

ارائه یک روش جدید فشرده‌سازی ویدیو مبتنی بر تبدیل موجک و شبکه‌های عصبی مصنوعی

محمد رحمانیان^(۱) - احمد حاتم^(۲) - محمدعلی شفیعیان^(۳)

(۱) کارشناس ارشد - گروه برق - مخابرات، دانشکده مهندسی برق، واحد بوشهر، دانشگاه آزاد اسلامی، بوشهر، ایران

(۲) استادیار - دانشکده مهندسی برق، دانشگاه هرمزگان، بندرعباس، ایران

(۳) کارشناس ارشد - دانشکده مهندسی برق، دانشگاه شیراز، شیراز، ایران

تاریخ پذیرش: ۱۳۹۴/۱۰/۲

تاریخ دریافت: ۱۳۹۴/۵/۲۹

چکیده: عمل فشرده‌سازی ویدیو با بهره‌گیری از افزونگی زمانی و فضایی موجود در رشته ویدیو انجام می‌گیرد. این افزونگی‌ها را می‌توان به ترتیب با اعمال یک تبدیل، تخمین و جبران‌سازی حرکت و کوانتیزاسیون به‌گونه‌ای کارآمد جبران‌سازی نمود. استانداردهای موجود برای فشرده‌سازی ویدیو مانند MPEG مبتنی بر تبدیل کسینوسی بلوکی می‌باشند که در آن تصویر ورودی باید به صورت بلوک‌هایی درآید. بنابراین همبستگی در عرض مرزهای بلوک حذف نشده و این امر منجر به مقدار قابل توجهی نویز بلوکی می‌گردد. با استفاده از تبدیل موجک می‌توان بر این پدیده غلبه نمود. اما چنین الگوی فشرده‌سازی از نظر زمان اجرا کمی کند است و برای غلبه بر این مشکل نیز از تبدیل موجک چندسطحی استفاده می‌شود. در واقع استفاده از تبدیل موجک باعث می‌شود که عمل تجزیه نمودن بر روی بُعد زمان نیز علاوه بر بُعد مکانی صورت گیرد. هدف از این مقاله دستیابی به یک الگوریتم فشرده‌سازی سریع‌تر و با نرخ فشرده‌سازی بالاتر می‌باشد. برای این منظور یک الگوریتم فشرده‌سازی ویدیو مبتنی بر تبدیل موجک سه سطحی معرفی و شبیه‌سازی شده است. برای انجام عمل پیش‌گویی از شبکه عصبی استفاده شده است. نتایج شبیه‌سازی‌ها مزایای استفاده از تبدیل موجک را آشکار می‌کنند. این نتایج نشان می‌دهند که الگوریتم پیشنهادی از نظر زمان اجرا سریع‌تر و از نظر نرخ فشرده‌سازی کارایی بهتری نسبت به استاندارد MPEG را از خود نشان می‌دهد. همچنین ویدیوی نهایی به دست آمده دارای کیفیت دیداری قابل قبولی برای چشم انسان می‌باشد و به دلیل نیاز به حجم کم حافظه می‌توان از آن در تجهیزات قابل حمل استفاده نمود.

کلمات کلیدی: فشرده‌سازی ویدیو، تخمین حرکت، جبران‌سازی حرکت، کدگذار و کدگشای ویدیو، استاندارد MPEG.

A Novel Video Compression Algorithm Using Wavelet Transform and Neural Network

Mohammad Rahmanian⁽¹⁾ - Ahmad Hatam⁽²⁾ - Mohammad Ali Shafieian⁽³⁾

(1) MSc. - Department of Electrical Engineering, Islamic Azad University, Boushehr Branch, Boushehr, Iran
mohammad_rah81@yahoo.com

(2) Assistant Professor - Department of Electrical and Computer Engineering, Hormozgan University,
Bandar Abbas, Iran
a.hatam@hormozgan.ac.ir

(3) MSc. - Department of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran
shafieian1381@yahoo.com

Videos are made up of a temporal sequence of frames and are projected at a proper rate to create the illusion of motion. This means that there exists a high correlation between adjacent temporal frames so that when projected at a proper rate, smooth motion is seen. Correlation between adjacent temporal frames is called interframe correlation. In order to decode compressed video bit stream uniformly by various platforms and devices, the bit stream format must be predefined. Thus, there must be a standard for a video compressor, which will enable all standard-compliant compressed video data to be decoded anywhere. The goal is to propose a new video compression algorithm based on wavelet transform and neural networks. Using wavelet transform leads to factorization in temporal as well as spatial domain. The goal in this paper is to achieve a compression algorithm which would be faster and has more compression ratio. Neural networks are used for prediction which is one of the most important functions in any video compression scheme. Furthermore, the proposed algorithm is compared with MPEG standard. Simulation results show the benefits of using wavelet transform which reveal that the proposed algorithm is faster and has better performance in some aspects compared to MPEG standard. The video which obtained from proposed algorithm is acceptable regarding human vision. since it needs less space for storing, it is suitable for portable devices.

Index Terms: Video compression, motion compensation, motion estimation, video encoder and decoder, MPEG standard

۱. مقدمه

در سال‌های اخیر، فعالیت‌های انجام شده در حوزه استانداردها و فن‌آوری‌های مربوط به سخت‌افزار تا حدی کامل شده‌اند که به راحتی بتوان سیگنال‌های ویدیویی که در فرمت‌های دیجیتال ذخیره شده‌اند را ارسال، ذخیره، پردازش و مشاهده نمود. بنابراین، فشرده‌سازی ویدیو یکی از ضروریاتی است که با آن مواجه هستیم و در سال‌های اخیر نیز پیشرفت قابل ملاحظه‌ای داشته است [۱]. نکته جالب توجه در این پیشرفت‌ها، مشارکت جامعه عمومی در آن‌ها می‌باشد زیرا گسترش کامپیوترها و رشد روزافزون فضای وب باعث سرازیر شدن حجم بالایی از اطلاعات به صورت لحظه‌ای به بخشی از کاربردهای تجاری یا خانگی شده است. بخش قابل توجهی از این اطلاعات برای نمایش دیداری در قالب متن، نمودار، تصویر و یا کاربردهای چندرسانه‌ای طراحی شده‌اند. تصاویر دیجیتال و ویدیوی دیجیتال به ترتیب به صورت تصاویر و رشته ویدیویی هستند که به فرمت باینری قابل خواندن توسط کامپیوتر شامل صفرها و یک‌ها تبدیل می‌شوند. معمولاً تصاویر یا ویدیوی دیجیتال با تبدیل سیگنال‌های پیوسته به فرمت دیجیتال به دست می‌آید اگرچه امروزه سیستم‌های کاملاً دیجیتال نیز رایج شده‌اند. فرکانسی که در آن اطلاعات ارسال، ذخیره، پردازش و در قالب دیجیتال نمایش داده می‌شود به سرعت در حال افزایش است و بنابراین طراحی روش‌های مهندسی برای ارسال، نگهداری و فراهم کردن و حتی تجمیع بصری اطلاعات از این اطلاعات به موضوعی مورد علاقه تبدیل شده است. با توجه به آنچه بیان شد، حجم وسیع اطلاعات و گسترش روزافزون آن لزوم استفاده از روش‌های فشرده‌سازی داده را بیشتر آشکار می‌سازد.

مفهومی که پردازش ویدیو را به موضوعی مورد علاقه تبدیل نموده است، آن است که این داده‌ها معمولاً چند بعدی هستند و می‌توانند به گونه‌ای مرتب شوند که برای سیستم بینایی انسان مناسب باشد. این مجموعه از تصاویر را می‌توان با استفاده از تکنیک‌های به وجود آمده توسط انسان، پردازش نمود. علاوه بر این، داده‌ها را می‌توان به گونه‌ای مرتب نمود که با زمان تغییر کنند و به کمک وسایل ذخیره‌سازی سریع‌تر، آن‌ها را به مجموعه‌ای از داده‌های ویدیوی دیجیتال تبدیل نمود. با داده‌کاوی در اطلاعات زمانی و مکانی که در یک ویدیو وجود دارد می‌توان تغییرات پدیده‌های فیزیکی موجود در ویدیو را تحلیل نمود. یکی از موضوعاتی که در پردازش ویدیو به آن پرداخته می‌شود، پیدا کردن مدل‌های آماری ساده برای توصیف تصاویر است. تصاویر طبیعی، ویژگی‌های مشترک معینی را دارا هستند. هدف از مدل‌سازی آماری، دریافت ویژگی‌های مشخص شده با تعداد کمی از پارامترها است، به طوری که این ویژگی‌ها می‌توانند به عنوان اطلاعات اولیه در کاربردهای پردازش تصویر و ویدیو، مانند فشرده‌سازی مورد استفاده قرار گیرند [۱]. یکی از راهکارها در این مسأله استفاده از حوزه تبدیل است و تاکنون انواع مختلفی از الگوریتم‌های پردازش تصویر ارائه شده‌اند که یک تصویر را به گونه‌ای تفکیک می‌کنند که در مؤلفه‌های مختلفی قابل تحلیل باشند و هر یک از آن‌ها اطلاعات موجود در آن مقیاس را به دست می‌دهند. یکی از ابزارهایی که برای تفکیک تصویر به مؤلفه‌های مختلف

وجود دارد، تبدیل موجک نام دارد که به عنوان ابزاری در الگوریتم پیشنهادی، معرفی و از آن برای فشرده‌سازی ویدیو استفاده شده است. کار بر روی شبکه‌های عصبی مصنوعی که معمولاً «شبکه‌های عصبی» نیز نامیده می‌شوند، از زمانی آغاز شده که متوجه شدیم مغز انسان محاسبات خود را به روش کاملاً متفاوتی از کامپیوترهای دیجیتال انجام می‌دهد. مغز انسان، یک کامپیوتر (سیستم پردازش اطلاعات) بسیار پیچیده، غیرخطی و موازی است و قادر است اجزای ساختاری خود که نورون نام دارند را طوری سازمان‌دهی کند که محاسبات خاصی مانند تشخیص الگو، ادراک و کنترل را چندین مرتبه سریع‌تر از سریع‌ترین کامپیوترهای دیجیتال انجام دهد. واضح است که شبکه عصبی قدرت محاسباتی خود را در درجه اول از ساختار گسترده موازی حجیم خود و در درجه دوم از توانایی یادگیری و تعمیم‌دهی به دست می‌آورد. طبیعت تعمیم‌دهی در شبکه‌های عصبی، آن را به ابزاری مناسب برای انجام عمل پیش‌گویی که یکی از اجزای اصلی فشرده‌سازی ویدیو است، تبدیل می‌کند که در الگوریتم پیشنهادی نیز استفاده شده است.

در روش‌های پیشین عمدتاً از تبدیل فرکانسی کسینوسی گسسته جهت کاهش میزان وابستگی‌های مکانی و فشرده‌سازی بیشتر ویدیو استفاده می‌شد که از آن جمله می‌توان استانداردهای H.261 و MPEG^۱ را نام برد. از آنجایی که چشم انسان واکنش بیشتری به فرکانس‌های پایین نسبت به فرکانس‌های بالا دارد، بنابراین فرکانس‌های بالا را با صفر نمایش داده و در نتیجه دامنه اطلاعات کاهش خواهند یافت. دنباله نهایی با روش‌های کدگذاری آنتروپی مانند کدینگ هافمن و RLE^۲ فشرده‌سازی می‌شود [۱]. از آنجایی که تاکنون در الگوریتم فشرده‌سازی ویدیو استفاده از شبکه عصبی کمتر مورد توجه قرار گرفته است در این مقاله به دنبال آن خواهیم بود که به کمک شبکه عصبی سطوح پایین‌تر تبدیل موجک را از روی سطوح بالاتر تخمین بزنیم و به عبارت دیگر از شبکه عصبی به عنوان قسمت پیش‌گویی الگوریتم استفاده نماییم. تاکنون مطالعات و پژوهش‌هایی در حوزه فشرده‌سازی ویدیو با استفاده از تبدیل موجک صورت گرفته است. در [۲] روش فشرده‌سازی ویدیو مبتنی بر تبدیل موجک سه بعدی معرفی شده است. تفاوت مهم این الگوریتم با سایر الگوریتم‌های کلاسیک در مسأله تجزیه زمانی است. این روش در عین اینکه کیفیت ویدیو را حفظ می‌کند، زمان پردازش و تجزیه رشته ویدیو را به میزان قابل ملاحظه‌ای کاهش می‌دهد و همچنین نیاز به حافظه بالا را هم برطرف می‌نماید. اکثر الگوریتم‌های فشرده‌سازی ویدیو، مبتنی بر روش‌های دو بعدی و به کارگیری تکنیک‌های تخمین و جبران‌سازی حرکت می‌باشند [۳] اما نیازمند زمان زیادی برای پردازش می‌باشند. از سوی دیگر، روش‌های کارآمدی مبتنی بر کدینگ موجک وجود دارد که قادرند بدون جبران‌سازی، افزونگی فضایی و زمانی را استخراج نمایند [۴]. هدف از تجزیه فضایی، استخراج افزونگی‌های فضایی در پیکسل‌های مجاور یک فریم است، درحالی‌که، تجزیه زمانی به دنبال استخراج افزونگی‌های زمانی در فریم‌های مجاور هم و بدون جبران‌سازی حرکت می‌باشد.

روش جدید مبتنی بر استفاده از SPIHT بیان شده است. در این روش که به نام SPIHT3D تغییر یافته یا همان MSPIHT3D شناخته شده است، آنتروپی به‌گونه‌ای متفاوت از SPIHT3D توزیع شده است و کدگذاری را نیز بهینه نموده است. نتایج شبیه‌سازی‌ها نشان می‌دهند که استفاده از این روش به PSNR و نرخ فشرده‌سازی بیشتری را به دست می‌دهد، ضمن آنکه بر زمان محاسبه نیز تأثیر نمی‌گذارد.

الگوریتم‌هایی برای جبران‌سازی حرکت و تبدیل بلوک مبتنی بر تبدیل موجک m باند در کدگذار ویدیو در [۱۱] معرفی شده است تا به این وسیله فشرده‌سازی بالاتر و پیچیدگی محاسباتی کمتری در جبران‌سازی و پیش‌گویی حرکت به دست آید. ایده‌ای جدید نیز برای کدگذاری ویدیو مبتنی بر تبدیل موجک در [۱۲] معرفی شده است. در این ایده، ابتدا هر فریم در رشته ویدیو با به‌کارگیری تبدیل موجک، به زیرباندی تجزیه می‌شود. پس از آن، الگوریتم با پیمایش زیرباندها طبق الگوی از پیش تعیین‌شده به جستجوی پیکسل‌های شاخص در زیرباندها می‌پردازد. این امر کارایی از دیدگاه نرخ - عوایج را در فشرده‌سازی با اتلاف بهبود می‌بخشد. نتایج عملی نشان می‌دهند که روش پیشنهادی در فشرده‌سازی با اتلاف و بی‌تلف کارایی بهتری را نسبت به استانداردهایی مانند SPIHT از خود نشان می‌دهد. در فشرده‌سازی با اتلاف، روش پیشنهادی در نرخ بیت‌های بالاتر کارایی بهتری را نسبت به نرخ بیت‌های پایین‌تر از خود نشان می‌دهد. علاوه بر این، نرخ بیت متوسط که برای فشرده‌سازی بی‌تلف در روش پیشنهادی مورد نیاز است، پایین است. شایان ذکر است که استفاده از تبدیل موجک در سایر جنبه‌های پردازش ویدیو نیز کاربرد دارد. در [۱۳] از تبدیل موجک برای بازیابی ویدیو مبتنی بر محتوا به منظور فشرده‌سازی انرژی ویدیو استفاده شده است. استفاده از تبدیل موجک، فشرده‌سازی بیشتری را در ضرایب انرژی از خود نشان داده است. با توجه به شناخت اولیه به دست آمده از روش‌های فشرده‌سازی موجود، هدف این مقاله آن است که بتوان با استفاده از تبدیلات کارآمد دیگری مانند تبدیل موجک در حالت ۳ بعدی [۱۴-۱۶] و نیز با استفاده از شبکه عصبی مصنوعی به ارتقا و بهبود الگوریتم‌ها و استانداردهای موجود پرداخته شود. این مقاله در ادامه از بخش‌های زیر تشکیل شده است. در بخش ۲ به معرفی اصطلاحات به‌کار رفته در الگوریتم پیشنهادی و نیز معرفی مختصر الگوریتم مذکور پرداخته شده است. سپس در بخش ۳ به تشریح اجزای بلوک دیاگرام الگوریتم پیشنهادی پرداخته شده است. شبیه‌سازی‌های انجام گرفته بر روی الگوریتم پیشنهادی و مقایسه آن با استاندارد MPEG در بخش ۴ مورد بررسی قرار گرفته است. در نهایت، در بخش ۵ به نتیجه‌گیری پرداخته شده است.

۲. معرفی اصطلاحات به‌کار رفته و الگوریتم پیشنهادی

یک ویدیو یک رشته زمانی است به این معنی که به صورت رشته‌ای از تصاویر می‌باشد که در بُعد زمان روی هم انباشته شده‌اند. از آنجایی که نرخ فریم‌های ویدیو بالاست (مثلاً بیش از ۱۵ فریم در ثانیه) و پارامترهای دوربین (طول لنز، زاویه دید، مکان) معمولاً بین فریم‌های سریع تغییر نمی‌کند، همبستگی زیادی بین فریم‌های مجاور زمانی

در [۵] از راهکارهای مبتنی بر جستجوی کامل برای کاهش محاسبات استفاده شده است. به منظور کاهش افزونگی در داده ویدیو، از تبدیل کسینوسی گسسته و تبدیل موجک گسسته استفاده شده است. یکی از مزایای تبدیل کسینوسی گسسته، یافتن انطباق ماکروبلوک‌ها برای مقادیر با فرکانس پایین و سپس افزایش آن برای مقایسه با فرکانس‌های بالاتر است و هدف استفاده از تبدیل موجک در فشرده‌سازی ذخیره داده ویدیو در یک فضای کم است. نتایج شبیه‌سازی‌ها نشان می‌دهند که تبدیل موجک گسسته کارایی بهتری را نسبت به تبدیل کسینوسی گسسته در یافتن انطباق ماکروبلوک‌ها از خود نشان می‌دهد.

یک روش مؤثر مبتنی بر تبدیل موجک برای فشرده‌سازی ویدیو و صوت در [۶] ارائه شده است که در آن یک مکانیزم مؤثر برای فشرده‌سازی سیگنال با کارایی بالا و نیز مقدار قابل قبولی کیفیت دیداری و شنیداری مورد بررسی قرار گرفته است. نتایج شبیه‌سازی نشان می‌دهند که مدل پیشنهادی نیازهای یک سیستم مخابراتی از دیدگاه زمان پردازش و کارایی فشرده‌سازی را برآورده می‌نماید. در الگوریتم پیشنهادی، به منظور دستیابی به سرعت و کارایی بالا و قابلیت پردازش با نرخ بیت پایین از مکانیزم فشرده‌سازی ویدیو به صورت تفاضلی نیز پشتیبانی می‌شود. با توجه به پهنای باند در دسترس، از دو پیکربندی در کدگذار/کدگشای پیشنهادی برای ویدیو استفاده شده است که در پهنای باندهای مختلف کارکرد مختلفی را از خود نشان می‌دهند. در [۷] نیز فشرده‌سازی ویدیو مبتنی بر تبدیل موجک و یک تکنیک تفاضلی فشرده‌سازی تصویر مورد بررسی قرار گرفته است. این الگوریتم از قابلیت کاهش مقیاس در تبدیل موجک و قابلیت کاهش فشرده‌سازی تفاضلی تصویر استفاده می‌کند. الگوریتم فشرده‌سازی تفاضلی تصویر که به اختصار DICA^۲ نامیده شده است [۸]، بر مبنای مفهوم ارسال تفاوت بین فریم‌ها به جای خود فریم‌ها می‌باشد. در این الگوریتم از مزیت وجود تشابه بین تصاویر فریم‌ها بهره برده شده است. نتایج شبیه‌سازی‌ها نشان می‌دهند که بیشترین نرخ فشرده‌سازی که با استفاده از موجک تفاضلی در این روش به دست می‌آید نسبت ۷۲:۱ است که در این حالت افت قابل ملاحظه‌ای در کیفیت دیداری تصویر نیز ایجاد نمی‌گردد. در [۹]، دو الگوریتم فشرده‌سازی ویدیو مبتنی بر تبدیل موجک پیشنهاد شده است. در این مقاله، دو قسمت مهم در فشرده‌سازی ویدیو یعنی تخمین حرکت و جبران‌سازی حرکت بر پایه حرکت بخش‌ها یا همان Segments صورت گرفته است. روش پیشنهادی در این مقاله مبتنی بر دو الگوریتم تبدیل موجک به نام‌های SPHIT^۱ و SWT^۵ می‌باشد. نتایج به دست آمده برای دو الگوریتم مذکور کارایی خوبی را از دیدگاه PSNR^۶ و MSE^۷ نشان می‌دهند که این امر این دو تکنیک را نسبت تبدیل DCT^۸ کارآمدتر می‌نماید. همچنین نتایج شبیه‌سازی‌ها نشان می‌دهند که مقدار PSNR و MSE در روش SPIHT نسبت به روش SWT بهتر است. علاوه بر این، بازسازی تصویر/ ویدیو با اطمینان بالایی انجام می‌شود ضمن اینکه کیفیت آن نیز حفظ می‌گردد. علاوه بر این، ایده استفاده از SPIHT برای فشرده‌سازی ویدیو در [۱۰] نیز مطرح شده است که در آن یک

۳. تشریح الگوریتم پیشنهادی

ابتدا رشته ویدیو ورودی وارد بلوک پیش‌پردازش می‌شود. در این بلوک عملیات اولیه مانند متعادل‌سازی هیستوگرام، تعیین نوع نمونه‌برداری و در صورت نیاز تبدیل ویدیوی رنگی به ویدیوی سیاه و سفید انجام می‌شود. هر یک از فریم‌های ویدیو ابتدا در بافری که در شکل (۱) به نام فریم فعلی نام‌گذاری شده است نگه‌داری می‌شوند تا پس از تعیین نوع هر فریم عمل فشرده‌سازی روی آن‌ها انجام گیرد. در این الگوریتم برای فشرده‌سازی فریم I از روشی مشابه استاندارد فشرده‌سازی JPEG2000 استفاده می‌شود. برای این منظور ابتدا فریم ورودی جبران‌سازی اختلاف می‌شود و سطح پیکسل‌های آن از یک عدد صحیح مشخص کاسته می‌شود که برای آن داریم [۱۷]:

$$I'_k[m, n] = I_k[m, n] - 2^{\text{size}_k} - 1 \quad (1)$$

که در آن منظور از size_k بیشینه مقدار ممکن از مؤلفه ورودی k ام است. به عبارت دیگر اگر یک مؤلفه ورودی دارای عمق پیکسل ۸ بیت باشد، آنگاه $2^{\text{size}_k} - 1 = 2^{8-1} = 128$ خواهد بود. این سطح DC پس از عمل ICT در گیرنده (کدگشا) به هر پیکسل اضافه می‌شود که برای آن داریم [۱۷]:

$$I_k[m, n] = I'_k[m, n] + 2^{\text{size}_k} - 1 \quad (2)$$

مؤلفه‌هایی که از تغییر سطح DC به دست آمده‌اند و دارای مؤلفه‌های قرمز، سبز و آبی هستند، بایستی به دستگاه مختصات دیگری تبدیل شوند. این امر از دیدگاه دستیابی به نرخ فشرده‌سازی بالاتر ضروری است. مؤلفه‌های قرمز، سبز و آبی نوعاً دارای همبستگی بالایی هستند. بنابراین فشرده‌سازی مستقل این مؤلفه‌ها در مختصات RGB چندان کارآمد نیست و این امر منجر به ایجاد رنگ‌های اشتباه می‌گردد. اما اگر این مؤلفه‌ها به یک مجموعه معادل متعامد تبدیل شوند، این امکان وجود دارد که مؤلفه‌ها را به طور مستقل از هم در مختصات جدید فشرده‌سازی نمود و به فشرده‌سازی با کارایی بالاتر دست یافت.

برای یک فشرده‌سازی با اتلاف تبدیل مؤلفه پیش‌رو^{۱۵} با عبارت زیر قابل بیان است [۱۷، ۱۸]:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.144 \\ -0.16875 & -0.33126 & 0.5 \\ 0.5 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} I_0 \\ I_1 \\ I_2 \end{bmatrix} \quad (3)$$

که در آن پارامترهای I_0 ، I_1 و I_2 همان مؤلفه‌های ورودی Y_0 ، Y_1 و Y_2 مؤلفه‌های با مختصات تغییر یافته هستند. باید توجه داشت که مؤلفه‌های ورودی رابطه (۳) به ترتیب مؤلفه‌های قرمز، سبز و آبی هستند. پس از آنکه عمل تبدیل روی آن انجام شد، مؤلفه‌های خروجی متناظر به ترتیب Y ، Cb و Cr می‌باشند. شایان ذکر است که در انتها، رشته ویدیو از مختصات رنگی RGB به مختصات $YCbCr$ تبدیل می‌شود.

پس از اعمال تبدیل مؤلفه پیش‌رو، هر فریم به قطعه‌هایی^{۱۶} و هر یک از قطعه‌ها توسط یک تبدیل موجک دوبعدی ۳ سطحی به زیرباندهایی تجزیه می‌شوند. برای انجام تبدیل موجک گسسته دوبعدی از بانک فیلتری FIR، Duabechies 9/7 استفاده می‌شود و نیز عکس عمل تبدیل موجک استفاده می‌شود.

وجود دارد و این فریم‌ها به هم شباهت دارند مگر آنکه اشیا مشخصی در یک صحنه خیلی سریع حرکت کنند. به همین دلیل گفته می‌شود که یک ویدئو دارای افزونگی زمانی می‌باشد [۱۷، ۱۸]. افزونگی زمانی از این جهت مهم است که به‌جای آنکه فریم‌ها به صورت مستقل کد شوند، می‌توان اختلاف آن‌ها را کد نمود. اگر افزونگی بین فریم‌ها زیاد باشد به این معنی است که تفاضل (اختلاف) بین فریم‌ها کم است و به این ترتیب آنتروپی کم می‌شود که این امر برای فشرده‌سازی نکته‌ای مثبت محسوب می‌شود [۱۸]. از همین ویژگی در الگوریتم پیشنهادی نیز استفاده شده است. یک رشته ویدیویی از تعدادی گروه تصویر^۹ تشکیل شده است که در هر گروه سه دسته تصویر وجود دارد که تصویر I، تصویر B و تصویر P نام دارند. تصویر I یا I-frame^{۱۱} (اینترا فریم) تصاویر کد شده به صورت اینترا^{۱۱} نیز نام دارند و در آن‌ها فقط افزونگی فضایی استخراج شده و بنابراین توسط خود آن‌ها و بدون استفاده از هیچ فریم مرجعی در دیگر گروه تصاویر کد می‌شوند. فریم‌های I به عنوان نقاط دسترسی تصادفی در جریان داده ویدیو استفاده می‌شوند و دارای کمترین نرخ فشرده‌سازی در الگوریتم پیشنهادی هستند. تصویر P یا P-frame^{۱۲} فریم‌هایی به صورت کدینگ پیشگویی رو به جلو^{۱۳} هستند که در آن فریم واقعی توسط فریم قبل از آن (I یا P) کد می‌شود. در تصاویر B یا B-frame^{۱۴}، فریم‌ها بر اساس دو فریم مرجع کد می‌شوند. یک فریم قبل و یک فریم بعد که می‌تواند فریم‌های I یا P باشد. کدگذاری دوطرفه یا میانبایی شده بیشترین فشرده‌سازی را ایجاد می‌کند. بلوک دیاگرام الگوریتم پیشنهادی برای فشرده‌سازی (کدگذاری ویدیو) در شکل (۱) و بلوک دیاگرام روش کدگشایی ویدیو در شکل (۲) نشان داده شده است. در این بلوک دیاگرام کدگذاری و کدگشایی هر سه فریم I، P و B نشان داده شده است. همان‌گونه که از این بلوک دیاگرام‌ها نیز پیداست، کدگذاری ویدیو با روش پیشنهادی همانند کدگذاری MPEG دارای طبیعت پیشگویی است. در این الگوریتم فریم‌های ویدیو در حوزه زمان ناهمبسته شده که این امر با محاسبه اختلاف بین فریم فعلی و فریم‌های قبلی و با استفاده از تخمین و جبران‌سازی حرکت قابل تحقق است. فشرده‌سازی ویدیو به این دلیل قابل تحقق است که فریم‌های تفاضل دارای محدوده دینامیکی از نظر شدت می‌باشد و تقریباً دارای توزیع لاپلاس هستند. در حالت کلی نقش یک تبدیل کاستن وابستگی‌های بین نمونه‌های ویدیو می‌باشد. الگوریتم پیشنهادی، یک الگوریتم مبتنی بر تبدیل موجک و شبکه عصبی می‌باشد. در استاندارد MPEG به‌جای موجک از تبدیل کسینوسی گسسته استفاده می‌شود. انتظار می‌رود که الگوریتم پیشنهادی دارای سرعت بیشتری از MPEG و نرخ فشرده‌سازی بهتری نسبت به MPEG باشد. البته لازم به ذکر است که برای مقایسه سرعت اجرای دو الگوریتم از یک کامپیوتر استفاده شده است.

در بخش بعد به بررسی بلوک‌های مختلف ساختار کدگذار و کدگشا در الگوریتم پیشنهادی برای فشرده‌سازی ویدیو پرداخته خواهد شد و هر یک از آن‌ها به تفصیل بررسی خواهند گردید

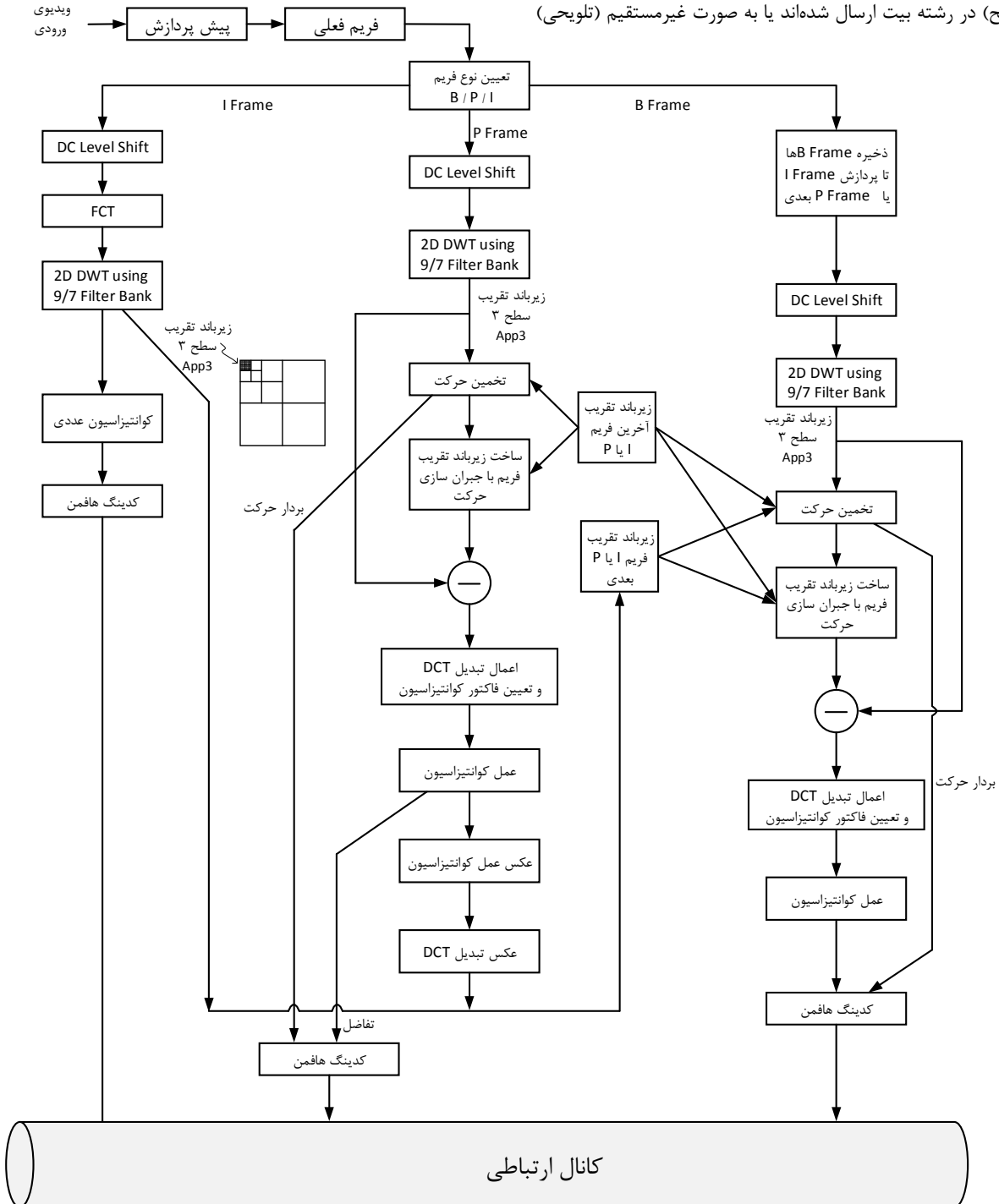
برای انجام عمل کوانتیزاسیون روی فریم I در الگوریتم پیشنهادی، گام کوانتیزاسیون $Q(b)$ در یک کوانتایزر یکنواخت برای زیرباند b از رابطه زیر محاسبه می‌شود [۱۷]:

$$Q(b) = 2^{R_b + \epsilon_b + \mu_b} (1 + \frac{\mu_b}{2^{11}}) \quad (4)$$

که در آن R_b محدوده دینامیکی نامی زیرباند b ، بهره ϵ_b (گین) زیرباند و μ_b مفسر و ماننسی ضرابی هستند که یا مستقیماً (به‌طور صریح) در رشته بیت ارسال شده‌اند یا به صورت غیرمستقیم (تلویحی)

برای باند LL ارسال شده‌اند. بهره (گین) در جهت پایین‌گذر برابر یک و در جهت بالاگذر برابر ۲ می‌باشد. بنابراین باندهای LH و HL در هر یک از سطوح دارای بهره ۲ می‌باشند در حالی که باند HH دارای بهره ۴ می‌باشد. با گام کوانتیزاسیون تعریف شده در رابطه ۴، یک ضریب در یک زیرباند با استفاده از رابطه زیر تعیین می‌شود [۱۷، ۱۸]:

$$X_{qb}(u, v) = \text{sign}(X_b(u, v)) \left\lceil \frac{|X_b(u, v)|}{Q(b)} \right\rceil \quad (5)$$



شکل (۱): بلوک دیاگرام کدگذار در الگوریتم پیشنهادی
 Fig. (1) : Encoder block diagram for proposed algorithm

الگوریتم پیشنهادی ابتدا یک تبدیل موجک گسسته سه سطحی روی فریم‌های I و P اعمال می‌شود و تخمین حرکت بر روی زیرباند تقریب سطح ۳ انجام می‌گیرد.

پس از آنکه حرکت یک ماکروبلوک تخمین زده شد، نوبت به تراز نمودن آن با ماکروبلوک مناسب در فریم مرجع و یافتن اختلاف بین ماکروبلوک فعلی و ماکروبلوک تراز شده در فریم مرجع می‌رسد. ماکروبلوک‌های مرجع را می‌توان از بافر زیر باند تقریب آخرین فریم I یا P به دست آورد. پس از ساخت زیر باند تقریب با جبران‌سازی حرکت و یافتن اختلاف بین آن و زیر باند تقریب سطح ۳ تبدیل کسینوسی گسسته روی آن اعمال شده و پس از اعمال عمل کوانتیزاسیون و عکس آن، عکس تبدیل کسینوسی گسسته نیز روی آن انجام می‌شود و فریم به دست آمده در بافر زیر باند تقریب فریم I یا P بعدی قرار می‌گیرد.

باید توجه داشت که مسأله تخمین حرکت در تجزیه به زیر باندهای (موجک‌های) فضایی و زمانی با آنچه در کدگذاری هیبرید وجود دارد، کمی متفاوت است. این امر به این دلیل است که در کدگذاری فضایی و زمانی که در آن قابلیت مقیاس‌پذیری وجود دارد، رشته‌ای با نرخ فریم پایین‌تر با فیلتر کردن فضایی و زمانی جبران‌سازی شده تولید می‌شوند و به این ترتیب فریمی که برای ارسال به گیرنده در نظر گرفته می‌شود در حالت ایده‌آل همان فریم با نرخ پایین است. در این حالت هر یک از حرکت‌های مصنوعی که توسط خطای حرکت ایجاد می‌شوند در خروجی فریم با نرخ پایین‌تر نیز مشاهده می‌شود.

با توجه به آنچه تاکنون بیان شد، می‌توان مشاهده نمود که یک حلقه بسته در سمت کدگذار ایجاد شده است. لازم به ذکر است که چنین حلقه‌ای نیز در کدگشا و برای بازسازی ویدیو وجود دارد.

در بلوک تبدیل کسینوسی گسسته عمل تبدیل روی بلوک‌های تفاضل بین زیرباند تقریب سطح ۳ و جبران‌سازی شده آن انجام می‌شود. ورودی تبدیل کسینوسی گسسته دوبعدی، بلوک‌های تفاضل بوده و خروجی آن ضرایب تبدیل کسینوسی گسسته اعمال شده بر آن می‌باشد. اگرچه بلوک‌های لوما دارای اندازه 16×16 پیکسل می‌باشد، اما تبدیل کسینوسی گسسته دو بعدی بر روی بلوک‌های 8×8 پیکسل انجام می‌شود. علاوه بر این، در این بلوک فاکتور کوانتیزاسیون نیز تعیین می‌شود که در بخش بعد پیرامون آن نیز صحبت خواهد شد. برای کوانتیزه نمودن فریم‌های P و B، هر بلوک 8×8 از ضرایب تبدیل کسینوسی گسسته با استفاده از کوانتیزاسیون یکنواخت کوانتیزه می‌شوند که گام‌های کوانتیزاسیون در آن‌ها با توجه به مدل کیفیت دیداری انسان تعیین می‌شود. در الگوریتم پیشنهادی، ماتریس کوانتیزاسیون برای کدگذاری دورن‌فریمی و بین‌فریمی همانند آنچه در استاندارد MPEG به کار می‌رود، در نظر گرفته شده است. نرخ فشرده‌سازی را می‌توان با ضرب ماتریس‌های کوانتیزاسیون در یک مقیاس (فاکتور) کوانتیزاسیون کنترل نمود. در استاندارد JPEG برای کوانتیزه کردن ضرایب تبدیل کسینوسی گسسته از رابطه زیر استفاده می‌شود [۱۷]:

$$y_q(k, l) = \left[\frac{y(k, l)}{Q(k, l)} + 0.5 \right] \quad 1 \leq k \quad \text{و} \quad 1 \leq l \quad (6)$$

به منظور ارتقای کارایی فشرده‌سازی بدون از دست‌دادن کیفیت دیداری، در الگوریتم پیشنهادی از وزن‌دهی فرکانسی اندازه گام‌های کوانتیزاسیون در کوانتیزه‌های یکنواخت نیز استفاده شده است [۱۷]. این وزن‌دهی فرکانسی به زیرباندهای مختلف مؤلفه لوما اعمال می‌شود و مبتنی بر مشاهده فواصل می‌باشد.

آخرین مرحله در کدگذاری فریم I در فشرده‌سازی مبتنی بر تبدیل موجک، کدگذاری آن‌تروپی می‌باشد. از آنجایی که در اکثر ضرایب جزئیات پس از عمل کوانتیزاسیون برابر با صفر می‌باشند، می‌توان از کدگذاری تکرار - طول 17 به منظور تبدیل ضرایب سمبل‌ها برای ذخیره یا ارسال استفاده نمود. می‌توان از هر یک از روش‌های کدگذاری آن‌تروپی مانند هافمن، حسابی، Golomb-Rice برای این منظور استفاده نمود. از کدگذاری تکرار - طول نوعاً برای کدگذاری منابع باینری مانند اسناد، نقشه‌ها و تصاویر رونوشت استفاده می‌شود که همه آن‌ها شامل پیکسل‌هایی با نقطه‌های سیاه و سفید هستند. در این منابع باینری، پیکسل سفید توسط عدد ۱ و یک پیکسل سیاه با عدد صفر نمایش داده می‌شود. هنگام کدگذاری چنین رشته باینری در خروجی می‌توان تعداد تکرارهای صفرها یا یک‌ها را شمارش نمود و به جای آنکه پیکسل‌های سیاه و سفید را به تنهایی کدگذاری نماییم، بهتر است که تعداد تکرارهای صفرها و یک‌ها را کدگذاری نماییم. یک تکرار - طول از صفرها نشان‌دهنده تعداد صفرهای پشت سر هم است که بین یک‌ها قرار دارند و به طور مشابه یک تکرار - طول از یک‌ها، تعداد یک‌های پشت سرهم است که بین صفرها وجود دارد. یک روش برای کدگذاری تکرار - طول‌ها استفاده از کلمه‌هایی با طول ثابت b متناظر با بیشترین تکرار - طول L-1 است که در آن $L=2^b$ می‌باشد. از آنجایی که از یک مجموعه از کلمه‌های کد استفاده می‌شود، بایستی به نوعی به کدگشا اعلام کرد که کدام تکرار - طول در ابتدا آمده است. به صورت نوعی، نخستین تکرار، یک تکرار سفید با طول صفر است؛ پس از آن تکرارها تغییر کرده و می‌توان کدگشایی را انجام داد. کدگذاری با طول متغیر دارای کارایی بیشتری نسبت به کدگذاری با طول ثابت است. بنابراین با کدگذاری هافمن می‌توان تکرارها را با کارایی بهتری کدگذاری نمود. به منظور کدگذاری فریم‌های P و B نخست حرکت هر ماکروبلوک در فریم فعلی با توجه به مرجع آن در فریم‌های قبل یا بعد تخمین زده می‌شود. در این بلوک، حرکت با استفاده از معیار مجموع قدرمطلق تفاضل تخمین زده می‌شود و علاوه بر این بردارهای حرکت برای هر ماکروبلوک تولید می‌گردد. این بردارهای حرکت به کدگذاری آن‌تروپی فرستاده می‌شوند تا پس از عمل کدگذاری ارسال یا ذخیره شوند. شایان ذکر است که عمل تخمین حرکت فقط برای مؤلفه لوما انجام می‌گیرد و برای مؤلفه‌های کروما از همان بردار حرکت به دست آمده برای مؤلفه لوما با یک مقیاس‌گذاری مناسب استفاده می‌شود که در انتخاب مقیاس نوع فرمت نمونه‌برداری نیز لحاظ می‌شود. فرآیند جستجو برای تخمین حرکت به عهده کدگذاری گذاشته می‌شود که به پردازنده و سخت‌افزار بلادرنگ مناسب برای این کار بستگی دارد. در

P یا B به دست می‌آیند که با اعمال عکس عمل تبدیل موجک گسسته دوبعدی فریم P یا B کدگشایی شده حاصل می‌شود. این فریم‌های بازسازی شده در بافر فریم‌های کدگشایی شده قرار داده می‌شوند و پس از آن با تغییر ترتیب نمایش آن‌ها آماده نمایش می‌گردند. به این ترتیب رشته ویدیو کدگشایی شده نمایش داده می‌شود.

۴. شبیه‌سازی و مقایسه نتایج

به‌منظور برآورد کارایی الگوریتم پیشنهادی، الگوریتم MPEG و الگوریتم پیشنهادی مورد شبیه‌سازی قرار گرفته و پارامترهای مورد نظر برای مقایسه با الگوریتم پیشنهادی از آن استخراج گردیده است. برای بررسی و شبیه‌سازی الگوریتم‌ها، تعدادی رشته ویدیو در نظر گرفته شده است و از یک کامپیوتر با پردازنده ۶۴ بیتی Intel Core i7 (2.9 GHz) با 12GB حافظه RAM استفاده شده است. برای شبیه‌سازی الگوریتم و نیز شبیه‌سازی استاندارد MPEG و مقایسه آن دو با یکدیگر از تعدادی رشته ویدیوی استفاده شده است که مشخصات هر یک از آن‌ها در جدول (۱) آورده شده است.

به دلیل بالا بودن حجم شبیه‌سازی و عدم دسترسی به یک سیستم پردازشی با سرعت بالا، روشی که در فشرده‌سازی استفاده گردیده است به این صورت می‌باشد که ابتدا بلوک‌های مختلف مانند بلوک‌های تخمین حرکت، ساخت فریم‌های مختلف، جبران‌سازی حرکت، تبدیل کسینوسی گسسته و کوانتیزاسیون، کدینگ و ... به صورت مستقل از یکدیگر شبیه‌سازی شده‌اند. پس از شبیه‌سازی هر بلوک و اطمینان از درستی عملکرد هر یک، این بلوک‌ها در کنار یکدیگر قرار گرفتند تا سیستم نهایی را تشکیل دهند.

در اجرا و شبیه‌سازی الگوریتم MPEG و الگوریتم پیشنهادی، پارامترهایی در نظر گرفته شده است تا بتوان به‌واسطه آن‌ها مقایسه‌ای را بین استاندارد MPEG و الگوریتم پیشنهادی انجام داد که برای این منظور پارامترهای PSNR، BPP^{۱۸} و زمان اجرا در نظر گرفته شده است. این نتایج برای استاندارد MPEG و الگوریتم پیشنهادی و برای رشته ویدیوی مختلف که نام آن‌ها در جدول (۱) آورده شده است، در قالب نمودار در ادامه آورده شده‌اند.

نتایج مربوط به PSNR نشان می‌دهند که روش پیشنهادی دارای مقدار کمتری PSNR نسبت به استاندارد MPEG می‌باشد به طوری که روش پیشنهادی نسبت به استاندارد MPEG به میزان متوسط ۳ dB افت در مقدار PSNR را نشان می‌دهد. مهم‌ترین دلیل این امر به خاطر بازسازی زیرباند‌های فریم با استفاده از شبکه عصبی می‌باشد. البته در طراحی شبکه دقت کافی صورت گرفته است.

علاوه بر این زمان اجرای شبیه‌سازی برای رشته ویدیوهای مختلف با کامپیوتری که مشخصات آن بیان شد، در جدول (۲) آورده شده است. همان‌گونه که در ابتدا نیز انتظار می‌رفت، الگوریتم پیشنهادی دارای زمان اجرای کمتری نسبت به استاندارد MPEG بوده و به عبارت دیگر سریع‌تر است.

که در آن $y(k, l)$ ضرایب تبدیل کسینوسی گسسته دو بعدی کوانتیزه شده و $y_q(k, l)$ ضرایب کوانتیزه شده متناظر به صورت یکنواخت می‌باشد. پارامتر $Q(k, l)$ همان گام کوانتیزاسیون برای (k, l) آمین ضریب است. علامت $[x]$ نشان‌دهنده عملیات گرد کردن (جزء صحیح یک عدد) می‌باشد. لازم به ذکر است که عکس عمل کوانتیزاسیون برای ضرایب تبدیل کسینوسی گسسته را می‌توان به صورت زیر انجام داد [۱۷]:

$$\hat{y}(k, l) = y_q(k, l) \times Q(k, l) \quad (7)$$

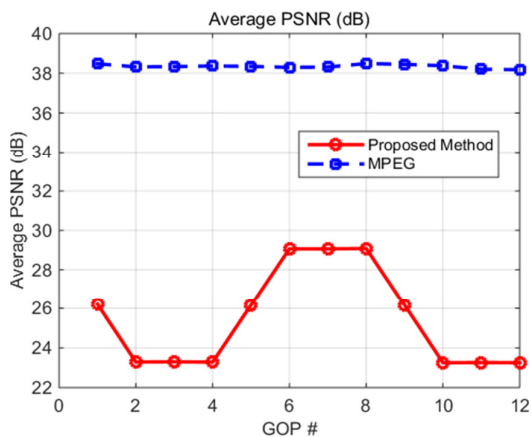
در حالت کلی، عکس عملیات انجام شده در قسمت کدگذاری در قسمت کدگشا بر روی زیرباند تقریب فریم‌ها اعمال می‌شود. برای این منظور ابتدا فریم ورودی که یک رشته بیت صفر و یک کدگشایی می‌شود که این عمل با توجه به نوع فریم صورت می‌گیرد. برای یک فریم I ابتدا یک عکس تبدیل موجک گسسته دو بعدی بر روی زیرباند تقریب کدگشایی شده اعمال می‌شود که پس از اعمال تبدیل مؤلفه معکوس و افزودن سطح DC فریم I کدگشایی شده حاصل می‌گردد. سپس بر روی این فریم سه سطح تبدیل موجک گسسته دو بعدی اعمال می‌گردد که از زیرباند‌های به دست آمده از اعمال این تبدیل برای بازسازی فریم‌های دیگر استفاده می‌شود. در صورتی که فریم کدگشایی شده در این مرحله فریم P یا B باشد ابتدا عکس عمل تبدیل کسینوسی گسسته و عکس عمل کوانتیزاسیون بر روی آن اعمال می‌شود، سپس بردارهای حرکت به دست آمده و زیر باند‌های تقریب فریم‌های I و P بازسازی شده عمل جبران‌سازی حرکت روی زیر باند تقریب صورت می‌گیرد. حاصل این دو عمل با یکدیگر جمع شده و در زیر باند‌های تقریب سطح ۳ فریم‌های P یا B قرار می‌گیرد. از این به بعد بازسازی فریم‌ها با استفاده از زیر باند‌های تقریب قرار داده شده در این بافر صورت می‌گیرد.

برای بازسازی سایر زیرباند‌ها در فریم‌های P یا B از فریم I کمک گرفته می‌شود. برای این کار پس از اعمال تبدیل موجک گسسته دو بعدی روی فریم I کدگشایی شده، با استفاده از زیر باند‌های افقی، عمودی و قطری به دست آمده در سطح ۳ و زیرباند تقریب سطح ۳ سه شبکه عصبی متناظر با این سه زیرباند را آموزش می‌دهیم. برای این منظور از یک شبکه عصبی پیش‌خور سه لایه با ۲۰ نرون در لایه ورودی، ۱۵ نرون در لایه پنهان و ۴ نرون در لایه خروجی استفاده شده است. از همین شبکه برای بازسازی زیرباند‌های افقی، عمودی و قطری سایر سطوح نیز استفاده می‌شود. پس از آموزش شبکه‌ای عصبی، زیرباند متناظر سطح ۳ متناظر با آن شبکه عصبی با استفاده از هر یک از شبکه‌های عصبی بازسازی می‌شوند و پس از آن تفاضل بین زیرباند‌های واقعی در سطح ۳ برای فریم I و فریم‌های حاصل از شبکه عصبی محاسبه می‌گردند و تصویر تفاضل به دست آمده برای زیرباند با میان‌یابی مناسب به زیرباند بازسازی شده متناظر که از شبکه عصبی آموزش دیده به دست آمده است اضافه می‌شود. به این ترتیب زیرباند‌های سطوح ۱، ۲ و ۳ کدگشایی شده برای هر یک از فریم‌های

Table (1) : Video sequences specifications used in the simulation

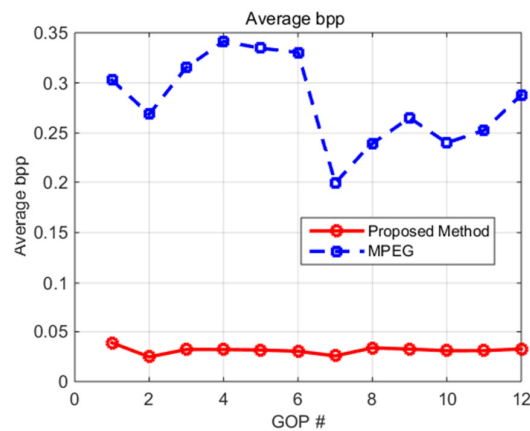
جدول (۱) : مشخصات رشته‌های ویدیوی در نظر گرفته شده برای شبیه‌سازی

نام رشته ویدیویی	تفکیک پذیری (پیکسل)	تعداد فریم	نرخ پخش فریم (fps)
Rhinos	۳۲۰ × ۲۴۰	۱۰۸	۳۰
Xylophone	۳۲۰ × ۲۴۰	۱۳۵	۳۰
Foreman	۱۷۶ × ۱۴۴	۲۹۷	۳۰
NEWS	۱۷۶ × ۱۴۴	۲۸۸	۳۰
Clair	۱۷۶ × ۱۴۴	۴۸۶	۳۰
Akio	۱۷۶ × ۱۴۴	۲۸۸	۳۰



(ب)

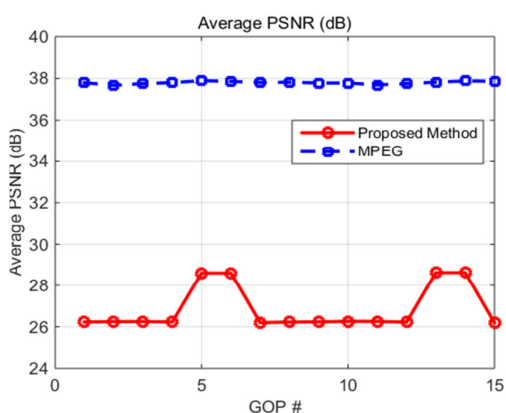
(b)



(الف)

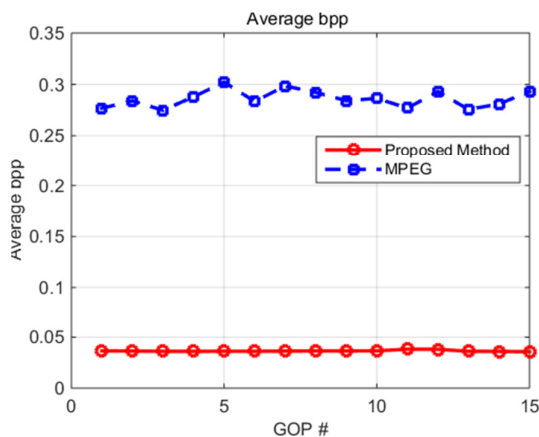
(a)

شکل (۳) : نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی Rhinos؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) PSNR متوسط
Fig (3) : Performance for every GOP in terms of GOP number for Rhinos video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR



(ب)

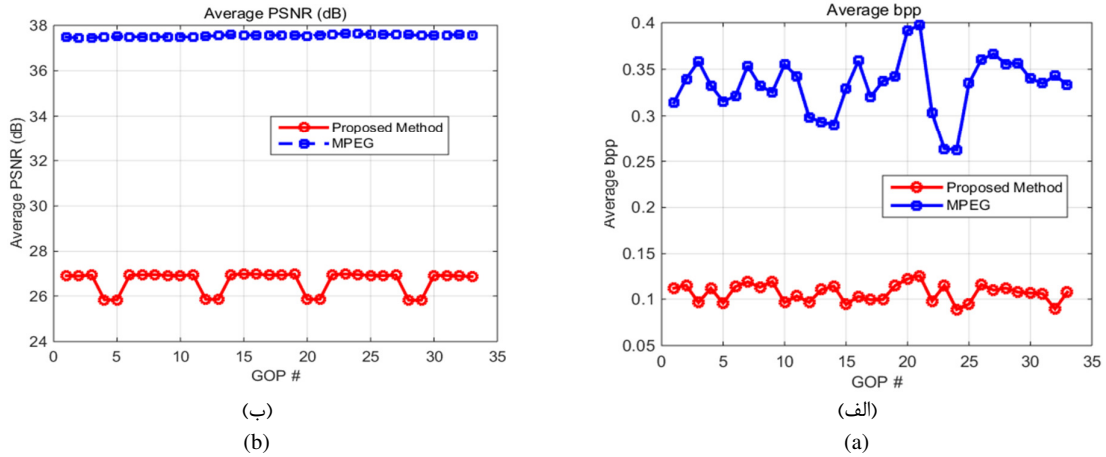
(b)



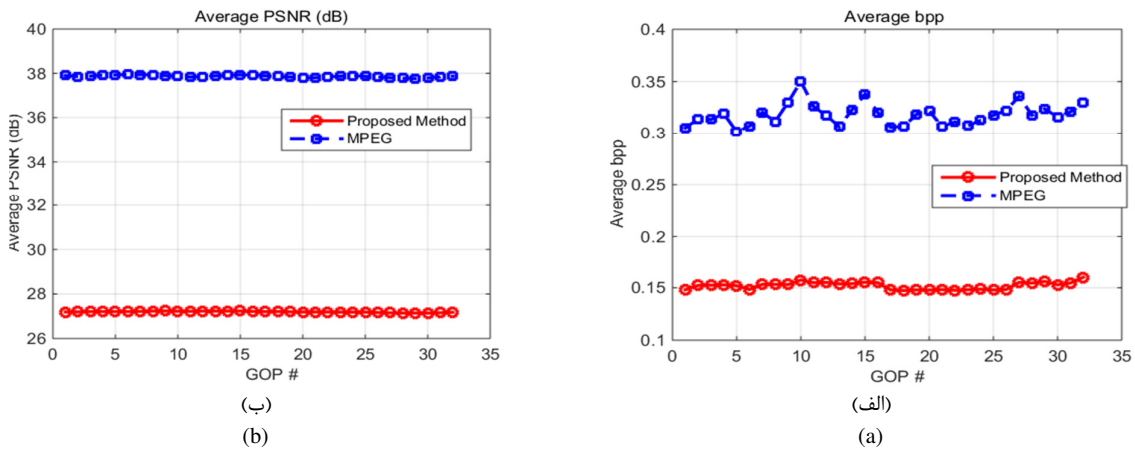
(الف)

(a)

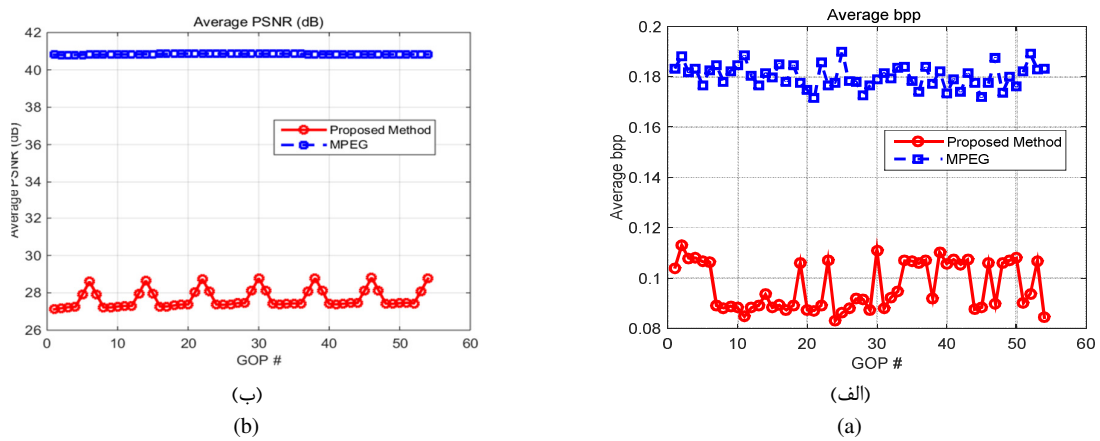
شکل (۴) : نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی Xylophone؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) PSNR متوسط
Fig (4) : Performance for every GOP in terms of GOP number for Xylophone video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR



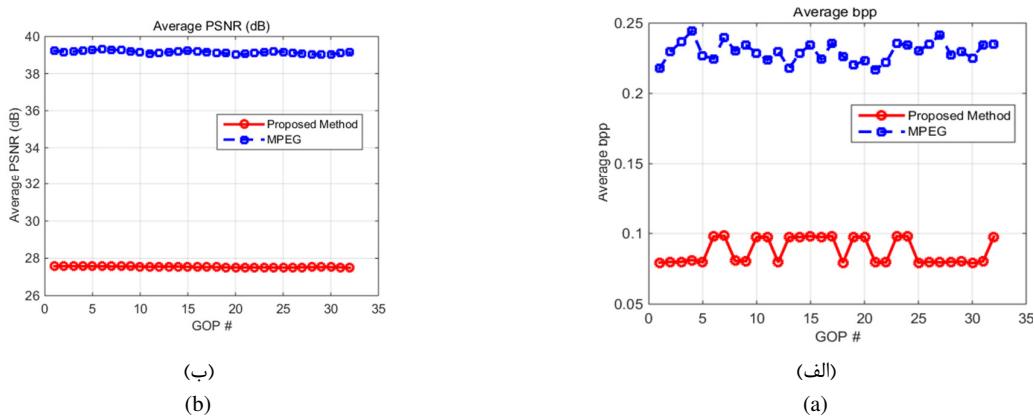
شکل (۵): نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی Foreman؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) متوسط PSNR
Fig (5) : Performance for every GOP in terms of GOP number for Foreman video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR



شکل (۶): نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی NEWS؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) متوسط PSNR
Fig (6) : Performance for every GOP in terms of GOP number for NEWS video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR



شکل (۷): نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی Clair؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) متوسط PSNR
Fig (7) : Performance for every GOP in terms of GOP number for Clair video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR



شکل (۸) : نمودار کارایی برای هر GOP برحسب شماره GOP برای رشته ویدیوی Akio؛ (الف) بیت بر پیکسل متوسط (BPP)، (ب) PSNR متوسط
 Fig (8) : Performance for every GOP in terms of GOP number for Akio video sequence; (a) Average bit per pixel (BPP), (b) Average PSNR

Table (2): Simulation run times for video sequences

جدول (۲): زمان اجرای شبیه‌سازی برای رشته ویدیوهای مختلف

زمان اجرا (دقیقه)		نام رشته ویدیویی
استاندارد MPEG	الگوریتم پیشنهادی	
۷/۳۳	۳/۲۸	Rhinos
۱۰/۰۴	۲/۹۵	Xylophone
۱۰/۹۳	۷/۶۱	Foreman
۱۰/۲۷	۶/۹۷	NEWS
۱۶/۲۵	۱۳/۸	Clair
۱۰/۴۱	۷/۶۵	Akio

مذکور نیز عمدتاً در دستگاه‌های موبایل مورد استفاده قرار می‌گیرد و با وجود آن که دارای حجم کمی هستند که برای این دستگاه‌ها مناسب است، کیفیت پایینی دارند.

پی‌نوشت:

- 1- Moving Picture Experts Group (MPEG)
- 2- Run Length Encoding (RLE)
- 3- Differential Image Compression Algorithm (DICA)
- 4- Set Partitioning in Hierarchical Trees (SPIHT)
- 5- Stationary Wavelet transform (SWT)
- 6- Peak Signal to Noise Ratio (PSNR)
- 7- Mean Square Error (MSE)
- 8- Discrete Cosine Transform (DCT)
- 9- Group of Pictures (GOP)
- 10- Intra-frame
- 11- Intracoded
- 12- Predictive coded picture or predicted frame
- 13- Forward Predictive Coding
- 14- Bidirectional or interpolated frame or BI-directionally predictive coded
- 15- Forward Component Transform (FCT)
- 16- Tiles
- 17- Run Length Coding (RLC)
- 18- Bit per Pixel (BPP)

۵. نتیجه‌گیری

در این مقاله یک الگوریتم برای فشرده‌سازی ویدیو مبتنی بر تبدیل موجک سه سطحی و شبکه عصبی، پیشنهاد و پیاده‌سازی گردید که از شبکه عصبی برای انجام عمل پیش‌گویی که یکی از بخش‌های اصلی ساختار یک الگوریتم فشرده‌سازی می‌باشد، استفاده شده است. در نهایت، پارامترهای کارایی الگوریتم پیشنهادی با استاندارد MPEG مقایسه گردیده که در برخی از پارامترها، همان‌گونه که انتظار می‌رفت، الگوریتم پیشنهادی نسبت به استاندارد MPEG کارایی بهتری را از خود نشان می‌دهد. نتایج شبیه‌سازی‌ها مزایای استفاده از تبدیل موجک را آشکار می‌کنند. این نتایج نشان می‌دهند که الگوریتم پیشنهادی از نظر زمان اجرا سریع‌تر و از نظر نرخ فشرده‌سازی کارایی بهتری نسبت به استاندارد MPEG را از خود نشان می‌دهد. همچنین ویدیوی نهایی به دست آمده دارای کیفیت دیداری قابل قبولی برای چشم انسان می‌باشد و به دلیل نیاز به حجم کم حافظه می‌توان از آن در تجهیزات قابل حمل استفاده نمود. به این ترتیب، روش پیشنهادی برای کاربردهایی که در آن‌ها سرعت پردازش و نرخ فشرده‌سازی بالا مورد نیاز است، اما کیفیت تصویر در درجه دوم اهمیت قرار دارد مناسب می‌باشد که از آن جمله می‌توان دستگاه‌های موبایل را نام برد. مشابه این امر را می‌توان در فرمت 3gp نیز مشاهده نمود. فرمت

References

- [1] Bovik, A., Handbook of Image and Video Processing, 2000, Academic Press.
- [2] E. Moyano, "A new fast 3D wavelet transform algorithm for video compression. in communications", Proceeding of the IEEE/PACIRM, Vol. 1, pp. 156-159, Victoria ,BC, 2001.
- [3] J.R. Ohm, "Three-dimensional subbandcoding with motion compensation", IEEE Trans. on Image Processing, pp. 559-571, 1994.
- [4] P.N.Topiwala, "Wavelet image and video compression", Kluwer Academic, 1998.
- [5] T. Aziz, D.R.J. Dolly, "Motion estimation and motion compensated video compression using DCT and DWT", International Journal of Emerging Technology and Advanced Engineering, Vol. 2, No. 12, pp. 667-671, 2014.
- [6] Y.S. Cheng, G.D. Huang, "Audio/video compression applications using wavelets", Proceedings of the IEEE/IJCNN, pp. 2214-2218, Honolulu, HI, 2002.
- [7] R. Adhami, "Video compression technique using wavelet transform", Proceeding of the IEEE/AERO, Vol. 4, pp. 449-455, Aspen, CO, 1996.
- [8] J. Bridge, "Differential image compression" .Dr. Dobbs's Journal, pp. 38-48, 1991.
- [9] M. Nema, L. Gupta, N.R. Trivedi, "Video compression using SPIHT and SWT wavelet", International Journal of Electronics and Communication Engineering, Vol. 5, No. 1, pp. 1-8, 2012.
- [10] D.Z. Athmane, et al., "New video compression using MSPiHT3D transformation", Proceeding of the MIC-Telecom, Istanbul, Turkey, 2014.
- [11] G.D. Bhokare, C.S. Gupta, V.M. Gadre, "Video compression using 3-band wavelet transform", in The National Conference on Communications, National Conference on Communications. pp. 321-326, 1996.
- [12] D.V. Babu, N.R. Alamelu, "A novel morpho codec for medical video compression based on lifting wavelet transform", Asian Journal of Scientific Research, Vol. 7, No. 1, pp. 85-93, 2014.
- [13] S.D. Thepade, N. Yadav, "Novel efficient content based video retrieval method using cosine-haar hybrid wavelet transform with energy compaction", Proceeding of the IEEE/ICCUBEA, pp. 615-619, Pune, Feb. 2015.
- [14] Y. Xu, L. XU. "The performance analysis of wavelet in video coding system", Proceeding of the IEEE/ICMT, pp. 878-883, Hangzhou, July 2011.
- [15] M. Lei, L. Hang, "Video compression algorithm based on wavele ttransform", Proceeding of the IEEE/, pp. 535-538, Shenyang, Aug. 2009.
- [16] R.Y. Omaki, G. Fujita, T. Onoye, I. Shirakawa, "Embedded zerotree wavelet based algorithm for video compression", Proceeding of the IEEE/TENCON, Vol. 2, pp. 1343-1346, Cheju Island, 1999.
- [17] K.S. Thyagarajan, "Still image and video compression with MATLAB", John Wiley & Sons, Inc, 2011.
- [18] Z. Li, M.S. Drew, J. Liu, "Fundamentals of multimedia", United States of America: Pearson Prentice Hall, 2004.