

## بکارگیری روش ترکیبی SPIHT-DCT با استفاده از مقیاس پذیری مکانی و

### مکانی- زمانی برای کد کردن تصاویر ویدئویی HDTV

وحید رضا صیرفیان<sup>۱\*</sup> سیامک طالبی<sup>\*\*</sup>

\*کارشناس ارشد، دانشکده مهندسی برق، دانشگاه آزاد اسلامی واحد میمه، میمه، ایران

\*\*کارشناس ارشد، دانشکده مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان

تاریخ دریافت: ۱۳۹۱/۰۴/۰۱

تاریخ پذیرش: ۱۳۹۱/۱۰/۱۱

#### چکیده

در این مقاله، کد کننده ترکیبی با استفاده از دو ویژگی مقیاس پذیری مکانی و مقیاس پذیری مکانی-زمانی برای کد کردن ویدئو با درجه تفکیک بالا ارائه شده است. در روش ترکیبی، قاب های ویدئویی Intra و Inter به دو روش متفاوت کد می شوند. قاب های Intra با استفاده از الگوریتم SPIHT<sup>۱</sup> که مبتنی بر تبدیل موجک است کد می شوند. قاب های Inter به روش معمول استاندارد MPEG-2 و بر اساس تبدیل DCT کد می شوند. با کد کردن ویدئو با درجه تفکیک بالا به دو روش مقیاس پذیری مکانی و مقیاس پذیری مکانی-زمانی، ویدئو از طریق دو یا سه لایه ارسال می شود. داده های ارسالی از لایه ها، ویدئویی با وضوح و کیفیت متفاوت به کاربر عرضه می کنند. به این ترتیب کاربر می تواند بر اساس نیاز خود سرویس مناسب را انتخاب کند.

در مقیاس پذیری مکانی لایه پایه و لایه ارتقا ساختار کدینگ یکسانی دارند. ولی در مقیاس پذیری مکانی-زمانی لایه ارتقا دوم، به دلیل اینکه فقط شامل قاب های Inter است، فقط مبتنی بر روش استاندارد کد می شود.

نتایج شبیه سازی های انجام شده روی ویدئوهای مختلف با درجه تفکیک بالا، بهبود کیفیت تصویر نهایی در روش ترکیبی پیشنهادی با مقیاس پذیری را در لایه های مختلف، نسبت به روش مبتنی بر استاندارد MPEG-2 نشان می دهد.

**کلید واژگان:** استاندارد MPEG-2، الگوریتم SPIHT، تبدیل

DCT، مقیاس پذیری، ویدئو با درجه تفکیک بالا.

#### ۱. مقدمه

کد کننده های ویدئویی مقیاس پذیر<sup>۲</sup>، کاربردهای زیادی در شبکه های مختلف ارتباطی دارند. هر شبکه نیز بر اساس سطوح مختلف سرویسی که به کاربر عرضه می کند ارزش گذاری<sup>۳</sup> و با ارسال ویدئو در کیفیت های مختلف و یا دقت مکانی یا زمانی متفاوت، کاربر می تواند بر اساس نیاز و محدودیت های گیرنده خود، سرویس مورد نظر را انتخاب می کند<sup>۴</sup>. این مورد برای ارسال ویدئو با درجه تفکیک بالا<sup>۵</sup> اهمیت بیشتری دارد.

بررسی کد کردن ویدئو با درجه تفکیک بالا با مقیاس پذیری از این جنبه دارای اهمیت است که عرضه کنندگان محصولات ویدئویی و تصویری، همچنین سرویس دهندگان رسانه های تصویری، به این نوع ویدئو توجه ویژه ای پیدا کرده اند و رویکرد همه جانبه ای به آن داشته اند<sup>۲</sup>. بنابراین به دلیل حجم بالای اطلاعات ارسالی و لزوم انطباق فنی گیرنده ها برای پخش تصاویر با درجه تفکیک بالا، قابلیت ارسال و دریافت تصاویر ویدئویی با اندازه و نرخ متفاوت، می تواند منجر به ارتقا کیفیت سرویس دهی سیستم شود.

از زمان مطرح شدن ارسال ویدئو با درجه کیفیت بالا، ارسال اطلاعات این نوع ویدئو از طریق دو یا چند لایه مطرح شد<sup>۳</sup>. لایه ای که بتواند اطلاعات لازم برای تصویر ویدئویی استاندارد<sup>۴</sup> را انتقال دهد و لایه ای که با افزوده شدن اطلاعات آن به تصویر ویدئویی استاندارد، ویدئو با درجه تفکیک بالا ارائه کند<sup>۳و۴</sup>. بنابراین کد کننده ویدئویی مقیاس پذیر به عنوان یک نیاز مطرح بوده است.

<sup>1</sup> Set Partitioning Hierarchical Trees

<sup>2</sup> Scalable

<sup>3</sup> QoS

<sup>4</sup> HDTV

<sup>5</sup> SDTV

هم چنین بکاربردن روش ترکیبی SPIHT-DCT برای ویدئوهای با درجه تفکیک بالا، باعث ارتقا کیفیت نسبت به کدک استاندارد MPEG-2 شده است [۱۷].

در این مقاله، روش کدکردن ترکیبی SPIHT-DCT برای ویدئوهای با درجه تفکیک بالا، HDTV، ابتدا بر اساس مقیاس پذیری مکانی<sup>۱</sup> و سپس مقیاس پذیری مکانی-زمانی<sup>۲</sup> پیشنهاد شده است. اطلاعات ویدئو در حالت اول از طریق دو لایه و در حالت دوم از طریق سه لایه مجزا ارسال می شود. در کدک مقیاس پذیر مکانی، کد کردن هر دو لایه مشابه هم انجام می شود. کد کردن قاب های Intra مبتنی بر تبدیل موجک و کدکردن قاب های Inter بر اساس تبدیل DCT انجام شده است. کدکننده مقیاس پذیر مکانی-زمانی به غیر از دو لایه پایه و ارتقا اول، لایه سوم می دارد که کدکننده آن فقط تحت استاندارد MPEG-2 عمل می کند. دلیل آن وجود قاب های Inter در این لایه است.

هم چنین برای ارزیابی تاثیر کدکننده ترکیبی، ویدئو با درجه تفکیک بالا، با شرایط یکسان، تحت استاندارد MPEG-2 نیز کد شده است.

ادامه این مقاله به این صورت سازماندهی شده است که در بخش دوم الگوریتم کدکردن ترکیبی پیشنهادی تشریح شده است. در بخش سوم کدکننده با مقیاس پذیری مکانی و در بخش چهارم الگوریتم کد کردن ترکیبی با مقیاس پذیری مکانی-زمانی برای ویدئو با درجه تفکیک بالا توضیح داده شده است. بخش پنجم به نتایج شبیه سازی-های انجام شده به هر دو روش مقیاس پذیری اختصاص دارد. در نهایت مقاله در بخش ششم جمع بندی شده است.

## ۲. کدکردن تصاویر با درجه تفکیک بالا به روش

### ترکیبی

برای کد کردن قاب های تصویری با درجه تفکیک بالا، ترکیب الگوریتم SPIHT مبتنی بر تبدیل موجک و تبدیل DCT به کار می رود. در روش ترکیبی SPIHT-DCT، کدکننده به دو بخش مجزا تقسیم می شود. قسمت اول کدکننده قاب های Intra است. این قسمت از کدکننده به طور کاملاً مجزا از روش مبتنی بر استاندارد MPEG-2 عمل می کند. این قاب ها با الگوریتم فشرده سازی SPIHT که مبتنی بر تبدیل موجک است، کد می شوند.

قاب های Inter که شامل قاب های P و B هستند، در بخش دیگر کدکننده بر اساس تبدیل DCT کد می شوند. قاب های Inter با

تصویر ویدئو با درجه تفکیک بالا از عرض و طول دارای وضوح دو برابر نسبت به تصویر ویدئو استاندارد است. اندازه قاب های ویدئوی HDTV ۱۲۸۰×۷۲۰ و ۱۹۲۰×۱۰۸۰ است. نسبت تصویر در ویدئو با درجه تفکیک بالا ۱۶:۹ است. این نسبت در تصویر ویدئو استاندارد برابر ۴:۳ است [۵].

بسیاری از الگوریتم های مورد استفاده در کدکردن ویدئو بر اساس تبدیل DCT پایه گذاری شده است [۶]. در استاندارد MPEG-2 نیز برای حذف افزونگی های مکانی موجود در تصویر از تبدیل DCT استفاده می شود [۷]. از سوی دیگر نتایج آزمایش های سال های اخیر نشان داده است که تبدیل موجک در فشرده سازی تصویر می تواند بر مشکلاتی که الگوریتم های مبتنی بر تبدیل بلوکی ایجاد می کنند، غلبه کند [۶]. هم چنین بکار بردن تبدیل موجک در کدکردن ویدئو می تواند نتایج کیفی بهتری نسبت به کدک های متداول مبتنی بر تبدیل DCT داشته باشد [۸ و ۹].

تبدیل موجک برای کدکردن تصویر با درجه تفکیک بالا نیز مورد استفاده قرار گرفته است [۶]. در روش های مختلف فشرده سازی، با افزایش تعداد پیکسل های یک بلوک تحت پردازش، از همبستگی بیشتر بین پیکسل ها استفاده می شود و فشرده سازی موثرتری انجام می شود [۱۰]. ولی در روش های مبتنی بر تبدیل DCT، بزرگ کردن بلوک تحت تبدیل، موجب پیچیدگی بیشتر می شود [۱۱]. تبدیل موجک از این خاصیت استفاده می کند و با اعمال تبدیل به کل تصویر یا قسمت های بزرگتر از تصویر، نتایج بهتری ارائه می کند. به همین دلیل در استاندارد JPEG2000 نیز از تبدیل موجک استفاده شده است [۱۱].

استفاده از مقیاس پذیری های کدک استاندارد، در روش های کدکردن مبتنی بر تبدیل موجک نیز بکار گرفته شده است [۱۲ و ۱۳]. استفاده از تبدیل موجک در کدکردن ویدئویی به اندازه وضوح CIF با مقیاس پذیری نسبت به استفاده از تبدیل DCT، منجر به ارتقا کیفیت در حدود ۴ dB شد [۱۲].

الگوریتم SPIHT روشی برای فشرده سازی تصویر بر اساس تبدیل موجک است [۱۴ و ۱۵]. با استفاده از این الگوریتم، فقط داده های مهم تصویر و ضرایب حساس تبدیل موجک انتقال پیدا می کند [۱۱]. مزایای استفاده از تبدیل موجک می تواند برای کد کردن قاب های Intra و تبدیل DCT برای قاب های Inter مورد استفاده قرار گیرد. این روش برای استاندارد ویدئویی H.263 نتایج قابل قبولی به همراه داشت [۱۶]. در این مرجع شبیه سازی ها در نرخ بیت های پایین حدود ۳۰۰ تا ۱۵۰۰ کیلو بیت در ثانیه و روی ویدئوهایی با اندازه CIF و QCIF، انجام شد.

<sup>1</sup> Spatial Scalability

<sup>2</sup> Spatial-Temporal Scalability

می دهد. با افزوده شدن داده های لایه ارتقا به داده های لایه پایه، ویدئوی با درجه تفکیک بالا در گیرنده دریافت می شود [۱۶].

کد کننده ترکیبی مقیاس پذیر مکانی شامل دو قسمت مجزا از هم است. ساختار کد کننده هر قسمت مشابه هم است.

در شکل (۱) ساختار بلوک دیاگرام کد کننده ترکیبی پیشنهادی با استفاده از ویژگی مقیاس پذیری مکانی برای کد کردن ویدئو با درجه تفکیک بالا نشان داده شده است.

در کد کننده لایه پایه، ابتدا وضوح قاب های تصویر کاهش پیدا می کند. قاب های Intra توسط قسمت کد کننده مبتنی بر الگوریتم SPIHT کد می شوند. این قسمت مجزا از کد کننده ی قاب های Inter است و مبتنی بر تبدیل DCT است. قاب های بازسازی شده I برای پیش بینی و جبران حرکت قاب های P و B بکار می روند.

قاب های کد شده توسط لایه پایه، پس از دکد شدن از طریق مسیر بازگشت در فرستنده به وضوح اصلی تصویر با درجه تفکیک بالا برگردانده می شوند. این قاب ها به همین صورت در گیرنده قابل بازسازی هستند. داده های لایه ارتقا، حاصل تفاضل قاب اصلی تصویر و داده های قابل بازسازی در گیرنده است. در واقع در این مرحله پس از حذف داده های ارسال شده، اطلاعات باقی مانده هر قاب به روش ترکیبی کد و ارسال می شوند.

هر کد کننده سری بیت های مجزایی ایجاد می کند. برای تصویر با وضوح پایین تر، سری بیت های ارسال لایه پایه مورد استفاده قرار می گیرد. سری ویدئویی حاصل از دکد کردن بیت های این لایه، ویدئویی با وضوح SDTV (۷۲۰×۴۸۰) به کاربر ارائه می دهد. سری دیگر که از طریق لایه ارتقا ارسال می شود، شامل اطلاعاتی است که با افزوده شدن به اطلاعات لایه پایه، ویدئوی HDTV (۷۲۰×۱۲۸۰ یا ۱۹۲۰×۱۰۸۰) با وضوح کامل را عرضه می کند. با استفاده از این ویژگی در گیرنده، یک ویدئو با دو وضوح متفاوت وجود دارد که کاربر با توجه به نیاز خود، سرویس دلخواه را انتخاب می کند.

تخمین و جبران سازی حرکت در هر لایه به طور مستقل از لایه دیگر محاسبه و ارسال می شود. به دلیل اختلاف وضوح تصویر در دو قسمت کد کننده، نمی توان از نتایج حاصل از پیاده سازی الگوریتم های پیش بینی و جبران حرکت و الگوریتم های تخمین بردار حرکت (Motion Vector) یک لایه برای لایه دیگر استفاده کرد.

بکاربردن الگوریتم های پیش بینی و جبران حرکت روی قاب مرجع ایجاد می شوند. پیاده سازی الگوریتم های پیش بینی و جبران حرکت روی قاب های تصویری منجر به ایجاد فرکانس های بالای مصنوعی می شود. اگر این فرکانس های بالا تحت تبدیل الگوریتم های مبتنی بر تبدیل موجک قرار گیرد، می تواند منجر به ایجاد اعوجاج شود. اعوجاج ایجاد شده در این قسمت مزایای استفاده از تبدیل موجک در کد کردن ویدئو را تحت تاثیر قرار می دهد [۱۶].

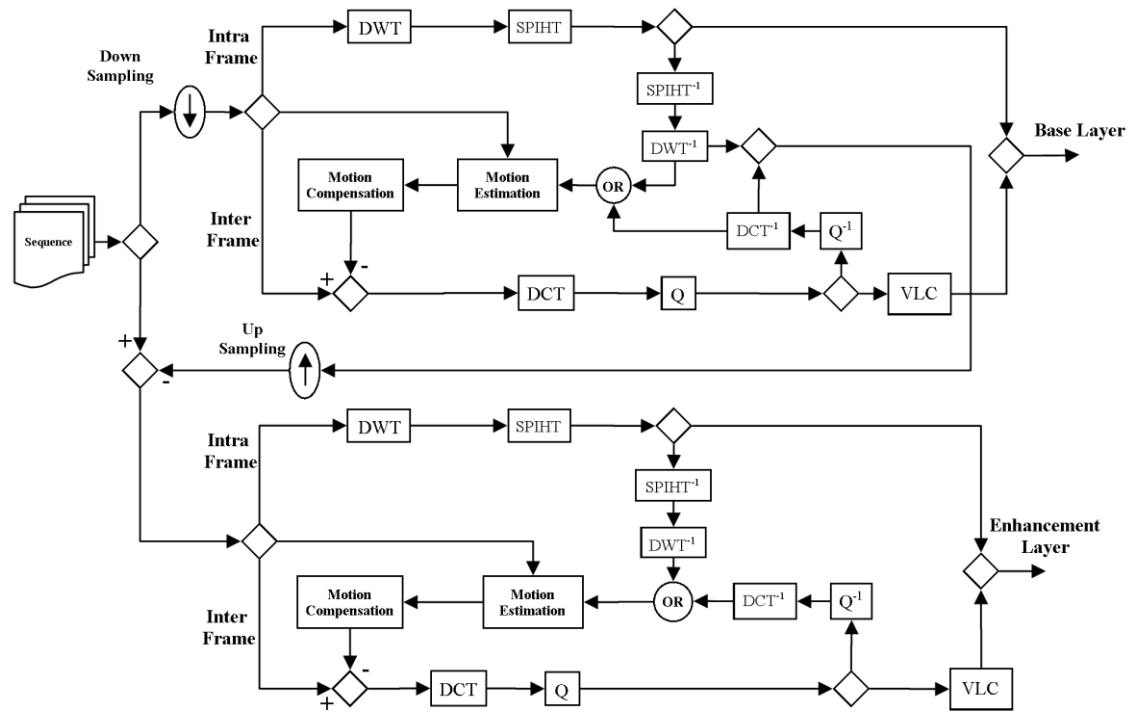
مرجع قاب های P و B، قاب I کد شده به روش مبتنی بر تبدیل موجک است. در کد کردن قاب های I با افزایش لایه های تجزیه در الگوریتم استفاده شده از همبستگی بین همه نمونه های تصویر استفاده می شود. ارتقا کیفیت قاب مرجع، منجر به بهبود کیفیت قاب های Inter می شود و می تواند در کیفیت کل ویدئوی ارسالی موثر واقع شود.

الگوریتم پیش بینی حرکت، یک مدل از قاب جاری را بر اساس داده های موجود در قاب های قبلی و بعدی کد شده ایجاد می کند. هدف از بکارگیری الگوریتم پیش بینی حرکت، مدل کردن قاب جاری با حداکثر دقت ممکن است تا اطلاعات ارسالی کاهش یابد. با کم کردن قاب مدل شده از قاب جاری، عمل جبران حرکت روی قاب انجام می شود [۱۱]. پساز جبران حرکت، قابی بدست می آید که حاوی اطلاعاتی است که دکد کننده برای بازسازی قاب جاری از روی قاب مرجع نیاز دارد. بهترین عملکرد فشرده سازی زمانی هنگامی بدست می آید که داده های قاب باقی مانده از الگوریتم جبران حرکت به کمترین مقدار برسد، در حالی که در دکد کننده بهترین کیفیت ممکن ایجاد شود. الگوی قرارگیری و کد شدن قاب ها را کد کننده مشخص می کند، در عین حال الگوی متداول ترتیب قاب ها به این صورت است:  $I B B P B B P B B$  [۵].

### ۳. روش کد کردن ترکیبی SPIHT-DCT با مقیاس

#### پذیری مکانی برای ویدئو با درجه تفکیک بالا

برای کد کردن ترکیبی ویدئوهای با درجه تفکیک بالا، مقیاس پذیری مکانی به کد کننده افزوده شده است. با استفاده از این ویژگی، دو لایه برای انتقال داده ها ایجاد می شود. لایه اول یا لایه پایه، داده های ویدئو با وضوح پایین تر را انتقال می دهد. لایه دوم داده های ارسال نشده برای ارتقا ویدئو به وضوح اصلی تصویر لازم استرا انتقال



شکل ۱: بلوک دیاگرام کد کننده ترکیبی با استفاده از مقیاس پذیری مکانی.

وضوح کامل را شامل می شود که از طریق لایه پایه ارسال نشده اند. در نتیجه مجموع اطلاعات سه لایه، ویدئو با درجه تفکیک بالا با وضوح و نرخ قاب در ثانیه ویدئوی اصلی است.

بلوک دیاگرام کد کننده ترکیبی پیشنهادی ویدئو با درجه تفکیک بالا با ویژگی مقیاس پذیری مکانی- زمانی در شکل (۲) نشان داده شده است.

در این کد کننده سه واحد جداگانه برای اجرای الگوریتم های پیش بینی و جبران حرکت وجود دارد. این الگوریتم ها برای قاب های با وضوح پایین تر به طور جداگانه محاسبه می شوند. بردارهای حرکتی این قاب ها هم مجزا از قسمت های دیگر محاسبه و ارسال می شوند. در لایه ارتقا اول نیز به دلیل افزایش وضوح تصویر باید الگوریتم پیش بینی و جبران حرکت جداگانه محاسبه شود. در لایه ارتقا دوم، قاب هایی ارسال می شوند که در لایه های قبلی کد و ارسال نشده اند. بنابراین این الگوریتم ها برای این قاب ها به طور مجزا انجام می شود.

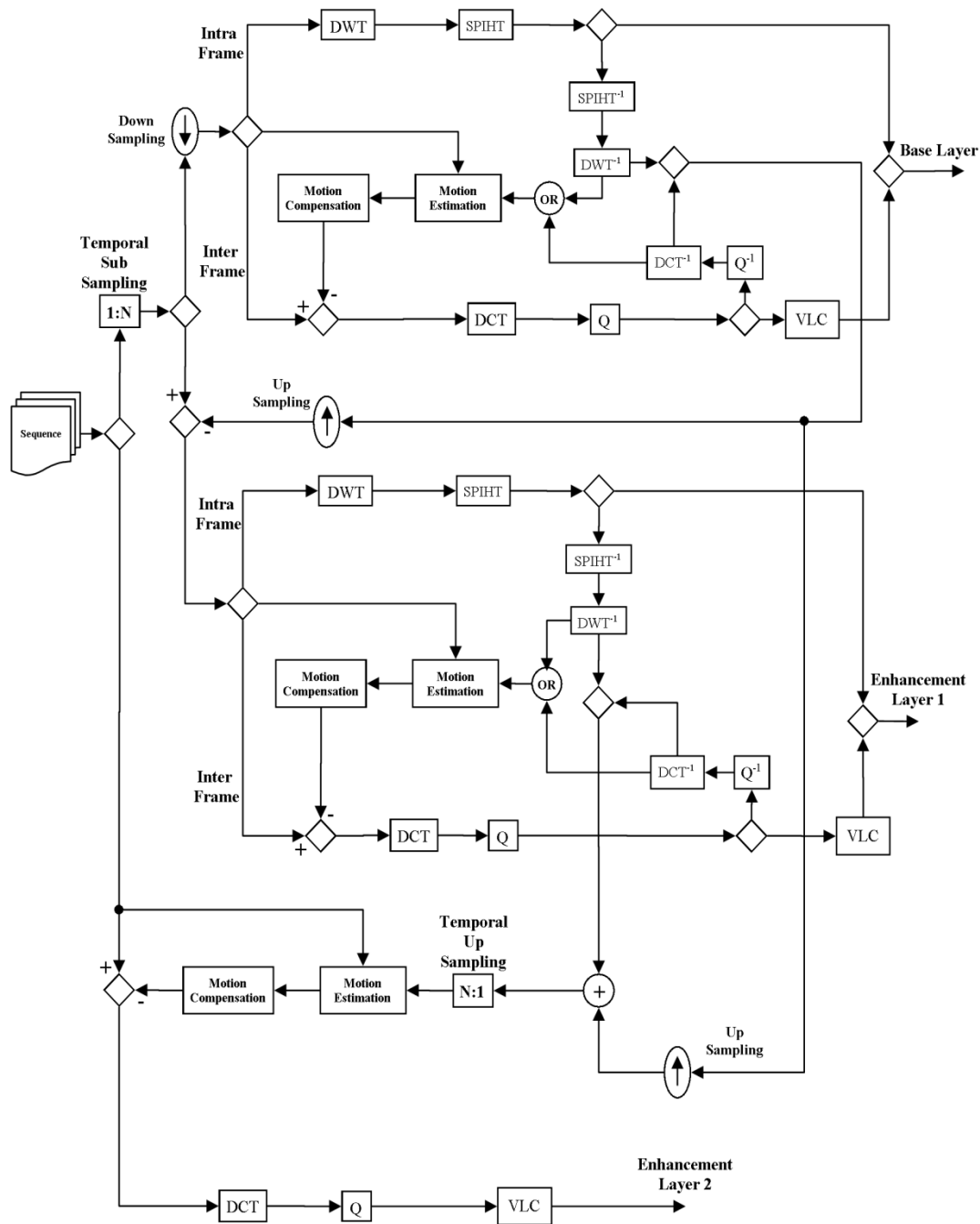
این کد کننده ترکیبی، ضمن ارتقا کیفیت قاب های مرجع که می تواند منجر به افزایش کیفیت سایر قاب های ویدئو شود، از مقیاس پذیری مکانی- زمانی نیز استفاده می کند. بنابراین کاربر می تواند رنج گسترده تری از سرویس ویدئویی در اختیار داشته باشد و بر اساس نیاز، سرویس مورد نظر خود را انتخاب نماید.

#### ۴. روش کد کردن ترکیبی SPIHT-DCT با مقیاس پذیری مکانی- زمانی برای ویدئو با درجه تفکیک بالا

در این قسمت به کد کننده ترکیبی پیشنهادی، ویژگی مقیاس پذیری مکانی- زمانی اضافه شده است. در مقیاس پذیری مکانی، قاب های تصویر ویدئو با اندازه وضوح متفاوت ارسال می شوند. حال آنکه در مقیاس پذیری زمانی نرخ قاب در ثانیه لایه های ارسالی نیز تغییر می کند. در واقع یک قاب ویدئویی با درجه تفکیک بالا در این کد کننده از سه لایه مختلف ارسال می شود.

لایه اول یا لایه پایه، اطلاعات پایین ترین سطح وضوح و نرخ قاب در ثانیه ویدئو را انتقال می دهد. داده های انتقالی توسط این لایه، ویدئو SDTV با اندازه قاب  $720 \times 480$  و نرخ ۲۵ قاب در ثانیه عرضه می کنند. در این لایه قاب های I، P، و برخی قاب های B قرار دارند. اطلاعات لایه دوم یا لایه ارتقا اول، با افزوده شدن به داده های لایه پایه، وضوح قاب های ارسالی لایه پایه را به اندازه واقعی می رسانند. در واقع ویدئو عرضه شده با مجموع اطلاعات لایه پایه و لایه ارتقا اول، نسبت به ویدئو با درجه تفکیک بالا، فقط نرخ قاب در ثانیه کمتری دارد.

اطلاعات انتقال داده شده توسط لایه سوم یا لایه ارتقا دوم، به اطلاعات دو لایه پایین تر افزوده می شوند. این لایه قاب هایی با



شکل ۲: بلوک دیاگرام کدکننده ترکیبی با استفاده از ویژگی مقیاس پذیری مکانی-زمانی.

## ۵. شبیه سازی ها و نتایج

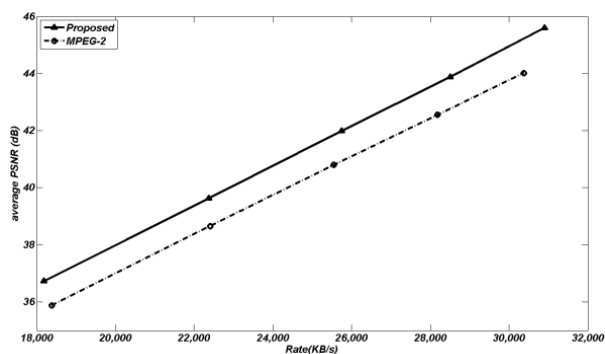
### ۵.۱. کدکننده ترکیبی با مقیاس پذیری مکانی

برای کدکردن ویدئو با درجه تفکیک بالا به روش ترکیبی پیشنهادی و با ویژگی مقیاس پذیری مکانی، از ویدئوهای با وضوح متفاوت استفاده شده است. اندازه وضوح ویدئوهای تحت شبیه سازی ۱۲۸۰×۷۲۰ و ۱۹۲۰×۱۰۸۰ است. شبیه سازی ها روی مولفه روشنایی

قاب های ویدئویی انجام شده است. هم چنین نرخ قاب های Intra در دو حالت متفاوت ۱ به ۵ و ۱ به ۹ در نظر گرفته شده است. در نرخ Intra ۱ به ۹ از همان الگوی قاب بیان شده استفاده شده است و در نرخ Intra ۱ به ۵ همان الگو با قطع پنجمین قاب به صورت I B B P B بکار رفته است. شکل های (۳ و ۴) نشان دهنده میانگین کیفیت ویدئو Shields به وضوح ۱۲۸۰×۷۲۰ در نرخ بیت های مختلف است. اختلاف بین میانگین کیفیت ویدئو کد شده با افزایش نرخ بیت تقریباً ثابت است و

مختلف را نشان می دهد. هم چنین نرخ بیت انتقال داده شده توسط لایه پایه نیز مشخص شده است. اندازه قاب های انتقال داده شده توسط لایه پایه  $720 \times 480$  است. لایه پایه حدود ۳۰ درصد مجموع بیت ارسالی را به خود اختصاص داده است.

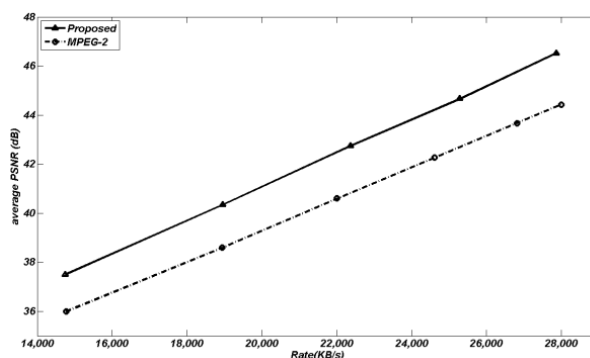
در شکل های (۵ و ۶) نمودار میانگین کیفیت ویدئو Mobcal به اندازه  $1280 \times 720$  بر حسب نرخ بیت های مختلف نشان داده شده است. در شکل ۵ با نرخ ایترا ۱ به ۵، افزایش نرخ بیت، اختلاف بین میانگین کیفیت روش پیشنهادی و روش استاندارد را افزایش داده است. با افزایش نرخ بیت، روند صعودی میانگین کیفیت ویدئو کد شده به روش استاندارد کاهش می یابد. برای ویدئو کد شده به روش ترکیبی پیشنهادی، روند افزایش کیفیت تقریباً ثابت است. شکل (۶) نشان می دهد که با نرخ ایترا ۱ به ۹، اختلاف ویدئو کد شده به روش پیشنهادی و استاندارد با افزایش نرخ بیت کمی افزایش می یابد و در نرخ بیت  $29900 \text{ KB/s}$  حدود  $1/4 \text{ dB}$  است.



شکل ۴: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

روند افزایش تقریباً مشابهی را دارد. در نرخ ایترا ۱ به ۵ بهبود میانگین کیفیت روش پیشنهادی نسبت به روش استاندارد بیش از مورد مشابه در نرخ ایترا ۱ به ۹ است. علت کاهش میانگین کیفیت ویدئو با کاهش نرخ قاب های ایترا، کاهش اثر قاب مرجع کد شده به روش مبتنی بر تبدیل موجک در سایر قاب ها است. هم چنین با توجه به نتایج حاصل از شبیه سازی روش پیشنهادی بدون در نظر گرفتن مقیاس پذیری، می توان این نکته را بیان کرد که روش ترکیبی پیشنهادی به همراه ویژگی مقیاس پذیری مکانی بهتر از روش استاندارد به همراه این ویژگی عمل می کند و با دو لایه شدن اطلاعات ارسالی عملکرد بهتری نسبت به روش استاندارد دارد [۱۷]. در نرخ ایترا ۱ به ۹ نتایج حاصل از روش مبتنی بر تبدیل موجک در نرخ بیت های بالا، نزدیک به نتایج روش ترکیبی است. در عین حال با نزدیک شدن به نرخ بیت  $30000 \text{ KB/s}$  میانگین کیفیت روش مبتنی بر تبدیل موجک اندکی کاهش می یابد.

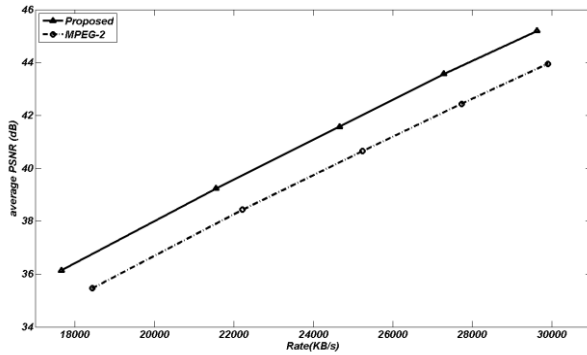
جدول (۱) نتایج شبیه سازی انجام شده روی ویدئو Shields به اندازه  $1280 \times 720$  را در نرخ بیت های تقریباً یکسان برای دو لایه



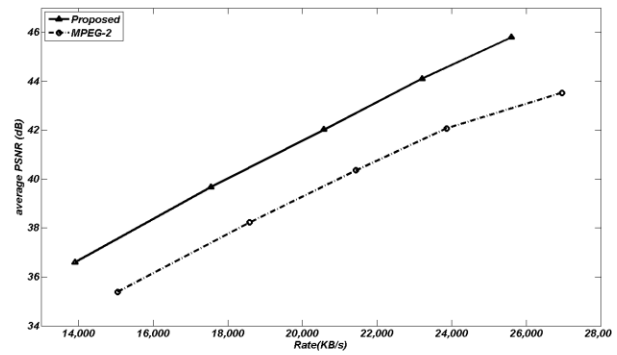
شکل ۳: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

جدول ۱: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه  $1280 \times 720$  برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

Sequence	Size	frame/sec	Intra Coding Rate	Compression Method	Bit Rate (KB/s)		Avg. PSNR (dB)
					SDTV	HDTV	HDTV
Shields_ter	1280*720	50	1/5	MPEG-2	5,848	18,937	38.59
				Proposed	5,469	18,947	40.34
			1/9	MPEG-2	8,627	25,541	40.8
				Proposed	8,425	25,751	41.99



شکل ۶: مقایسه میانگین کیفیت لایه های ویدئو Mobcal به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.



شکل ۵: مقایسه میانگین کیفیت لایه های ویدئو Mobcal به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

جدول ۲: مقایسه نتایج شبیه سازی ویدئو Mobcal به اندازه  $1280 \times 720$

برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی.

Sequence	Size	frame/sec	Intra Coding Rate	Compression Method	Bit Rate (KB/s)		Avg. PSNR (dB)
					SDTV	HDTV	HDTV
Mobcal_ter	1280*720	59.94	1/5	MPEG-2	7,466	23,862	42.07
				Proposed	6,883	23,207	44.1
			1/9	MPEG-2	10,419	29,905	43.93
				Proposed	10,060	29,625	45.19

با افزایش نرخ بیت، اختلاف میانگین کیفیت ابتدا کمی کاهش و سپس اندکی افزایش می یابد.

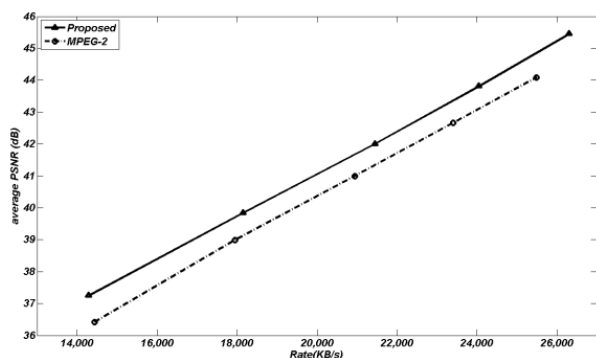
شکل (۸) در نرخ ایترا ۱ به ۹، نشان می دهد که اختلاف میانگین کیفیت دو روش کد کردن نسبت به نرخ ایترا بالاتر، کاهش یافته که به دلیل میرا بودن اثر بهبود کیفیت قاب کد شده به روش تبدیل موجک در قاب های بعدی است. در نرخ بیت برابر، میانگین کیفیت ویدئو کد شده در نرخ ایترا ۱ به ۵ حدود  $1/6$  dB بیشتر از ویدئو کد شده در نرخ ۱ به ۹ است.

جدول (۳) نتایج میانگین کیفیت ویدئو Shields به اندازه  $1920 \times 1080$  در نرخ بیت تقریباً یکسان نشان می دهد. حجم اطلاعات ارسالی توسط لایه پایه هم در هر نرخ بیت مشخص شده است. در نرخ ایترا ۱ به ۵ و در نرخ بیت حدود  $20000$  KB/s اختلاف میانگین کیفیت دو روش کد کردن به حدود  $1/5$  dB می رسد. در نرخ ایترا ۱ به ۹ و در نرخ بیت حدود  $21000$  KB/s اختلاف میانگین کیفیت دو روش به حدود  $1$  dB می رسد. لایه پایه حدود ۲۹ تا ۳۱ درصد حجم اطلاعات ارسالی را تشکیل می دهد.

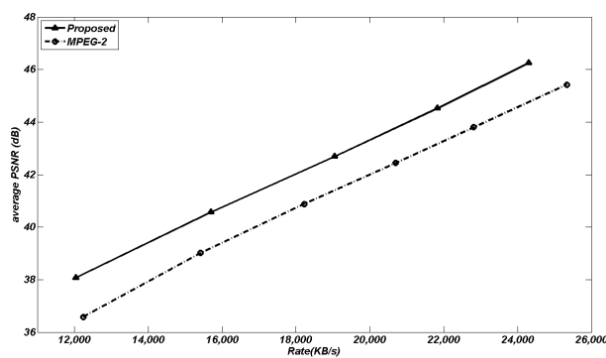
در جدول (۲) مقایسه میانگین کیفیت ویدئو Mobcal به اندازه  $1280 \times 720$  کد شده به سه روش پیشنهادی و تبدیل موجک و استاندارد در نرخ بیت تقریباً یکسان نشان داده شده است. حجم اطلاعات لایه پایه به اندازه وضوح  $720 \times 480$  در مقایسه با حجم بیت ارسالی برای ویدئو با درجه تفکیک بالا نیز نشان داده شده است. اطلاعات ارسالی لایه پایه حدود ۳۰ درصد از حجم ارسالی برای انتقال ویدئو کامل را تشکیل می دهد.

شکل های (۷ و ۸) نمودارهای مقایسه میانگین کیفیت برای ویدئو Shields به اندازه  $1920 \times 1080$  را نشان می دهد. این ویدئو به دو روش ترکیبی و استاندارد و با ویژگی مقیاس پذیری مکانی در دو نرخ بیت متفاوت کد شده است.

شکل (۷) در نرخ ایترا ۱ به ۵ است و شبیه سازی از  $12000$  KB/s تا حدود  $26000$  KB/s انجام شده است. اختلاف میانگین کیفیت بین دو روش کد کردن ترکیبی پیشنهادی و روش استاندارد به  $1/5$  dB هم می رسد. نتایج حاصل شده نشان می دهد که



شکل ۸: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1920 \times 1080$  در نرخ ایترا ۱ به ۹ کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.



شکل ۷: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1920 \times 1080$  در نرخ ایترا ۱ به ۵ کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.

جدول ۳: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه  $1920 \times 1080$  برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش پیشنهادی و استاندارد با مقیاس پذیری مکانی.

Size	frame/sec	Intra Coding Rate	Compression Method	Bit Rate (KB/s)		Avg. PSNR (dB)
				SDTV	HDTV	HDTV
1920*1080	25	1/5	MPEG-2	5,949	20,696	42.44
			Proposed	5,333	19,051	42.69
		1/9	MPEG-2	7,012	20,941	41
			Proposed	7,052	21,443	42

جدول (۴)، میانگین کیفیت ویدئو را در یک نرخ بیت در ثانیه تقریباً یکسان و با نرخ Intra متفاوت نشان می دهد.

شکل های ۱۱ و ۱۲ مقایسه میانگین کیفیت ویدئوی Mobcal به اندازه وضوح  $1280 \times 720$  با نرخ ۵۰ قاب در ثانیه است. مقایسه در دو نرخ Intra متفاوت ۱ به ۵ و ۱ به ۹ انجام شده است. شبیه سازی کد کردن این ویدئو نیز روی ۵۰ قاب انجام شده است.

با مشاهده نتایج این ویدئو، با افزایش نرخ بیت در ثانیه میانگین کیفیت لایه های ویدئو کد شده به روش ترکیب نسبت به ویدئو کد شده به روش استاندارد افزایش پیدا می کند. با کاهش نرخ قاب Intra به نرخ ۱ به ۹، اختلاف کیفیت دو روش کاهش می یابد. این نتایج نشان می دهد با افزایش قاب های Inter وابسته، اثر ارتقا قاب مرجع در آنها میرا می شود. جدول (۵)، نتایج شبیه سازی ویدئو به هر دو روش ترکیبی و استاندارد را در نرخ بیت ارسالی تقریباً یکسان نشان می دهد. نتایج در دو نرخ متفاوت Intra نشان داده شده است. لایه پایه در نرخ Intra ۱ به ۵ حدود ۲۰ درصد حجم اطلاعات ارسالی را تشکیل می دهد. در نرخ ۱ به ۹ این مقدار به ۱۹ درصد می رسد.

اختلاف میانگین کیفیت لایه های ویدئو Shields با اندازه وضوح  $1920 \times 1080$  با نرخ ۲۵ قاب در ثانیه در دو نرخ Intra متفاوت در شکل های ۱۳ و ۱۴ نشان داده شده است. اختلاف بین ویدئو کامل و

## ۵.۲. کد کننده ترکیبی با مقیاس پذیری مکانی-زمانی

عملکرد کد کننده ترکیبی با مقیاس پذیری مکانی-زمانی با کد کننده استاندارد MPEG-2 مقایسه شده است. ویدئوهای با درجه تفکیک بالا با وضوح متفاوتی برای شبیه سازی ها در نظر گرفته شده است.

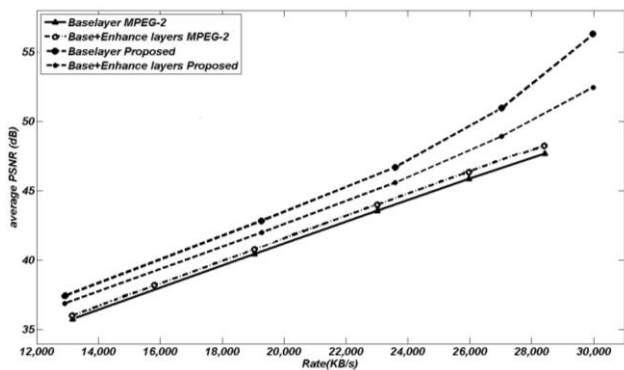
شکل های ۹ و ۱۰، میانگین کیفیت ویدئوی Shields با اندازه وضوح  $1280 \times 720$  را نسبت به نرخ های بیت بر ثانیه مختلف در دو نرخ Intra ۱ به ۵ و ۱ به ۹ نشان می دهند. در این شکل ها میانگین کیفیت ویدئو کد شده به دو روش متفاوت با هم مقایسه شده است. شبیه سازی روی این ویدئو برای ۵۰ قاب انجام شده است.

ترکیب لایه پایه و لایه ارتقا اول کد شده به روش ترکیبی پیشنهادی بیشترین کیفیت را دارد. ویدئو این دو لایه با نرخ ۲۵ قاب در ثانیه است. در واقع نصف نرخ قاب در ثانیه ویدئو اصلی است. با افزایش نرخ بیت، روند افزایش کیفیت سریع تر می شود. نرخ بیت ارسالی برای لایه پایه و لایه ارتقا اول همان نرخ بیت ارسالی کل ویدئو است که ترکیب این دو لایه بخشی از آن را شامل می شود. در مجموع اختلاف میانگین کیفیت ویدئو کامل کد شده به روش ترکیبی نسبت به روش استاندارد، با افزایش نرخ بیت در ثانیه زیاد می شود.

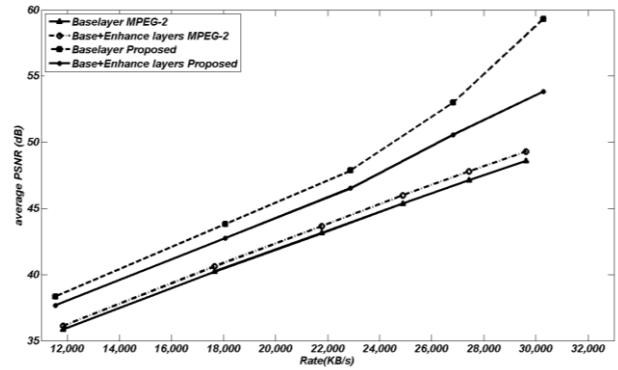


و در نرخ **Intra** متفاوت نشان داده شده است. از این جدول می توان درصد اشغال شده پهنای باند توسط لایه های پایه و ارتقا اول را بدست آورد. ویدئو لایه پایه در هر دو نرخ **Intra** حدود ۹ درصد پهنای باند را اشغال می کند. میانگین کیفیت ویدئو با درجه تفکیک بالا و ویدئو حاصل از لایه های پایه و ارتقا اول، در نرخ بیت ارسال تقریباً یکسان نیز، مشخص است. لایه پایه برای این ویدئو همان اندازه وضوح **SDTV** یا  $720 \times 480$  را ارسال می کند. هم چنین ویدئو حاصل از لایه پایه و ترکیب لایه پایه و لایه ارتقا اول نرخ ۱۳ قاب در ثانیه را دارد.

ویدئو حاصل از مجموع لایه پایه و لایه ارتقا در روش ترکیبی بیشتر از اختلاف لایه های مشابه در روش استاندارد است. با افزایش نرخ بیت ارسال، این اختلاف در روش ترکیبی افزایش می یابد، در صورتی که در روش استاندارد، این اختلاف تقریباً ثابت است. نتایج نشان می دهد که با افزایش اندازه وضوح، در نرخ بیت ارسال یکسان، میانگین کیفیت ویدئو نیز افزایش پیدا می کند. در نرخ **Intra** ۱ به ۹ اختلاف میانگین کیفیت ویدئو کد شده به روش استاندارد و ترکیبی نسبت به نرخ **Intra** ۱ به ۵ کاهش یافته است. جدول (۶) نتایج شبیه سازی ویدئو را در نرخ بیت های تقریباً یکسان



شکل ۱۰: میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

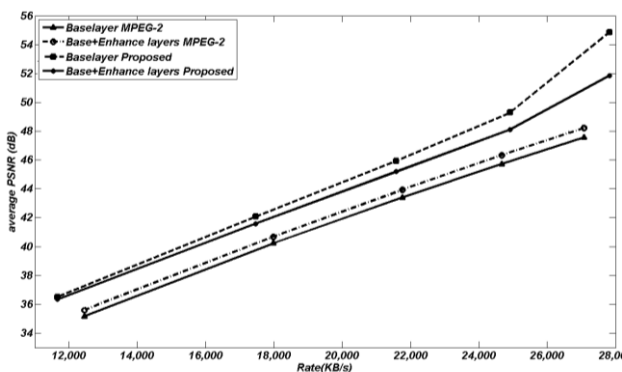


شکل ۹: میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

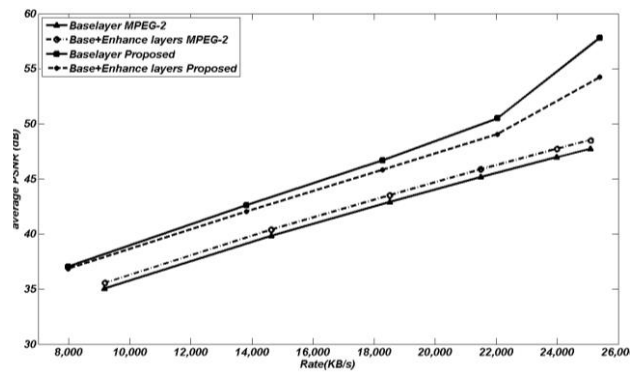
جدول ۴: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه  $1280 \times 720$  برای نرخ ارسال بیت

تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی

Sequence	fps	Intra Rate	compression Method	Bit Rate (KB/s)			Avg. PSNR (dB)	
				Base	Base+ Enhance1	HDTV	Base+ Enhance1	HDTV
Shields	50	1/5	MPEG-2	4,378	14,275	21,774	43.13	43.67
			Proposed	4,207	15,405	22,884	47.86	46.52
		1/9	MPEG-2	4,173	13,607	23,005	43.56	44.02
			Proposed	4,066	14,216	23,591	46.72	45.6



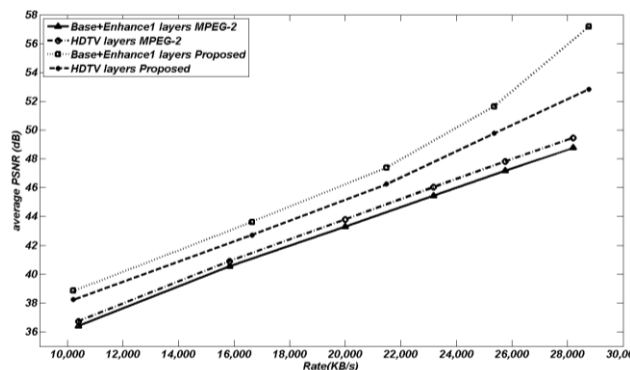
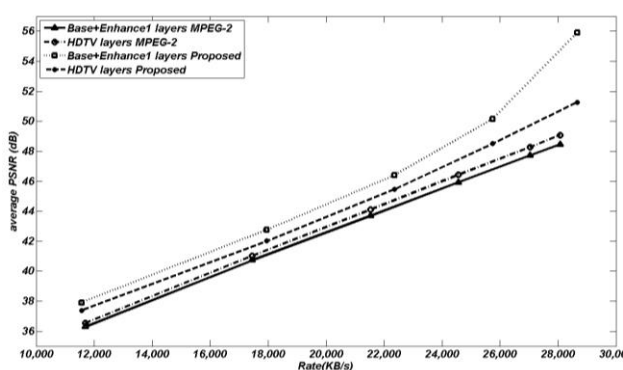
شکل ۱۲: میانگین کیفیت لایه های ویدئو Mobcal به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۹ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.



شکل ۱۱: میانگین کیفیت لایه های ویدئو Mobcal به اندازه  $1280 \times 720$  در نرخ ایترا ۱ به ۵ کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

جدول ۵: مقایسه نتایج شبیه سازی ویدئو Mobcal به اندازه  $1280 \times 720$  برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش ترکیبی و استاندارد با مقیاس پذیری مکانی-زمانی.

Sequence	fps	Intra Rate	compression Method	Bit Rate (KB/s)			Avg. PSNR (dB)	
				Base	Base+ Enhance1	HDTV	Base+ Enhance1	HDTV
Mobcal	50	1/5	MPEG-2	4,408	14,270	20,515	42.88	43.53
			Proposed	3,953	14,151	20,286	46.69	45.81
		1/9	MPEG-2	4,813	15,474	24,668	45.73	46.33
			Proposed	4,629	15,776	24,917	49.29	48.11



شکل ۱۴: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 1080$  در نرخ اینترا ۱ به ۹ کد شده به دو روش پیشنهادی و استاندارد.

شکل ۱۳: مقایسه میانگین کیفیت لایه های ویدئو Shields به اندازه  $1280 \times 1080$  در نرخ اینترا ۱ به ۵ کد شده به دو روش پیشنهادی و استاندارد.

جدول ۶: مقایسه نتایج شبیه سازی ویدئو Shields به اندازه  $1280 \times 1080$  برای نرخ ارسال بیت تقریباً یکسان لایه های مختلف کد شده به دو روش پیشنهادی و استاندارد.

Sequence	fps	Intra Rate	compression Method	Bit Rate (KB/s)			Avg. PSNR (dB)	
				Base	Base+ Enhance1	HDTV	Base+ Enhance1	HDTV
Shields	25	1/5	MPEG-2	2,203	16,510	25,758	47.16	47.84
			Proposed	2,236	16,965	25,357	51.65	49.77
		1/9	MPEG-2	1,497	10,302	17,448	40.73	41.02
			Proposed	1,520	10,809	17,949	42.77	42.01

هایی می شود که از آن به عنوان قاب مرجع استفاده می کنند. استفاده از ویژگی مقیاس پذیری، ضمن انعطاف پذیر کردن کدکننده، باعث ارتقا کیفیت سرویس نیز می شود. مقایسه نتایج کدکننده پیشنهادی برای ویدئو با درجه تفکیک بالا با سایر کدک های استاندارد مانند MPEG-4 می تواند زمینه ساز تحقیقات در آینده باشد.

### مراجع

[1].M. Domenski, A. Luczuk, S. Mackwiak, "On Improving MPEG Spatial Scalability", in proceedings of 2000 international conference of image processing ICIP Vancouver, Canada, vol. 2, pp. 848-851, 2000.

### ۶. نتیجه گیری

در این مقاله، کدکننده ترکیبی برای ویدئو با درجه تفکیک بالا پیشنهاد شده است. این کدکننده در دو حالت مختلف طراحی و شبیه سازی شده است. ویدئو با درجه تفکیک بالا ابتدا با استفاده از مقیاس پذیری مکانی و سپس با بکار بردن مقیاس پذیری مکانی-زمانی کد و ارسال شده است. بکارگیری روش پیشنهاد شده در این کدکننده، منجر به بهبود کیفیت ویدئو ارسالی نسبت به کدکننده مبتنی بر استاندارد MPEG-2 می شود. قاب های Intra با استفاده از الگوریتم SPIHT کد می شوند و کیفیت آن نسبت به قاب های Intra کد شده به روش استاندارد افزایش پیدا می کند. این بهبود منجر به افزایش کیفیت قاب

- [2]. L. Lima, Manerba F, Adami N, Signoroni A, Leonardi R, "Wavelet-Based Encoding for HD Applications", IEEE, ICME, pp. 1351-1354, 2007.
- [3]. T. Chiang, D. Anastassiou, "Hierarchical HDTV/SDTV Compatible Coding Using Kalman Statistical Filtering", IEEE Trans. On circuits for video Tech. Vol. 9, No. 3, 1999.
- [4]. M. Domenski, A. Luczak, S. Mackowaik, R. Swierczynfiski, "Hybrid coding of video with spatio-temporal Scalability using subband decomposition", in Proceedings 1999 of SPIE, vol. 3653, pp. 1018-1025, 1999.
- [5]. C. Poynton, "Digital Video and HDVT Algorithms and Interfaces", Los Altos, USA, pp. 111-126, 2003.
- [6]. S. Uzun, A. Amira "Real-time 2-D Wavelet Transform Implementation for HDTV Compression", in Elsevier Science, Real-Time Imaging 11, pp. 151-165, 2005.
- [7]. ISO/IEC International Standard 13813-2, information Technology Generic Coding of Moving Pictures and Associated Audio Information, 1995.
- [8]. J. Fowler, M. Tagliasacchi, B. Pesquet, "Wavelet-Based Distributed Source Coding of Video", In Proc. of the European signal Processing Conference, EUSIPCO, Antalia, Turkey, 2005.
- [9]. M. Weeks, M. A. Bayoumi, "Three-Dimensional Discrete Wavelet Transform Architectures", IEEE transactions on signal processing, 2002, vol. 50, No. 8.
- [10]. G. Kwon, Hyo-Kak Kim, Yoon Kim, Sung-Jea Ko, "An Efficient POCS-based Post-processing Technique Using Wavelet Transform in HDTV", IEEE Transactions Consumer Electronics, Vol. 51, No. 4, Pages: 1283-1288, 2005.
- [11]. M. Ghanbari, "Standard Codecs: Image Compression to Advanced Video Coding", IEE Telecommunication 49, London, U.K., 2003.
- [12]. B. Wu, X. Ji, D. Zhoa, W. Gao, "Wavelet Based distributed video coding with spatial scalability", circuits and systems, ISCAS, 2008 IEEE international symposium, pp. 3458-3461, 2008.
- [13]. N. Adami, M. Brescianini, Riccardo Leonardi, A. Signoroni. "A Fully Scalable Video Coding Scheme with Homologous Inter-scale Prediction.", ST Journal of Research, Vol. 3, No. 2, Pages 19-35, 2006.
- [14]. A. Said, W.a. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", in IEEE trans. circuits system video technology, vol.6, no. 3, pp. 243-250, 1996.
- [15]. N. Bozinovic, J. Konrad, "Modeling Motion for Spatial Svalability", Acoustics, Speech and Signal Processing ICASSP, Proceedings, 2006 IEEE Intenational conference, Vol. 2, PP. 29-32.
- [16]. R. Dianat, M Ghanbari, F. Marvasti "A Low Bit Rate Hybrid Wavelet-DCT Video Codec" IEEE Tran. On circuits and systems for Video Technology, VOL. 15, NO. 7, July 2005
- [17]. وحیدرضا صیرفیان، سیامک طالبی، "کدکردن تصاویر ویدئویی HDTV با استفاده از روش ترکیبی SPIHT-DCT"، هفدهمین کنفرانس مهندسی برق ایران، ICEE2009، تهران، دانشگاه علم و صنعت، ش. ۲۲۵۴، بهار ۱۳۸۸.

بکارگیری روش ترکیبی SPIHT-DCT با استفاده از مقیاس پذیری مکانی و مکانی-زمانی برای کد کردن تصاویر ویدئویی HDTV