی آموزش از راه دور در صد قابل توجهی از یک فریم شامل پیکسل هایی است که در طول یک سکانس بدون تغییر می باشند، اما به علت وجود تعدادی فریم که در تصویر تغییر پیدا می کنند نمی توان این سکانس ها را مستقیماً به کدک شماره ۱ ارسال نمود.

#### ۴-۷ تابع تخریب

همانگونه که در فصل پیش توضیح داده شد برای آنکه میزان کاهش کیفیتی که در روند فشرده سازی تصاویر ویدیویی صورت می گیرد به کیفیت محتوای آموزش آسیب وارد نکند و همچنین میزان افت کیفیت تصویر در مناطقی از فریم صورت بگیرد که کمترین توجه به آن مناطق می باشد و مناطقی از فریم که از لحاظ ارزش جزئیات تصویر مهمتر می باشند دچار افت کیفیت نگردند، از تابع تخریب استفاده می نماییم. در این تابع بسته به میزان تغییرات بردار روشنایی گریدهای هم جوار یک گرید، به هر گرید یک امتیاز تخریب تعلق می گیرد. هر چه این امتیاز پایین تر باشد میزان تخریب توسط تابع

فشرده‌سازی می‌تواند بیشتر باشد. تخریب بیشتر توسط تابع فشرده‌سازی می‌تواند باعث افزایش نرخ فشرده‌سازی و کاهش حجم ویدیو گردد. در جدول ذیل مقدار تابع تخریب برای یکی از ویدیوهای پایگاه داده به نمایش گذاشته شده است.

۰,۰۱۱	۰,۰۳۸	۰,۰۱۱۱	۰,۰۱۷۹	۰,۰۱۷۹	۰,۰۱۰۶	۰,۰۴۲	۰,۰۰۱۴
۰,۰۱۱	۰,۰۱۰۳	۰,۰۳۷۴	۰,۰۶۳۸	۰,۰۶۳۵	۰,۰۳۶۲	۰,۰۱۰۶	۰,۰۰۱۸
۰,۰۰۱	۰,۰۰۹۰	۰,۰۳۶۳	۰,۰۶۹۲	۰,۰۸۳۷	۰,۰۶۲۸	۰,۰۳۲۴	۰,۰۰۹۱
۰,۰۰۴	۰,۰۰۸۵	۰,۰۳۰۷	۰,۰۵۸۸	۰,۰۷۷۹	۰,۰۶۶۰	۰,۰۴۰۹	۰,۰۱۳۸
۰,۰۰۳	۰,۰۲۰	۰,۰۰۴۴	۰,۰۱۲۹	۰,۰۳۲۳	۰,۰۴۰۳	۰,۰۳۴۵	۰,۰۱۳۴

حال بسته به مقدار تابع تخریب برای هر گرید، نرخ فشرده‌سازی برای افزونگی‌های مکانی در کدک شماره ۱ بیشتر خواهد بود. می‌توان سه درجه برای این کاهش افت در نظر گرفت. درجه اول بدون کاهش (۱۰۰٪ کیفیت اولیه)، درجه دوم کاهش متوسط (۹۰٪ کیفیت اولیه) و درجه سوم کاهش زیاد (۸۰٪ کیفیت اولیه).

برای فشرده‌سازی تصویر هر گرید با توجه به مقدار تابع فشرده‌سازی صورت می‌گیرد. در ذیل برای هر گرید مقدار تابع تخریب و مقدار حجم هر گرید قبل و بعد از فشرده‌سازی نشان داده شده است.

جدول ۴-۷: تاثیر تابع تخریب بر حجم گریدها

شماره پنجره	مقدار تابع تخریب	حجم اولیه تصویر کیلو بایت	حجم پس از فشرده سازی	درصد فشرده سازی
۱-۱	۰,۰۰۱۱	۴,۷۷	۴,۴۲	٪۷
۲-۱	۰,۰۰۳۸	۳,۷۸	۳,۰۸	٪۱۹
۳-۱	۰,۰۱۱۱	۴,۲۲	۳,۵۳	٪۱۷
۴-۱	۰,۰۱۷۹	۳,۰۸	۲,۵۷	٪۱۶
۵-۱	۰,۰۱۷۹	۲,۵۳	۱,۸۳	٪۲۷
۶-۱	۰,۰۱۰۶	۲,۸۹	۲,۲۲	٪۱۴
۷-۱	۰,۰۰۴۲	۴,۱۹	۳,۵۴	٪۱۶
۸-۱	۰,۰۰۱۴	۲,۹۶	۲,۴۲	٪۱۹
۱-۲	۰,۰۰۱۱	۱,۱۴	۰,۷۶۱	٪۳۶
۲-۲	۰,۰۱۰۳	۱,۵۱	۱,۱۵	٪۲۴
۳-۲	۰,۰۳۷۴	۲,۵۷	۳,۵۳	-٪۳۷
۶-۲	۰,۰۳۹۲	۲,۴۷	۲,۴۰	٪۲
۷-۲	۰,۰۱۰۶	۳,۶۳	۳,۳۶	٪۸
۸-۲	۰,۰۰۱۸	۱,۳	۰,۹۶۸	٪۲۶
۲-۳	۰,۰۰۰۱	۱,۱۴	۰,۷۶۱	٪۳۶

٢-٣	٠,٠٠٩	١,٣٢	٠,٩٧٣	%٢٧
٧-٣	٠,٣٢٤	١,٤٥	١,٤	-%١٥
٨-٣	٠,٠٠٩١	١,٢٩	٠,٩٢٥	%٢٩
١-٤	٠,٠٠٠٤	١,٣٣	١,٠١	%٢٥
٢-٤	٠,٠٠٨٥	١,٣٢	٠,٩٧٣	%٢٧
٣-٤	٠,٠٣٠٧	١	٠,٤٧	%٣٣
٤-٤	٠,٠٥٨٨	١,٣١	١,٢٨	%٢
٨-٤	٠,٠١٣٨	١,٢٩	٠,٩٢٥	%٢٩
١-٥	٠,٠٠٠٣	٣,٠٤	٣,٠١	%١
٢-٥	٠,٠٠٠٢	٢,٩٢	٢,٥	%١٣
٣-٥	٠,٠٠٤٤	٣,١٥	٣,١٣	%١
٤-٥	٠,٠١٢٩	٣,٠٨	٢,٧١	%١٢
٥-٥	٠,٠٣٢٣	١,٩٣	١,٩٣	%٠
٧-٥	٠,٠٣٤٥	٢,٠٥	٢,٥١	-%٢٢
٨-٥	٠,٠١٣٤	٢,٨٤	٢,٤٢	%١٥
مجموع	-	٧١,٥٢	٤٣,١٠٤	%١٢

حال در این مرحله زیر سکانس های باقی مانده عبارتند از : زیر سکانس های پرتحرک و بسیار مهم تصویر که دارای ارزشمندترین پیکسل های تصویر می باشند. این زیرسکانس ها را توسط کدک شماره- ۲ فشرده سازی می کنیم و به علت اهمیت محتوای این زیر سکانس ها، سعی کرده ایم هیچ کاهش کیفیتی در این زیرسکانس ها صورت نگیرد و با همان کیفیت اولیه فشرده سازی شوند. در جدول ذیل حجم ۱۰ ویدیو پیش از فشرده سازی و پس از فشرده سازی به تفکیک سکانس ها قرار داده شده است.

جدول ۴-۸: تاثیر متد فشرده سازی بر حجم سکانس ها

شماره تصویر ویدیویی	شماره سکانس	حجم اولیه (مگا بایت)	حجم پس از اعمال کدک	درصد کاهش حجم
ویدیو شماره ۱	سکانس ۱	۰,۲۵	۰,۱۷۸	%۵۳
	سکانس ۲	۱,۳۸	۱,۱۲	%۱۹
	سکانس ۳	۲,۳۶	۰,۱۶	%۹۳
ویدیو شماره ۲	سکانس ۱	۱,۵۲۶	۰,۱۵۱	%۹۰
ویدیو شماره ۳	سکانس ۱	۱,۳۵	۰,۱۵۱	%۸۸
	سکانس ۲	۲,۴۴	۱,۸۲	%۲۶
ویدیو شماره ۴	سکانس ۱	۱,۴۵۱	۰,۱۴	%۹۳
ویدیو شماره ۵	سکانس ۱	۱,۳۵۶	۰,۱۶۳	%۸۸
ویدیو شماره ۶	سکانس ۱	۰,۹۲۷	۰,۱۶۳	%۸۳
	سکانس ۲	۲,۹۶	۰,۱۷	%۹۴
ویدیو شماره ۷	سکانس ۱	۱,۵۱	۰,۱۶۸	%۸۹
ویدیو شماره ۸	سکانس ۱	۱,۴۱۱	۰,۱۴۳	%۹۰

ویدیو شماره ۹	سکانس ۱	۰,۱۶۷	۱۶۷	%۰
	سکانس ۲	۳,۶۲	۰,۱۸۳	%۹۵
ویدیو شماره ۱۰	سکانس ۱	۱,۵۰۳	۰,۱۷۶	%۸۸
	سکانس ۲	۳,۵۲	۰,۱۸۵	%۹۵
مجموع		۲۵,۳۱۲	۵,۲۳۸	%۷۹

#### ۸-۴ ارزیابی نرخ فشرده سازی تصاویر ویدیویی

در این مرحله می خواهیم نرخ فشرده سازی تصاویر ویدیویی پایگاه داده را با روش پیشنهادی بدست آوریم. می توان بیان کرد که نرخ فشرده سازی هر تصویر ویدیویی رابطه ی مستقیمی با میزان تغییرات و پیچیدگی پیکسل های هر فریم و مابین فریمی وجود دارد و نرخ فشرده سازی وابسته به نوع تصویر می باشد. ما در این قسمت با انتخاب ۱۰ ویدیو از پایگاه داده و اعمال روش فشرده سازی پیشنهادی میزان کاهش حجم داده ها را بدست می آوریم. سپس بوسیله نرم افزارهای مطرح فشرده سازی ویدیویی این تصاویر را با فرمت های گوناگون فشرده سازی نموده و در ذیل به مقایسه ی نرخ فشرده سازی هر فرمت و مقایسه با نرخ فشرده سازی روش پیشنهادی پرداخته ایم.

در جدول زیر ۱۰ تصویر ویدیویی توسط فرمت های مطرح فشرده سازی و روش پیشنهادی کاهش حجم داده اند و نرخ فشرده سازی هر فرمت برای هر تصویر ویدیویی نمایش داده شده است.

جدول ۴-۹: مقایسه نرخ فشرده سازی فرمت های گوناگون									
	*H264	**H264/avc (intel)	نرخ فشرده سازی	***Wmv9	نرخ فشرده سازی	****H263	نرخ فشرده سازی	فرمت پیشنهادی	نرخ فشرده سازی

ویدیو شماره ۱	۱,۸۵۳	۲,۰۸۶	-%۱۲	۱,۲۶۴	%۳۲	۱,۸۸۴	-%۱	۱,۴۵۸	%۲۱
ویدیو شماره ۲	۱,۵۲۶	۱,۴۱	%۸	۰,۸۳۴	%۵۶	۱,۳۵۸	%۱۱	۰,۱۵۱	%۹۰
ویدیو شماره ۳	۱,۵۸۳	۱,۴۵۲	%۹	۰,۸۶۵	%۵۱	۱,۴۰۹	%۱۱	۱,۹۷۱	-%۲۴
ویدیو شماره ۴	۱,۴۵۱	۱,۴۲۸	%۲	۰,۸۱۸	%۵۴	۱,۳۳۶	%۸	۰,۱۴	%۹۳
ویدیو شماره ۵	۱,۳۵۶	۱,۵۲۲	-%۱۲	۰,۸۹۶	%۳۳	۱,۴۴۵	-%۶	۰,۱۶۳	%۸۸
ویدیو شماره ۶	۱,۴۶۵	۱,۶۳۱	-%۱۱	۰,۹۵۱	%۳۶	۱,۵۶۶	-%۶	۰,۳۳	%۸۷
ویدیو شماره ۷	۱,۵۱۰	۱,۶۳۱	-%۸	۰,۹۴۵	%۳۸	۱,۵۲۴	-%۱	۰,۱۶۸	%۸۹
ویدیو شماره ۸	۱,۴۱۱	۱,۴۰۱	%۱	۰,۷۹۵	%۵۴	۱,۳۵۰	%۵	۰,۱۴۳	%۹۰
ویدیو شماره ۹	۱,۵۴۰	۱,۷۷۷	-%۱۵	۱,۱۷	%۲۵	۱,۶۸۵	-%۱۰	۰,۳۸۷	%۷۵
ویدیو شماره ۱۰	۱,۵۰۳	۱,۶۷۸	-%۱۱	۰,۹۶۷	%۳۶	۱,۶۱۱	-%۷	۰,۳۶۱	%۷۶
مجموع	۱۵,۱۹۸	۱۶,۰۱۶	-%۵	۹,۵۰۵	%۳۸	۱۵,۱۶۸	%۱	۵,۲۳۸	%۶۶

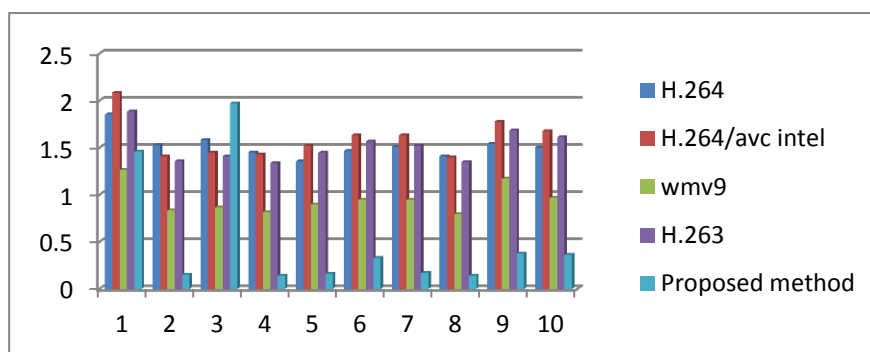
\*توسط فرمت Avi فشرده سازی شده است.

\*\* توسط فرمت Avi فشرده سازی شده است

\*\*\* توسط فرمت wmv فشرده سازی شده است. [20]

\*\*\*\* توسط فرمت Flv فشرده سازی شده است

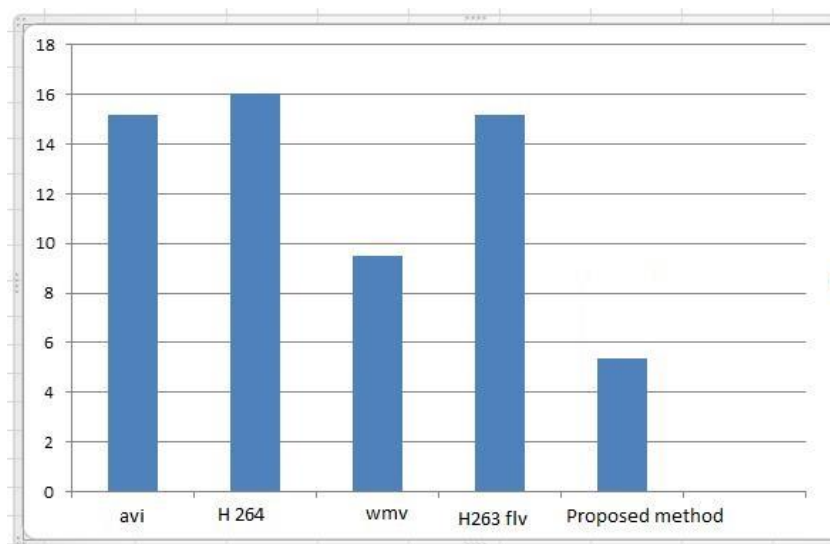
در تصویر ذیل فرمت های گوناگون فشرده سازی برای تصاویر ویدیویی به مقایسه گذاشته شده است.



شکل شماره ۴-۱۴ مقایسه حجم کدک های متفاوت فشرده سازی



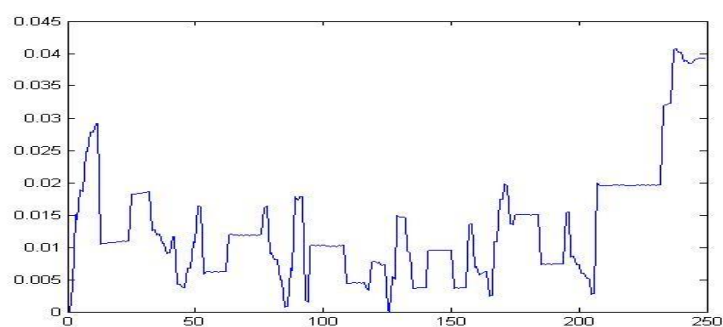
در تصویر زیر میزان حجم کلیه تصاویر ویدیویی که با فرمت های گوناگون فشرده سازی شده اند برای مقایسه قرار داده شده است.



شکل ۴-۱۵: مقایسه حجم نهایی ویدیو با فرمت های گوناگون

#### ۴-۹ ارزیابی کیفیت تصاویر ویدیویی قبل و بعد از اعمال فشرده سازی

برای مقایسه تغییرات مقادیر پیکسل، مقادیر تفاوت پیکسل های هر فریم از تصویر اصلی و تصویر فشرده سازی شده را محاسبه می نماییم و سپس قدر مطلق تفاوت را بدست می آوریم. در شکل شماره ۴-۱۶ نمودار تغییرات میانگین پیکسل های سکانس اول از ویدیویی شماره ۱ را با فریم متناظر در تصویر ویدیویی فشرده سازی شده با روش پیشنهادی نمایش داده شده است.



شکل شماره ۴-۱۶ مقادیر میانگین تغییرات روشنایی پیکسل ها در سکانس اول از ویدیوی اول پایگاه داده ها

همانگونه که مشاهده می شود دامنه‌ی تغییرات مقادیر روشنایی پیکسل ها برای این سکانس کمتر از ۰,۰۴ می باشد که این مقدار بسیار ناچیز بوده و توسط ناظر قابل درک نمی باشد.

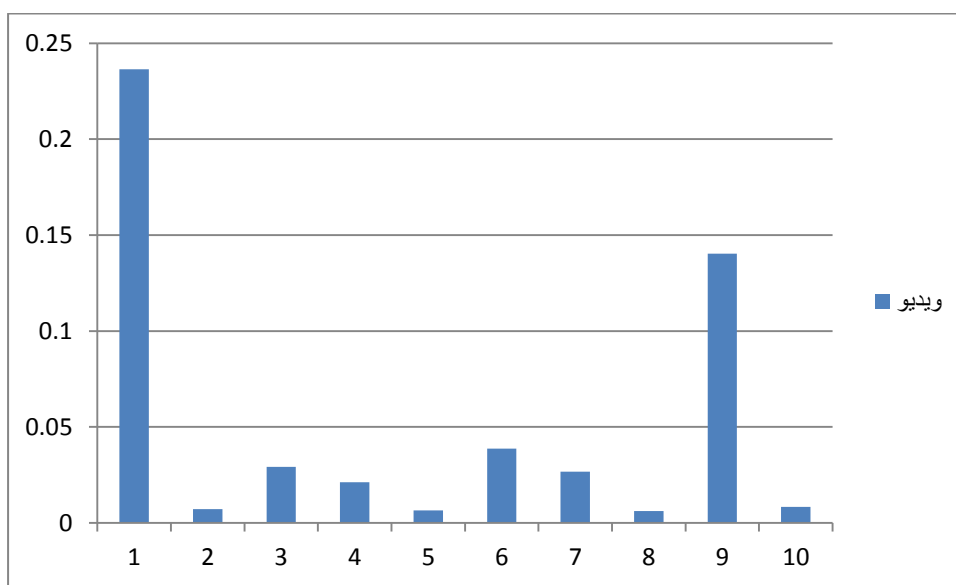
در جدول ذیل میزان میانگین تغییرات برای سکانس های مورد استفاده قرار داده شده است.

جدول شماره ۴-۱۰ میزان میانگین تغییرات روشنایی سکانس ها پس از فشرده سازی

شماره تصویر ویدیویی	شماره سکانس	میانگین تغییرات
ویدیو شماره ۱	سکانس ۱	۰,۰۱۲۹
	سکانس ۲	۰,۷۶۸۶
	سکانس ۳	۰,۰۲۲۶
ویدیو شماره ۲	سکانس ۱	۰,۰۰۷۲
ویدیو شماره ۳	سکانس ۱	۰,۰۵۳۰۲
	سکانس ۲	۱,۱۷۸۸
ویدیو شماره ۴	سکانس ۱	۰,۰۲۱۲
ویدیو شماره ۵	سکانس ۱	۰,۰۰۶۵
ویدیو شماره ۶	سکانس ۱	۰,۰۵۲۶۲
	سکانس ۲	۰,۰۱۳۰

ویدیو شماره ۷	سکانس ۱	۰,۰۲۶۸
ویدیو شماره ۸	سکانس ۱	۰,۰۰۶۲
ویدیو شماره ۹	سکانس ۱	۱,۷۰۲۳
	سکانس ۲	۰,۰۲۳۳
ویدیو شماره ۱۰	سکانس ۱	۰,۰۰۶۴
	سکانس ۲	۰,۰۰۸۵

در تصویر ذیل می توان میانگین تغییرات را برای هر ویدیو مشاهده نمود.



تصویر شماره ۴-۱۷ مقایسه ی میزان کاهش کیفیت در ویدیو ها پس از فشرده سازی

#### ۴-۹-۱ نتیجه گیری

همانگونه که در تصویر شماره ۴-۱۷ مشاهده می شود بیشترین میزان تغییرات در مقادیر پیکسل های تصاویر مربوط به تصاویری است که با کدک شماره ۳ فشرده سازی شده اند. (ویدیوی شماره ۱ و ۹) در

این تصاویر با استفاده از تابع تخریب، مقادیر تغییر یافته در گریدهایی از فریم که کمتر مورد توجه ناظر می باشند، قرار گرفته اند. می توان ادعا نمود کاهش کیفیت این تصاویر در مقایسه با کدک‌های دیگر هوشمندانه‌تر بوده و به جای تغییر مقادیر پیکسل‌ها بر مبنای افزایش نرخ فشرده سازی، بر حسب کاهش کیفیت بر مبنای توجه کمتر ناظر به گریدهای فریم صورت می پذیرد.

در روش پیشنهادی کاهش کیفیت در تصاویر ویدیویی در کدک شماره ۱ به علت عدم تغییر در محتوای تصاویر نامحسوس بوده و مقادیر جزئی تغییر در هر فریم در کیفیت تصویر ناچیز می باشد. در کدک شماره ۳ نیز بعلاوه توجه ناظر به گریدهای با ارزش، کاهش کیفیت در گریدهایی با تغییرات کم، از دید ناظر پوشیده می باشد و به کیفیت تصویر از دید ناظر آسیبی وارد نمی نماید.

فصل پنجم:

نتیجہ گیری و کارهای آتی

## فصل پنجم

### ۵-۱ نتیجه گیری

مهمترین عامل در فشرده سازی تصاویر ویدیویی حذف افزونگی های زمانی و مکانی مابین پیکسل های یک ویدیو می باشد. با حذف این افزونگی ها می توان نرخ فشرده سازی را تا حد بسیار زیادی افزایش داد. وجود این افزونگی ها به وابستگی و ارتباط بین پیکسل های یک فریم و همچنین مقادیر پیکسلدر فریم های متوالی دانست. می توان از این دیدگاه ویدیو ها را دسته بندی نمود. ویدیوهای پرتحرک دارای وابستگی کمتری مابین پیکسلها نسبت به ویدیوهای کم تحرک می باشند. تصاویر ویدیویی آموزش از راه دور نیز در دسته ی ویدیوهای کم تحرک قرار میگیرند. در صورت عدم لحاظ میزان وابستگی های پیکسل های یک ویدیو در انتخاب نوع فرمت فشرده سازی می تواند ما را از دستیابی به نرخ فشرده سازی بالاتری محروم نماید. می توان نتیجه گرفت برای فشرده سازی بهتر باید نوع ویدیو را در انتخاب فرمت فشرده سازی لحاظ نمود. فشرده سازی تصاویر ویدیویی آموزش از راه دور با فرمت تصاویر پرتحرک باعث کاهش نرخ فشرده سازی این ویدیوها می گردد. تصاویر ویدیویی آموزش از راه دور اکثرا در بستر وب انتقال پیدا میکنند و انتخاب یک فرمت مناسب در دستیابی به نرخ فشرده سازی بالاتر می تواند در کیفیت انتقال این تصاویر تاثیر بسزایی داشته باشند.

روش فشرده سازی پیشنهادی با توجه به نوع کم تحرک تصاویر ویدیویی در ابتدا افزونگی های زمانی و مکانی فریم ها را تا حد امکان کاهش می دهد تا به نرخ فشرده سازی بالاتری دست پیدا نماید. فشرده سازی های ویدیو که دارای نرخ مطلوب تری می باشند دارای کاهش کیفیت، به علت تغییرات در مقادیر پیکسل ها می باشند. در فرمت های متداول فشرده سازی این کاهش کیفیت با معیار افزایش نرخ فشرده سازی صورت می گیرد و تصویر ویدیویی پس از فشرده سازی دارای افت کیفیت، بدون در نظر گرفتن

محتوای تصویر می‌باد. در روش فشرده سازی پیشنهادی این کاهش کیفیت برای افزایش نرخ فشرده سازی با استفاده از تابع تخریب صورت می‌پذیرد. تابع تخریب مقادیر پیکسل‌ها را با در نظر گرفتن فاصله مکانی یک پیکسل با پیکسل‌های دارای تغییرات عمده که قسمت‌های همه فریم می‌باشند اعمال می‌گردد. هر چه یک پیکسل از کانون تغییرات فاصله‌ی بیشتری داشته باشد تابع تخریب، آن قسمت از تصویر را با نرخ بالاتری فشرده سازی می‌کند.

در تصاویر آموزش از راه دور در هنگامی که استاد در حال نوشتن در قسمتی از فریم می‌باشد و یا در قسمتی از فریم در حال تغییرات در محتوای تصویر باشد توجه بیننده به مکان‌های دیگر تصویر کمتر بوده و می‌توان پیکسل‌های آن مناطق را با نرخ بالاتری فشرده سازی نمود بدون آنکه بیننده متوجه تغییر کیفیت در تصویر شود.

در تصاویر ویدیویی آموزش از راه دور درصد قابل توجهی از فریم‌های تصویر بدون تغییر می‌باشند که در این زمان استاد در حال توضیح بر روی تصویر موجود می‌باشد و مقادیر پیکسل‌ها دچار تغییر نمی‌گردند که می‌توان با استفاده از کدک شماره ۱ که در روش پیشنهادی توضیح داده شد این فریم‌ها را با جایگزینی تک فریم فشرده سازی نمود. در این مرحله نیز تغییرات پیکسل در فریم بسیار ناچیز بوده و با جایگزینی تک فریم هیچگونه تاثیری در کیفیت تصویر ویدیویی نخواهیم داشت.

می‌توان نتیجه گرفت به علت وجود چنین فریم‌هایی در تصاویر ویدیویی آموزش از راه دور با استفاده از روش فشرده سازی پیشنهادی در این دسته از ویدیوها نرخ فشرده سازی بالایی را خواهیم داشت بدون آنکه به کیفیت تصویر لطمه‌ای وارد گردد.

در فریم‌هایی که تغییرات مهمی صورت می‌گیرد نیز کاهش بجای آنکه در کل فریم اعمال گردد تنها در قسمتی از تصویر که از دید بیننده مورد توجه کمتری قرار دارد صورت می‌پذیرد که بر کیفیت روش

فشرده سازی پیشنهادی می افزاید. در روش فشرده سازی پیشنهادی کاهش کیفیت هوشمندانه تر از روش های دیگر فشرده سازی برای تصاویر ویدیویی آموزش از راه دور صورت می پذیرد.

همانگونه که در فصل پیش مشاهده شد نرخ فشرده سازی روش پیشنهادی نسبت به روش های متداول فشرده سازی بر روی تصاویر ویدیویی آموزش از راه دور از نرخ بسیار مطلوب تری برخوردار می باشد. این نرخ فشرده سازی بالاتر بدون آن است که افت کیفیت از دید ناظر قابل تشخیص باشد. این روش با کاهش چشمگیر حجم یک تصویر ویدیویی آموزش از راه دور پهنای بالاتری را در ارسال این دسته از ویدیوها به کاربران می دهد و می توانند تصاویر را با سرعت بالاتر و افت کیفیت کمتری در بستر وب انتقال دهند.

## ۵-۲ کارهای آتی

روش پیشنهادی در این پایان نامه بر پایه‌ی انتقال تصاویر ویدیوی آموزشی بر بستر وب بصورت غیر برخط می باشد. این انتقال با حجم کمتر می تواند به انتقال این تصاویر با سرعت بالاتر و با استفاده از پهنای باند کمتر صورت پذیرد، اما درصد بیشتری از تصاویر ویدیویی در بستر وب بصورت برخط بوده و همزمانی مابین ارسال و دریافت تصاویر توسط استاد و دانشجو وجود دارد. می توان به عنوان پیشنهاد برای ادامه ی این مسیر انتقال برخط تصاویر ویدیویی آموزش از راه دور بر بستر وب را بیان نمود. این انتقال به علت بر خط بودن نمی تواند از محتوای تغییرات فریم ها بیش از چند فریم آگاهی داشته باشد و تاخیر در اینگونه انتقالات نمی تواند بیش از چند فریم باشد. می توان در روش برخط از تکنیک‌هایی مانند پیش بینی و تصحیح خطا برای فشرده سازی تصاویر ویدیویی کم تحرک برای دستیابی به نرخ فشرده سازی بالاتری استفاده نمود.



همچنین می‌توان بیان کرد در روش فشرده سازی پیشنهادی از سه کدک استفاده شده است، در صورتیکه می‌توان با استفاده از ویژگیهای بیشتری از تصویر بجز میانگین روشنایی فریم‌ها، تعداد کدک‌های مورد استفاده در روش فشرده سازی را بیشتر نمود تا نرخ فشرده سازی تصاویر ویدیویی آموزش از راه دور بالاتر رود.

- [1] G. Jiang, "Distance Learning Technologies and an Interactive Multimedia Educational System," IEEE, 2001.
- [2] G. Varotto, "Low Delay Video Coding," 2006.
- [3] Rafael C. Gonzalez, Richard Eugene Woods, Digital Image Processing, Pearson/Prentice Hall, 2008.
- [4] Studio encoding parameters of digital television for stan-, ITU-R Recommendation BT.601-5.
- [5] Tiina Jarske and Yrj Neuvo, "Adaptive DPCM with median type predictors," IEEE Transactions on Consumer Electronics, vol. Vol. 37, no. No. 3, August 1991.
- [6] R. J. Clarke, "Image and Video Compression: A Survey," John Wiley & Sons, Inc., vol. Vol. 10, p. 20–32, 1999.
- [7] S. Acharjee, "A New Fast Motion Vector Estimation Algorithm for Video Compression," in IEEE/OSA/IAPR International Conference on Infonnatics, Electronics & Vision, 2012.
- [8] J. Fabrizio, "Motion compensationbasedontangentdistancepredictionfor video compression," Signal Processing: Image Communication, vol. 27, p. 153–171, 2012.
- [9] K. Jack, Video Demysti ed. A Hand book for the Digital Engineer, VA: LLH Technology Publishing, 2001.
- [10] I. E. Richardson, H.264 and MPEG-4 Video Compression, Video Coding for Next generation Multimedia, John Willey & Sons Ltd, 2003.
- [11] D. A. Huffman, "A method for the construction of minimum redundancy codes," Proc. of the IRE, vol. 40, pp. p 1098-1101, 1952.
- [12] J. Z. a. A. Lempel, "A Universal Algorithm for Sequential compression," IEEE Transaction on Information Theory, vol. 23, no. 3, pp. 337-343, 1977.
- [13] T. A. Welch, "A Technique for High Performance Data Compression," IEEE Computer, vol. 17, no. 6, 1984.

- [14] I. 11172-2, "Information technology, Coding of moving pictures and associated audio for digital storage media at up to 1.5Mbit/s video," MPEG1, Geneva, Switzerland, 1994.
- [15] I. 13818-2, "Information technology, Generic coding of moving pictures and associated audio information: video," MPEG2, Geneva, Switzerland, 1996.
- [16] B. Furht, A Survey of Multimedia Compression Techniques and Standards, Florida: Academic Press, 1995.
- [17] Morteza Zahedi, Alireza P. Nouri, "Introduce a new low motion video database," International Journal of Computer Application, 2013.
- [18] موسسه فرهنگی هنری پردازش هوشمند علانم, فرمت های ویدیویی, فاتح. م. 1392.
- [19] I. E. Richardson, The H.264 Advanced Video Compression Standard, ISBN: 978-0-470-51692-8, 2010, John Wiley & Sons, Ltd.
- [20] Srinivasan, Sridhar, et al. "Windows Media Video 9: overview and applications." *Signal Processing: Image Communication* 19.9 (2004): 851-875.

## **Abstract**

Video compression is one of the most important areas of research and commercial domains in computer science. With the advent of video in a web context, due to limited communication bandwidth, the growing demand for video compression has emerged. One of the most important areas to transfer video contents over the web platform is in the field of distance learning. The video contents are either online or pre-recorded to be transferred on the web. Due to the limited bandwidth and high size of video frames, video compression is a necessary solution. By reducing the size of a video frame, it can be easily transferred over the web, and associate the remaining bandwidth to sound or other additional information. The challenge in the areas of video compression codecs is to provide higher quality frames with smaller size and the most compact possible.

In this thesis, the focus is towards to video compression for sedentary video frames in distance learning applications and a codec is proposed to compress and transfer with a higher rate.

In this codec, all the redundancies of the video frames are removed at the first place. Then, by the help of image segmentation in each frame, redundancy is eliminated due to the fact that some parts of the frame does not change in a frame sequence. This action increases the compression rate of the video without having to harm the image pixel value and reduces the quality of the produce.

It can be claimed that video compression rate for distance learning is increased against to other methods which are applicable to stirring video frames.

The main objective of this thesis is to provide a unique video compression for distance learning application. The proposed codec for video compression has a high compression rate and will not affect the quality of moving parts of the video frames.

**Key word:** video compression, distance learning, video stream, artificial intelligent, new codec



Shahrood University of Technology

**Faculty of information technology and computer**

**Speed up the streaming video for distance learning**

Alireza Pashamohammad Nouri

Supervisor:

Dr.Morteza Zahedi

**FOR THE DEGREE OF  
MASTER OF SCIENCE**

September 2014