کاهش پیچیدگی محاسباتی واحد پیش‌گویی بین‌قابی در استاندارد کد کننده‌ی ویدئویی پربازده مبتنی بر ویژگی‌های مکانی و زمانی

محمدرضا رمضان‌پور[۱]، استادیار؛ ریحانه خورسند[۲]، استادیار

۱- گروه مهندسی کامپیوتر - واحد مبارکه - دانشگاه آزاد اسلامی - اصفهان - ایران - ramezanpour@mau.ac.ir

۲- گروه مهندسی کامپیوتر - واحد دولت آباد - دانشگاه آزاد اسلامی - اصفهان - ایران - r.khorsand@iauda.ac.ir

چکیده: استاندارد کد کننده‌ی ویدئویی پربازده، آخرین استاندارد بین‌المللی فشرده‌سازی ویدئو می‌باشد که اولین نسخه آن در سال ۲۰۱۳ منتشر شد. در این استاندارد به دلیل افزایش تعداد بلوک‌های مورد ارزیابی و افزایش تعداد حالت‌های پیش‌گویی، پیچیدگی محاسباتی افزایش یافته است. لذا، استفاده از روش‌هایی که باعث کاهش حالت‌های پیش‌گویی و کاهش پیچیدگی محاسباتی شوند، ضروری است. برای این منظور، در این مقاله جهت کاهش پیچیدگی محاسبات روشی ارائه شده است که با توجه به افزونگی زمانی بین فریم‌ها، تعداد بلوک‌های مورد ارزیابی و با توجه به افزونگی مکانی تعداد حالت‌های پیش‌گویی درون‌قابی کاهش می‌یابد. ابتدا بلوک با بیشترین هم‌بستگی که بیشترین شباهت به بلوک جاری دارد و در اطراف بلوک هم مکانش در فریم مرجع می‌باشد به‌عنوان همبسته‌ترین بلوک معرفی شده و با توجه به اندازه آن بلوک، در مورد اندازه بلوک جاری تصمیم‌گیری می‌شود و از ارزیابی تعدادی از بلوک‌ها که شانس کمتری جهت انتخاب شدن دارند صرف‌نظر می‌گردد. سپس با توجه به شباهت پیکسل‌ها در جهت افقی یا عمودی تعداد حالت‌های پیش‌گویی کاهش می‌یابد. روش پیشنهادی با ۱۰ ویدئو با رزولوشن‌های مختلف که دارای بافت‌های متفاوت می‌باشند ارزیابی شده است. نتایج شبیه‌سازی نشان می‌دهد که روش پیشنهادی می‌تواند زمان کد کردن را به‌طور متوسط ۲۸/۸ درصد کاهش دهد، در حالی که نرخ بیت دارای افزایش ناچیزی در حدود ۰/۹ درصد است.

واژه‌های کلیدی: استاندارد کد کننده ویدئویی پربازده، افزونگی زمانی و مکانی، پیش‌گویی بین‌قابی، پیش‌گویی درون‌قابی.

Reduction of Computational Complexity of Inter Prediction Unit in High Efficiency Video Coding (HEVC) Standard Based on Spatial and Temporal Features

M. R. Ramezanpour[1], Assistant Professor; R. Khorsand[2], Assistant Professor

1- Department of Computer Engineering, Mobarakeh Branch, Islamic Azad University, Mobarakeh, Isfahan, Iran, Email: ramezanpour@mau.ac.ir

2- Department of Computer Engineering, Dolatabad Branch, Islamic Azad University, Isfahan, Iran, Email: r.khorsand@iauda.ac.ir

Abstract: High efficiency video coding standard is the latest standard of international video compression whose first version was released in 2013. In this standard, computational complexity has increased due to the increase in the number of evaluated blocks and prediction modes. Therefore, it is necessary to use methods that reduce the computational complexity and prediction modes. In this paper, in order to reduce the computational complexity, a method has been proposed in which due to the temporal redundancy between frames, the number of evaluated blocks and spatial redundancy, the number of intra-frame prediction modes decreases. First, the block with the highest correlation and similarity to the current block and is around the co-located block within the reference frame is considered as the most correlated block and according to the size of that block, the size of the current block is determined, and evaluating a number of blocks having less chance to be selected is ignored. Then, due to the similarity of pixels in the horizontal or vertical direction, the number of prediction modes decreases. The proposed method is evaluated with 10 different video resolutions having different textures. The simulation results show that the proposed method can reduce the coding time 28.8 % on average, while the BD-rate has a slight increase of 0.9 %.

Keywords: High efficiency video coding standard, Temporal and Spatial Redundancy, Inter Prediction, Intra Prediction.

تاریخ ارسال مقاله: ۱۳۹۸/۰۴/۲۷

تاریخ اصلاح مقاله: ۱۳۹۸/۰۸/۲۶

تاریخ پذیرش مقاله: ۱۳۹۸/۰۹/۰۲

نام نویسنده مسئول: ریحانه خورسند

نشانی نویسنده مسئول: ایران - اصفهان - دولت آباد - دانشگاه آزاد اسلامی واحد دولت آباد - دانشکده خوارزمی - گروه مهندسی کامپیوتر.

۱- مقدمه

در سال ۲۰۰۳، استاندارد فشرده‌سازی ویدئویی H.264/AVC (استاندارد کد گذاری ویدئوی پیشرفته[۱])، توسط دو سازمان بین‌المللی ITU-T[۲] (اتحادیه بین‌المللی ارتباطات) و ISO/IEC[۳] (سازمان بین‌المللی استانداردسازی/کمیته بین‌المللی الکتروتکنیک) معرفی شد [۱]. علی‌رغم این‌که این استاندارد نسبت به استانداردهای قبلی دارای توانایی بیش‌تری در فشرده‌سازی ویدئو است ولی پیدایش فرمت‌های ویدئویی بالاتر از HD[۴] نیاز به فشرده‌سازی فراتر از H.264/AVC را ایجاد نموده است. از سوی دیگر افزایش ترافیک شبکه ناشی از درخواست‌های مرتبط با ویدئو نیز به عنوان یک دغدغه مهم باعث شد که استاندارد کد کننده ویدئو پر بازده، یعنی HEVC[۵] معرفی گردد [۲].

در استاندارد HEVC افزایش تعداد بلوک‌های مورد ارزیابی و هم‌چنین افزایش تعداد حالت‌های پیش‌گویی را می‌توان از دلایل افزایش دو برابری فشرده‌سازی نسبت به استاندارد H.۲۶۴/AVC بیان کرد. اما از طرفی این عوامل باعث شده است که پیچیدگی محاسباتی استاندارد HEVC افزایش یابد [۳-۴] و یکی از بخش‌هایی که بار محاسباتی زیادی را تحمیل نموده است واحد پیش‌گویی بین‌قابی است. لذا، از جمله اهداف ضروری در این استاندارد، استفاده از روش‌هایی است که باعث کاهش حالت‌های پیش‌گویی و پیچیدگی محاسباتی می‌شوند و فراهم آوردن این امر به عنوان یک چالش عمده قرار گرفته و در رأس پژوهش‌های محققانی است که در زمینه‌ی استاندارد کد کننده‌ی ویدئویی پربازده فعالیت می‌کنند.

موارد ذکر شده در بالا، عامل اصلی ایجاد این انگیزه شد که در این مقاله یک الگوریتم به منظور کاهش پیچیدگی محاسباتی در استاندارد کد کننده‌ی ویدئویی پربازده ارائه گردد که برای این منظور از محدود کردن تعداد حالت‌های پیش‌گویی (درون‌قابی و بین‌قابی) و نیز عدم محاسبه‌ی هزینه‌ی تمامی حالت‌ها، استفاده می‌کند که این امر منجر به کاهش سربار محاسباتی و در نتیجه‌ی آن بهبود پیچیدگی محاسباتی در استاندارد کدکننده‌ی ویدئویی پربازده می‌شود. الگوریتم پیشنهادی، مبتنی بر همبستگی مکانی و زمانی پیکسل‌ها است و با کاهش تعداد حالت‌های پیش‌گویی باعث کاهش پیچیدگی محاسباتی شده است.

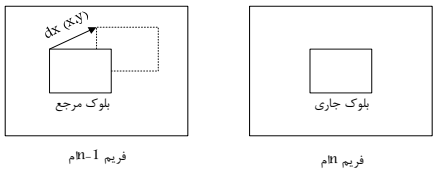
ساختار مقاله در ادامه به صورت زیر می‌باشد: در بخش ۲، پیش‌زمینه‌ی تحقیق که شامل انواع روش‌های پیش‌گویی و ساختار درختی استاندارد HEVC است تا حدی که بقیه بخش‌های مقاله قابل فهم باشند، توضیح داده شده است. سپس در بخش ۳، پیشینه‌ی تحقیق تشریح خواهد شد. در بخش ۴، روش پیشنهادی با جزییات کامل بیان خواهد شد. نتایج شبیه‌سازی روش پیشنهادی در بخش ۵ ارائه شده و در نهایت در بخش ۶ نتیجه‌گیری بیان خواهد شد.

۲- پیش‌زمینه

۲-۱- انواع روش‌های پیش‌گویی در استاندارد HEVC

روش‌های پیش‌گویی که توسط استاندارد HEVC پشتیبانی می‌شود، نسبت به استانداردهای قبلی قابلیت انعطاف بیش‌تری را دارند و امکان پیش‌گویی‌های دقیق‌تر و در نتیجه فشرده‌سازی مؤثرتری را فراهم می‌آورد. عملیات پیش‌گویی به طور کلی به دو دسته‌ی پیش‌گویی درون‌قابی و بین‌قابی تقسیم می‌شوند[۵] که در ادامه توضیح هر یک آورده شده است.

- پیش‌گویی بین‌قابی: در پیش‌گویی بین‌قابی، پیش‌گویی توسط فریم‌هایی که قبلاً کدگذاری و بازسازی شده‌اند، صورت می‌گیرد [۶]. در این روش تصویر به بلوک‌های غیر هم‌پوشان تقسیم شده و به دنبال هر بلوک در فریم قبلی می‌گردد. هر بلوک در یک محدوده‌ای از فریم قبلی جستجو گردیده و بهترین مکانی که با بلوک جاری منطبق باشد را پیدا می‌کند، میزان جابه‌جایی بلوک جاری نسبت به بلوک منطبق شده با یک برداری به نام بردار حرکت نشان داده می‌شود. همان‌طور که در شکل ۱ نشان داده شده است، بردار حرکت شامل دو مقدار x و y است که مقدار x میزان تغییر حرکت روی محور xها و مقدار y میزان تغییر حرکت روی محور yها را نشان می‌دهد. علاوه بر بردار حرکت یک ماتریس جبران نیز ارسال می‌شود که ضرایب ماتریس اختلاف بین بلوک جاری و بهترین بلوک منطبق را نشان می‌دهد. به دلیل محاسبات زیاد جهت جستجوی هر بلوک، استاندارد HEVC از دو حالت ادغام[۶] و پرش[۷] نیز استفاده می‌کند. حالت پرش حالتی است که بردار حرکت (۰و۰) باشد و تمام ضرایب ماتریس جبران نیز صفر گردد، در این وضعیت حالت پرش فعال می‌شود و از ارسال بردار حرکت خودداری می‌کند. در صورتی که حالت پرش انتخاب گردد فرآیند جستجو متوقف می‌گردد. در صورتی حالت ادغام انتخاب می‌گردد که بردار حرکت بلوک جاری مانند بردار حرکت یکی از بلوک‌های همسایه باشد، در این حالت روند جستجو متوقف می‌گردد و از بررسی بقیه بلوک‌ها صرف نظر می‌گردد [۷].



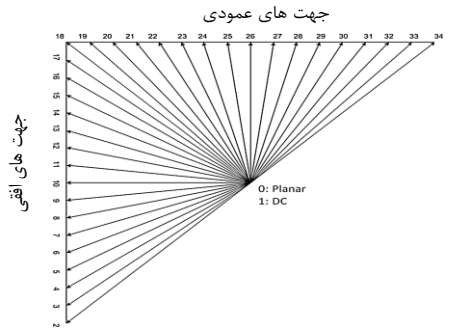
شکل ۱: چگونگی پیش‌گویی بین‌قابی و جستجو در بلوک‌های فریم قبلی در استاندارد HEVC

- پیش‌گویی درون‌قابی: استاندارد HEVC برای حذف افزونگی مکانی از پیش‌گویی درون‌قابی استفاده می‌نماید. در این پیش‌گویی از پیکسل‌های بلوک‌های مجاور که قبلاً کد شده‌اند جهت پیش‌گویی استفاده می‌گردد. برای به دست آوردن بهترین جهت پیش‌گویی از ۳۵ جهت مختلف استفاده می‌نماید. که از ۳۵ حالت،

یک حالت برای پیش‌گویی Planar (در این حالت از یک تابع خطی هموار کننده که از پیکسل‌های بالا و چپ بلوک استفاده می‌کند برای پیش‌گویی استفاده می‌شود)، یک حالت برای پیش‌گویی DC (در این حالت از میانگین تمام پیکسل‌های مرجع جهت پیش‌گویی استفاده می‌شود) و ۳۳ حالت برای پیش‌گویی زاویه‌ایی می‌باشد که در شکل ۲ نشان داده شده است و در پایان حالتی که دارای کم‌ترین هزینه می‌باشد را انتخاب می‌نماید. این ۳۵ جهت روی بلوک‌هایی با اندازه‌های مختلف از ۶۴×۶۴ تا ۴×۴ بررسی می‌گردد تا بلوک با اندازه مناسب و حالت بهینه را انتخاب نماید. جهت انتخاب بهترین حالت و اندازه بلوک از تابع هزینه نشان داده شده در رابطه ۱ استفاده می‌گردد.

$$J = SSD + \lambda . R \qquad (۱)$$

به‌طوری‌که SSD مجموع مجذور اختلاف بین پیکسل‌های بلوک کد کننده اصلی با بلوک جاری می‌باشد و λ ضریب لاگرانژ می‌باشد و R تعداد بیت‌های حقیقی است که برای کد کردن بلوک مورد نیاز است [۸].

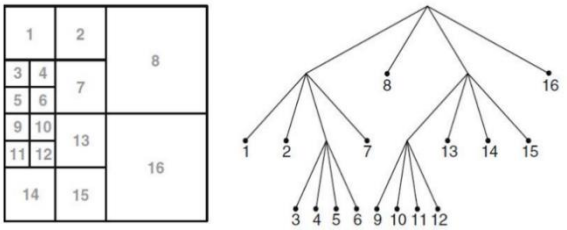


شکل ۲: نمایش ۳۵ حالت برای پیش‌گویی درون‌قابی [۸]

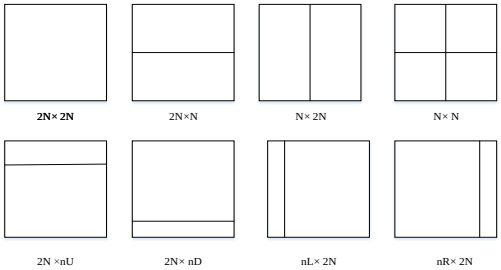
۲-۲- ساختار درختی استاندارد HEVC

استاندارد HEVC مشابه استانداردهای قبلی ابتدا یک تصویر را به بلوک‌های غیر هم‌پوشان به نام CTU[۸] تقسیم می‌نماید که اندازه CTU در آخرین نسخه نرم‌افزار مرجع HEVC یعنی HM-۱۶.۸ برابر ۶۴×۶۴ می‌باشد. CTU می‌تواند به واحدهای هم‌اندازه کوچک‌تر به نام بلوک‌های کد کننده تقسیم شود. در یک ساختار درختی چهارتایی که CTU به عنوان ریشه آن درخت در نظر گرفته می‌شود، اندازه بلوک‌های کد کننده می‌تواند از ۶۴×۶۴ تا ۸×۸ باشد. هر بلوک کد کننده ناحیه‌ای از تصویر را مشخص می‌کند که حالت پیش‌گویی آن‌ها یکسان است. برای هر بلوک کد کننده یک یا چند بلوک پیش‌گویی تعریف می‌شود که داخل یک بلوک پیش‌گویی، فرآیند پیش‌گویی یکسانی انجام می‌شود [۹]. استاندارد HEVC، ۸ حالت مختلف برای تقسیم شدن یک بلوک کد کننده به بلوک پیش‌گویی تعریف کرده است که در شکل ۳ نشان داده شده است.

HEVC دو حالت تقسیم شدن برای بلوک‌هایی که به صورت درون‌قابی کد می‌شوند تعریف کرده است و برای بلوک‌هایی که به صورت بین‌قابی کد می‌شوند، هشت حالت مختلف تقسیم شدن را تعریف کرده است که در شکل ۴، نشان داده شده است، همان‌طور که مشاهده می‌شود این هشت حالت به نام‌های ۲N×۲N، N×۲N، ۲N×N، N×N، ۲N×nU، ۲N×nD، nL×۲N و nR×۲N هستند که چهار حالت اول را حالت‌های متقارن می‌نامند زیرا در آن‌ها اندازه های بلوک‌های کدکننده یکسان است و چهار حالت بعدی را حالت‌های نامتقارن گویند زیرا در آن‌ها اندازه‌های بلوک‌های کد کننده یکسان نیست [۱۰].



شکل ۳: مثالی از شکسته شدن یک CTU به بلوک‌های کد کننده به همراه ساختار درختی چهارتایی آن [۹]



شکل ۴: چگونگی تقسیم‌بندی حالت‌های پیش‌گویی بین‌قابی در استاندرد HEVC [۱۰]

۳- پیشینه تحقیق

در طی سالیان اخیر به منظور طراحی الگوریتم‌هایی که منجر به کاهش پیچیدگی محاسباتی در استاندارد کد کننده‌ی ویدئویی پربازده می‌شوند، روش‌های زیادی ارائه شده است که در این بخش به معرفی برخی از این الگوریتم‌ها پرداخته می‌شود.

Saurty و همکاران [۱۱] به منظور کاهش تعداد ارزیابی بلوک‌های پیش‌گویی با اندازه مختلف روشی مبتنی بر حد آستانه ارائه دادند. در روش پیشنهادی آن‌ها ابتدا بلوک با اندازه ۲N×۲N بررسی شده و براساس مقدار باقی‌مانده‌ی اختلاف بلوک پیش‌گویی شده و بلوک تصویر اصلی تصمیم به ارزیابی سایر اندازه‌های بلوک می‌گیرد. در صورتی که میزان باقی‌مانده پیکسل‌ها بعد از پیش‌گویی از یک حد آستانه‌ای بیشتر باشد از بررسی بلوک‌ها با اندازه N×۲N، ۲N×N، ۲N×nU، ۲N×nD، nL×۲N و nR×۲N صرف نظر می‌نماید و بلوک ۲N×۲N به چهار بلوک با اندازه N×N شکسته می‌شود. در صورتی که میزان باقی‌مانده پیکسل‌ها بعد از پیش‌گویی از یک حد آستانه‌ای دیگری کمتر باشد روند بررسی متوقف می‌گردد و سایر حالت‌ها بررسی نمی‌گردد. نتایج شبیه‌سازی نشان داده که این الگوریتم توانسته حدود ۶۲/۲ درصد زمان کد کردن را بهبود بخشد و نرخ بیت به میزان ۱/۱۴ درصد افزایش داشته

است. عیب این روش این است که به حد آستانه‌ها وابستگی زیادی دارد و این حد آستانه‌ها مقدار ثابتی در نظر گرفته شده است.

در مرجع [۱۲]، به منظور کاهش پیچیدگی محاسباتی در استاندارد کد کننده‌ی ویدئویی پربازده یک الگوریتم تصمیم‌گیری برای پیش‌گویی درون‌قابی ارائه شده است که با محدود کردن تعداد حالت‌های پیش‌گویی منجر به افزایش سرعت کد کردن و در نتیجه کاهش پیچیدگی محاسباتی می‌گردد. در این روش بر اساس نوع تقسیم‌بندی حالت‌های پیش‌گویی و نیز پارامتر کوانتیزاسیون، یک حد آستانه تعیین می‌شود. سپس محدوده‌ی سطح خاکستری با حد آستانه مقایسه شده و در صورتی که محدوده‌ی سطح خاکستری بیش‌تر از حد آستانه باشد توسط فیلتر سوبل اطلاعات لبه‌ی تصویر محاسبه می‌شود. این اطلاعات بیان می‌کنند که بلوک کد کننده به چه تعداد و چه اندازه‌ای باید تقسیم گردد و از تقسیم شدن زیاد بلوک کد کننده که منجر به افزایش پیچیدگی محاسباتی می‌گردد، جلوگیری می‌نماید. نتایج شبیه‌سازی نشان داده که این الگوریتم توانسته حدود ۳۹/۵ درصد زمان کد کردن را بهبود بخشد و نرخ بیت به میزان ۱/۹۸ درصد افزایش داشته است.

در مرجع [۱۳]، یک چارچوب برای تصمیم‌گیری سریع برای تعیین بهترین حالت ارائه شده است. در این روش، در مرحله‌ی اول، برای شناسایی دقیق‌تر مناطق حرکت یکنواخت، مناطق با ساختار صاف، حداقل تغییر فریم برای تصمیم‌گیری سریع حالت‌های پیش‌گویی، ارتباط بین کاندیدها و بلوک‌های کد کننده بررسی می‌شود. سپس در مرحله‌ی دوم، یک مدل که مبتنی بر ارتباط بین کم‌ترین و بیش‌ترین تابع هزینه‌ی فریم میانی و جمع قدرمطلق تفاضل حالت‌ها می‌باشد، طراحی می‌شود. این مدل برای پیش‌گویی مقدار جمع قدرمطلق تفاضل که منجر به تصمیم‌گیری سریع بهترین حالت می‌شود، به کار برده می‌شود. نتایج شبیه‌سازی نشان داده که این الگوریتم توانسته حدود ۴۶ درصد زمان کد کردن را بهبود بخشد و نرخ بیت به میزان ۱/۹۷ درصد افزایش داشته است.

در مرجع [۱۴]، یک الگوریتم برای تصمیم‌گیری سریع برای بلوک کد کننده ارائه شده است که روی واحدهای پیش‌گویی درون‌قابی و بین‌قابی اعمال شده است که در آن با تجزیه و تحلیل مناطقی که دارای بافت همگن هستند، برای اندازه‌ی بلوک تصمیم‌گیری می‌کند. نتایج شبیه‌سازی نشان داده که این الگوریتم توانسته حدود ۲۵ درصد زمان کد کردن را بهبود بخشد و نرخ بیت به میزان ۰/۸ درصد افزایش داشته است.

Vanne و همکاران [۱۵]، به منظور کاهش پیچیدگی محاسباتی در استاندارد HEVC تعداد ارزیابی بلوک‌های پیش‌گویی با اندازه‌های مختلف را کاهش داده‌اند. در روش پیشنهادی، سه تکنیک ارائه شده است که در تکنیک اول ابتدا حالت‌های پرش و ۲N×۲N بررسی شده در صورتی که حالت پرش به عنوان بهترین حالت انتخاب شود از بررسی سایر حالت‌های متقارن و نامتقارن صرف نظر می‌گردد. در تکنیک دوم بازه‌ی اندازه‌های حالت‌های متقارن و نامتقارن را کاهش می‌دهند و در تکنیک سوم بازه‌ی اندازه بلوک‌های متقارن و نامتقارن را با توجه به مقدار پارامتر کوانتیزاسیون انتخاب می‌کنند. نتایج شبیه سازی آنها نشان داده است که حداکثر ۵۱ درصد زمان محاسبات را کاهش داده به طوری که ۱/۳ درصد نرخ بیت افزایش می یابد.

Hu و Yang [۱۶]، یک مدل برای کاهش پیچیدگی محاسباتی پیش‌گویی بین‌قابی در استاندارد HEVC ارائه دادند. در مدل پیشنهادی، ابتدا پیش‌گویی روی بلوک ۲N×۲N انجام شده و تبدیل DCT روی ضرائب باقی مانده اعمال می‌کنند تا از ضرائب تبدیل DCT میزان هم-بستگی مکانی-زمانی بلوک جاری را به دست آورند. سپس بر اساس هم‌بستگی مکانی-زمانی بین بلوک‌های پیش‌گویی، اندازه بلوک جاری جهت پیش‌گویی تخمین زده می‌شود و بلوک‌هایی که شانس کم‌تری جهت پیش‌گویی مناسب دارند حذف می‌گردند. آنها ادعا کردند که روش پیشنهادی می‌تواند حدود ۶۰ درصد زمان کد کردن را کاهش داده در حالی‌که نرخ بیت در کیفیت یکسان به میزان ۱/۰۶ درصد افزایش می‌یابد.

در این مقاله یک الگوریتم به منظور کاهش پیچیدگی محاسباتی در استاندارد کد کننده‌ی ویدئویی پربازده معرفی شده است که با کاهش تعداد حالت‌ها و تعداد بلوک‌های مورد ارزیابی در واحد پیش‌گویی بین-قابی منجر به کاهش پیچیدگی محاسباتی و در نتیجه‌ی آن کاهش زمان کد کردن می‌شود.

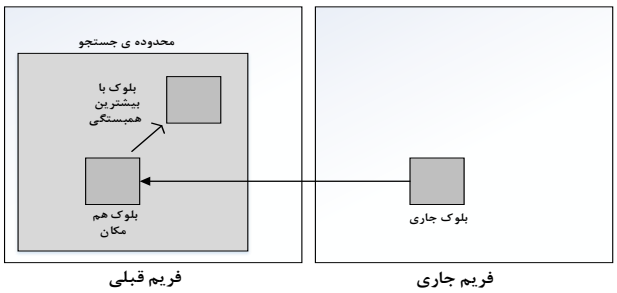
۴- روش پیشنهادی

یکی از قسمت‌هایی که در استاندارد HEVC به دلیل پیچیدگی محاسباتی بالا روش‌هایی برای بهینه‌سازی آن ارائه شده است، واحد پیش‌گویی بین‌قابی است. استاندارد HEVC به منظور افزایش کارایی پیش‌گویی بین‌قابی، تعداد بلوک‌های کد کننده مورد ارزیابی را افزایش داده است. در این استاندارد، ابتدا فریم به واحدهایی به نام بلوک کد کننده تقسیم شده و سپس هر بلوک توسط پیش‌گویی بین‌قابی یا درون-قابی پیش‌گویی می‌شود. واحد کد کننده برای هر بلوک، پیش‌گویی را بر اساس داده‌های کد شده قبلی از فریم جاری (پیش‌گویی درون‌قابی) و یا توسط داده‌های کد شده قبلی در فریم‌های دیگر (پیش‌گویی بین-قابی) به دست می‌آورد. در پیش‌گویی بین‌قابی، برای یک بلوک ۶۴×۶۴ پیکسلی تعداد اندازه‌های مختلف مورد پیش‌گویی از ۹ حالت در استاندارد H.264/AVC به ۳۵ حالت افزایش یافته است که موجب افزایش پیچیدگی محاسباتی آن نیز می‌شود به طوری که برای بلوک ۶۴×۶۴ پیکسلی ۶۹۷ بلوک با اندازه‌های مختلف (از ۶۴×۶۴ تا ۴×۴) پیش‌گویی می‌گردد و بهترین اندازه بلوک پیش‌گویی انتخاب می‌گردد، این پیچیدگی محاسباتی باعث می‌شود که این استاندارد برای کاربردهای بی‌درنگ مناسب نباشد. در روش پیشنهادی هدف این است که با استفاده از همبستگی مکانی و زمانی بین پیکسل‌ها تعداد بررسی بلوک‌ها با اندازه‌های مختلف را کاهش داده به طوری که برای کاربرده‌های بی‌درنگ مناسب گردد. در ادامه جزییات روش پیشنهادی بیان می‌شود.

۴-۱- الگوریتم تصمیم‌گیری بهترین اندازه بلوک پیش‌گویی

در استاندارد HEVC از پیش‌گویی بین‌قابی جهت حذف افزونگی زمانی استفاده می‌گردد و به جستجوی بهترین اندازه بلوک پیش‌گویی در یک محدوده‌ای از فریم قبلی می‌پردازد. بررسی اندازه‌های مختلف بلوک پیش‌گویی هزینه محاسباتی زیادی را تحمیل می‌کند که در روش پیشنهادی با حذف برخی از اندازه‌های بلوک پیش‌گویی که شانس کم-تری جهت انتخاب بهترین بلوک پیش‌گویی دارند به کاهش پیچیدگی محاسباتی کمک می‌نماییم. از آن‌جایی که همبسته‌ترین بلوک در فریم قبلی با بلوک جاری از لحاظ بافت شباهت بیش‌تری به هم دارند می‌توان از اطلاعات بلوک همبسته در فریم قبلی استفاده کرد تا بتوان برخی از اندازه‌های مورد ارزیابی بلوک پیش‌گویی را کاهش داد. همان‌طور که در شکل ۵، نشان داده شده است. در روش پیشنهادی ابتدا بلوک با بیشترین همبستگی در فریم قبل که به بلوک جاری شباهت دارد را به عنوان همبسته‌ترین بلوک تعریف می‌گردد. میزان شباهت بین دو بلوک از رابطه ۲ محاسبه می‌شود.

$$SAD = \sum_{i=1}^{2N\times 2N}\sum_{j=1} abs(M[i,j]-N[i,j]) \qquad (۲)$$



شکل ۵: انتخاب بلوک با بیشترین همبستگی نسبت به بلوک جاری

که در آن M[i,j] و N[i,j] شدت پیکسل‌ها به ترتیب در بلوک جاری و بلوک مورد مقایسه می‌باشد. بلوکی که کم‌ترین مقدار SAD[۹] را داشته باشد به عنوان همبسته‌ترین بلوکی که با بلوک جاری انطباق دارد انتخاب می‌شود. اگر در همبسته‌ترین بلوک حالت ۲N×۲N به عنوان بهترین بلوک پیش‌گویی انتخاب شده باشد، با توجه به شبیه بودن تصویر می‌توان نتیجه گرفت که بهترین حالت بلوک جاری نیز حالت ۲N×۲N خواهد بود و می‌توان از بررسی سایر حالت‌های ۲N×N یا N×۲N صرف‌نظر کرد. برای ارزیابی فرض فوق، احتمال انتخاب شدن حالت ۲N×۲N برای یک بلوک را در حالتی که همبسته‌ترین بلوک آن در فریم قبلی نیز دارای حالت ۲N×۲N باشد را محاسبه کرده و در جدول ۱ نشان داده شده است. برای محاسبه احتمال فوق از ۵ ویدئو با رزولوشن مختلف و برای شبیه‌سازی این ۵ ویدئو از آخرین نسخه نرم‌افزار مرجع HEVC یعنی HM-۱۶.۸ استفاده شده است. به منظور مقایسه دقیق با کارهای دیگران شرایط آزمایش با شرایط بیان شده توسط گروه JCT-VC مطابقت داده شد. بیش‌ترین اندازه‌ی بلوک پیش‌گویی ۶۴×۶۴ و کم‌ترین اندازه‌ی آن ۴×۴ در نظر گرفته شد. هم‌چنین پارامتر کوانتیزاسیون مقادیر ۲۲، ۲۷، ۳۲، ۳۷ تنظیم شد. چنان‌چه در جدول ۱ نشان داده شده است، احتمال فرض فوق در بیش از ۹۸ درصد حالات صحیح است، در نتیجه در روش پیشنهادی اگر همبسته‌ترین بلوک دارای حالت ۲N×۲N باشد، می‌توان از بررسی سایر حالت‌های ۲N×N یا N×۲N صرف نظر کرد.

زمانی که تصویر دارای بافت همسان باشد، مقدار پیکسل‌ها نزدیک به هم‌دیگر هستند و اختلاف بین مقدار پیکسل‌ها ناچیز است که می‌توان جهت بافت تصویر را بدین وسیله تشخیص داد. جهت بافت در یک تصویر می‌تواند عمودی و یا افقی باشد. حال برای محاسبه‌ی جهت بافت بلوک، همان‌طور که در شکل ۶ نشان داده شده است اختلاف میانگین پیکسل‌ها یک بار به صورت عمودی و بار دیگر به صورت افقی محاسبه می‌شود و سپس مجموع هر کدام به دست آورده می‌شود.

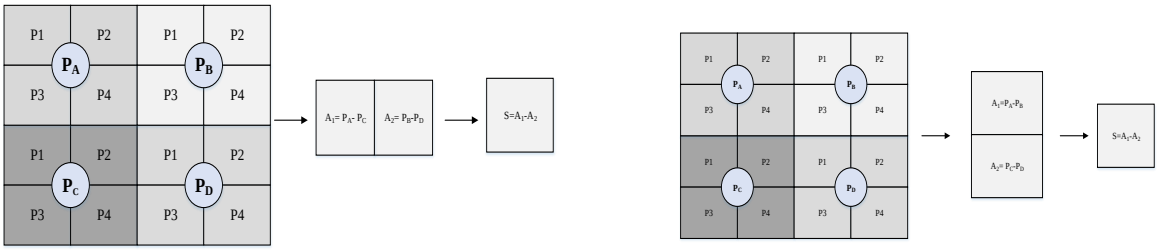
جدول ۱: احتمال انتخاب شدن حالت ۲N×۲N در زمانی که در همبسته‌ترین بلوک در فریم قبلی بلوکی با اندازه ۲N×۲N باشد

نام ویدئو	QP=۲۲	QP=۲۷	QP=۳۲	QP=۳۷
Basketballpass	٪۹۸	٪۹۹	٪۹۶	٪۹۸
RaceHorses	٪۹۹	٪۹۷	٪۹۲	٪۹۳
Blowingbubbles	٪۹۶	٪۹۰	٪۸۹	٪۹۲
Parkscene	٪۹۹	٪۹۳	٪۹۶	٪۹۵
Partyscene	٪۹۷	٪۹۸	٪۹۸	٪۹۱
Basketballdrill	٪۹۳	٪۸۷	٪۹۶	٪۹۵
Fourpeople	٪۸۹	٪۹۲	٪۹۷	٪۸۶
Kimono1	٪۸۷	٪۹۳	٪۹۵	٪۹۱
Peopleonstreet	٪۹۲	٪۸۹	٪۹۴	٪۹۵
Traffic	٪۹۳	٪۹۴	٪۸۸	٪۹۱

اگر بتوان با توجه به بافت تصویر تخمین زد که کدام یک از حالت-های شکسته شدن تصویر به‌صورت افقی یا عمودی برای پیش‌گویی آن ناحیه از تصویر مناسب است، می‌توان تا حدی از پیچیدگی محاسباتی کاست. برای به دست آوردن بافت یک بلوک پیش‌گویی از اختلاف بین پیکسل‌ها می‌توان استفاده کرد. زمانی که تصویر دارای بافت افقی باشد مقدار شدت پیکسل‌ها در جهت افقی نزدیک به هم هستند و در نتیجه اختلاف بین مقدار شدت پیکسل‌ها ناچیز خواهد بود. برای محاسبه بافت، در روش پیشنهادی مطابق شکل ۷، ابتدا بلوک به چهار بخش کوچک‌تر شکسته می‌شود و سپس با استفاده از رابطه‌ی ۳ میانگین هر قسمت محاسبه می‌گردد.

$$Pi = \frac{۱}{۴}\sum_{j=۱}^{۴} P\,j \quad ,i=\{A,B,C,D\} \qquad (۳)$$

در رابطه‌ی ۳، Pi میانگین مقدار پیکسل‌های هر قسمت می‌باشد.



الف) ب)

شکل ۶: مثالی از نحوه‌ی به دست آوردن اختلاف پیکسل‌ها: الف) به صورت عمودی، ب) به صورت افقی

پس از محاسبه میانگین هر قسمت اختلاف میانگین هر قسمت یک بار به صورت افقی و بار دیگر به صورت عمودی محاسبه می‌شود و هر حالت که اختلاف کم‌تری داشته باشد نشان دهنده این است که در آن جهت پیکسل‌ها به هم نزدیک تر هستند و دارای مقادیر مشابه‌تری هستند. به عبارت دیگر اگر اختلاف دو بلوک به صورت افقی کم‌تر از حالت عمودی باشد، حالت N×۲N جهت پیش‌گویی استفاده می‌شود. هم‌چنین اگر اختلاف دو بلوک به صورت عمودی کم‌تر از حالت افقی باشد حالت ۲N×N جهت پیش‌گویی استفاده می‌شود و از بررسی حالت N×۲N صرف‌نظر می‌شود.

برای ارزیابی فرض فوق احتمال انتخاب شدن حالت ۲N×N زمانی که اختلاف پیکسل‌ها در ناحیه عمودی کم‌تر می‌باشد، از رابطه‌ی ۴ بررسی می‌شود.

$$P_a = \frac{N_E}{N_T} \qquad (۴)$$

به‌طوری‌که در این رابطه N_E بیان‌گر تعداد بلوک‌های ۲N×N می‌باشد به شرطی که اختلاف عمودی کم‌تر از افقی باشد و N_T بیان‌گر تعداد کل بلوک‌های ۲N×N می‌باشد. هم‌چنین احتمال انتخاب شدن حالت N×۲N زمانی که اختلاف پیکسل‌ها در ناحیه افقی کم‌تر باشد، از رابطه‌ی ۵ بررسی می‌شود.

$$P_b = \frac{N_D}{N_T} \qquad (۵)$$

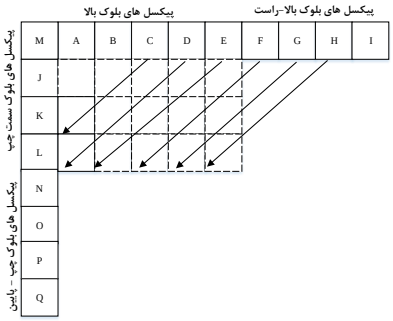
به‌طوری‌که در این رابطه N_D بیان‌گر تعداد بلوک‌های N×۲N می‌باشد به شرطی که اختلاف افقی کم‌تر از عمودی باشد و N_T بیان‌گر تعداد کل بلوک‌های N×۲N می‌باشد. برای محاسبه احتمالات فوق، شرایط آزمایش مطابق شرایط بیان شده در جدول ۱ است و نتایج به‌دست آمده در جدول ۲ آورده شده است. چنان‌چه در جدول ۲ نمایش داده شده است می‌توان نتیجه گرفت که فرض فوق در بیش از ۸۵ درصد حالت‌ها صحیح است. در نتیجه در روش پیشنهادی بهترین حالت برای پیش‌گویی در ناحیه‌ای از بلوک پیش‌گویی قرار دارد که جهت بافت در آن قسمت باشد. پس در این صورت زمانی که اختلاف میانگین پیکسل‌های دو بلوک مجاور در ناحیه افقی کم‌تر باشد فقط حالت N×۲N محاسبه می‌شود و در صورتی که میانگین پیکسل‌های دو بلوک مجاور در ناحیه عمودی کم‌تر باشد حالت 2N×N محاسبه می‌شود.

جدول ۲: احتمالات انتخاب حالت‌های N×۲N و ۲N×N در پیش‌گویی بین‌قابی روی ویدئوهای مختلف

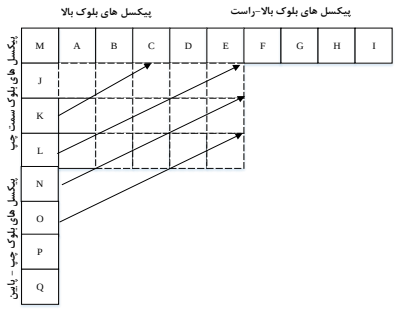
نام ویدئو	QP=۲۲		QP=۲۷		QP=۳۲		QP=۳۷	
	P_b	P_a	P_b	P_a	P_b	P_a	P_b	P_a
Basketballpass	٪۹۰	٪۹۲	٪۹۱	٪۹۳	٪۸۷	٪۹۶	٪۹۵	٪۹۲
RaceHorses	٪۹۳	٪۸۶	٪۸۸	٪۸۹	٪۹۲	٪۹۷	٪۸۶	٪۸۹
Blowingbubbles	٪۹۵	٪۹۱	٪۹۰	٪۸۷	٪۹۳	٪۹۵	٪۹۱	٪۹۱
Parkscene	٪۹۴	٪۹۵	٪۹۳	٪۹۲	٪۸۹	٪۹۴	٪۹۵	٪۹۶
Partyscene	٪۹۲	٪۹۱	٪۹۰	٪۹۳	٪۹۴	٪۸۸	٪۹۱	٪۹۸
Basketballdrill	٪۹۳	٪۸۷	٪۹۶	٪۹۵	٪۸۶	٪۸۸	٪۸۹	٪۹۲
Fourpeople	٪۸۹	٪۹۲	٪۹۷	٪۸۶	٪۹۱	٪۹۰	٪۸۷	٪۹۳
Kimono1	٪۸۷	٪۹۳	٪۹۵	٪۹۱	٪۹۵	٪۹۳	٪۹۲	٪۸۹
Peopleonstreet	٪۹۲	٪۸۹	٪۹۴	٪۹۵	٪۹۱	٪۹۰	٪۹۳	٪۹۴
Traffic	٪۹۳	٪۹۴	٪۸۸	٪۹۱	٪۸۷	٪۹۶	٪۹۵	٪۸۹

استاندارد HEVC پس از پیش‌گویی بین‌قابی از پیش‌گویی درون‌قابی نیز استفاده می‌کند تا بررسی کند که آن بلوک دارای همبستگی زمانی بیش‌تری است یا همبستگی مکانی آن بیش‌تر است که این کار باعث می‌شود که تعداد پیکسل‌های مورد نیاز برای ذخیره‌سازی کاهش

یابد. در پیش‌گویی درون‌قابی از پیکسل‌های مجاور کد شده در همان فریم استفاده می‌شود تا بلوک جاری را پیش‌گویی نماید. از ۳۵ جهت مختلف در پیش‌گویی درون‌قابی استقاده می‌گردد تا بهترین جهت برای پیش‌گویی آن بلوک به دست آید. این ۳۵ جهت را می‌توان به دو دسته‌ی افقی و عمودی تقسیم کرد (۱۷ حالت زاویه‌ای، یک حالت DC و یک Planar)، جهت‌های افقی از پیکسل‌های بلوک سمت چپ و پایین-چپ جهت پیش‌گویی استفاده می‌کنند. در حالی‌که جهت‌های عمودی از پیکسل‌های بلوک‌های بالا و بالا-راست جهت پیش‌گویی استفاده می‌نماید که در شکل ۷ نشان داده شده است.



الف)



ب)

شکل۷: نحوه‌ی پیش‌گویی توسط دو جهت عمودی و افقی : الف) پیش-گویی در جهت افقی، ب) پیش‌گویی در جهت عمودی

پیش‌گویی در جهت بافت تصویر بهترین پیش‌گویی را تولید خواهد کرد. می‌توان از بافت تصویر جهت کاهش تعداد جهت‌های مورد بررسی در پیش‌گویی درون‌قابی استفاده کرد. در روش پیشنهادی با توجه به بافت محاسبه شده در مرحله قبل، یکی از دسته‌ها جهت پیش‌گویی افقی یا عمودی برای پیش‌گویی یک بلوک استفاده خواهد شد. در صورتی که جهت بافت عمودی باشد، پیکسل‌های بلوک مورد ارزیابی در جهت عمودی بیش‌تر شبیه به هم‌دیگر هستند، در نتیجه فقط با حالت-هایی که در جهت عمودی قرار دارند پیش‌گویی درون‌قابی انجام می‌شود و تعداد حالت‌ها از ۳۵ حالت به ۱۹ حالت کاهش می‌یابد و در صورتی که جهت بافت افقی باشد، یعنی پیکسل‌ها در جهت افقی بیش‌تر شبیه به هم‌دیگر هستند، فقط با حالت‌هایی که در جهت افقی قرار دارند پیش‌گویی درون‌قابی انجام می‌شود. در این حالت نیز ۱۷ جهت افقی به همراه DC و Planar برای پیش‌گویی انتخاب می‌گردد. برای ارزیابی فرض فوق، احتمال انتخاب شدن جهت‌های عمودی برای پیش‌گویی درون‌قابی در زمانی که جهت بافت عمودی باشد، از رابطه‌ی ۶ بررسی می‌شود.

$$Pc = \frac{N_H}{N_T} \qquad (۶)$$

به طوری‌که N_H بیان‌گر تعداد جهت‌های عمودی به شرطی که جهت بافت تصویر عمودی باشد و N_T تعداد کل جهت‌ها می‌باشد.

هم‌چنین احتمال انتخاب شدن جهت‌های افقی برای پیش‌گویی درون‌قابی در زمانی که جهت بافت افقی باشد، از رابطه‌ی ۷ بررسی می‌شود.

$$Pd = \frac{Nv}{N_T} \qquad (۷)$$

به‌طوری‌که در این رابطه N_V بیان‌گر تعداد جهت‌های افقی به شرطی که جهت بافت تصویر افقی باشد و N_T تعداد کل جهت‌ها می‌باشد. برای محاسبه احتمالات فوق مطابق با شرایطی که پیش‌تر بیان گردیده عمل شده است و نتایج در جدول ۳ آورده شده است. چنان‌چه در جدول ۳ نمایش داده شده است می‌توان نتیجه گرفت که فرض فوق در اغلب حالت‌ها صحیح می‌باشد در نتیجه در روش پیشنهادی بهترین حالت برای پیش‌گویی درون‌قابی در صورتی که جهت بافت عمودی باشد در بین جهت‌های عمودی قرار دارد و در صورتی که جهت بافت افقی باشد در بین جهت‌های افقی قرار دارد.

در شکل ۸ فلوچارت روش پیشنهادی آورده شده است و به‌طور خلاصه در زیر مراحل روش پیشنهادی بیان شده است.

- یافتن بلوک با بیشترین هم‌بستگی در فریم مرجع.
- اگر حالت همبسته‌ترین بلوک ۲N×۲N است، فقط برای بلوک جاری حالت‌های ۲N×۲N بررسی می‌گردد.
- در غیر این‌صورت، بافت تصویر محاسبه می‌گردد (ابتدا با استفاده از رابطه‌ی ۳ میانگین هر بخش محاسبه شده و سپس اختلاف میانگین هر بخش به صورت عمودی و به صورت افقی محاسبه می‌شود و جهت با کمترین مقدار به عنوان بافت آن بلوک در نظر گرفته می‌شود).
- اگر بافت تصویر افقی باشد حالت N×۲N انتخاب می‌شود و از حالت ۲N×N صرف‌نظر می‌گردد. اما اگر بافت تصویر عمودی باشد حالت ۲N×N انتخاب می‌شود و از حالت N×۲N صرف‌نظر می‌گردد.
- اگر بافت تصویر افقی باشد جهت‌های افقی برای پیشگویی درون‌قابی انتخاب می‌گردد و از بررسی جهت‌های عمودی صرف‌نظر می‌شود. اما اگر بافت تصویر عمودی باشد جهت‌های عمودی انتخاب می‌گردد و از بررسی جهت‌های افقی صرف‌نظر می‌شود.

جدول ۳: احتمال انتخاب بهترین حالت در ناحیه افقی / عمودی بلوک پیش‌گویی زمانی که جهت بافت تصویر افقی/عمودی باشد

نام ویدئو	QP=۲۲		QP=۲۷		QP=۳۲		QP=۳۷	
	P_c	P_d	P_c	P_d	P_c	P_d	P_c	P_d
Basketballpass	٪۹۰	٪۹۲	٪۹۱	٪۹۵	٪۸۸	٪۹۱	٪۸۷	٪۹۲
RaceHorses	٪۹۱	٪۸۷	٪۸۹	٪۸۷	٪۹۱	٪۸۹	٪۹۲	٪۸۷
Blowingbubbles	٪۹۵	٪۹۲	٪۹۰	٪۹۲	٪۹۵	٪۹۰	٪۹۰	٪۹۲
Parkscene	٪۹۸	٪۹۶	٪۹۱	٪۹۸	٪۹۳	٪۹۱	٪۸۹	٪۹۶
Partyscene	٪۹۰	٪۹۱	٪۹۵	٪۹۱	٪۹۰	٪۹۵	٪۹۴	٪۹۱
Basketballdrill	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۹۰	٪۹۱	٪۹۵	٪۹۱
Fourpeople	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹
Kimono1	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹
Peopleonstreet	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹	٪۸۹
Traffic	٪۹۵	٪۸۷	٪۸۸	٪۹۳	٪۹۰	٪۹۱	٪۹۵	٪۸۹

۵- نتایج شبیه‌سازی

در این بخش نتایج شبیه‌سازی روش پیشنهادی بیان خواهد شد. در ابتدا پارامترهای مورد اندازه‌گیری تعریف می‌شود، و شرایط آزمایش بیان خواهد شد. سپس نتایج روش پیشنهادی با انجام آزمایش بر روی ۵ ویدئو با رزولوشن‌های مختلف که توسط گروه JCT-VC معرفی شده است بیان خواهد شد.

۵-۱- تنظیمات شبیه‌سازی

از محیط تست HM-۱۶.۸ جهت پیاده‌سازی روش پیشنهادی استفاده شده است. هر ویدئو چهار بار با مقادیر مختلف پارامتر کوانتیزاسیون ۲۲، ۲۷، ۳۲و ۳۷ تست می‌شوند که توسط گروه JCT-VC بیان شده است. بیش‌ترین اندازه بلوک پیش‌گویی ۶۴×۶۴ تنظیم می‌شود و کم‌ترین اندازه‌ی بلوک پیش‌گویی ۴×۴ در نظر گرفته می‌شود. هم‌چنین بیش-ترین اندازه بلوک تبدیل برابر ۳۲×۳۲ و کم‌ترین آن ۴×۴ تنظیم می‌شود. لازم به ذکر است این شرایط در تمام حالت‌های تست و هم‌چنین جهت مقایسه با کارهای دیگر یکسان تنظیم می‌گردد.

۵-۲- کارایی الگوریتم پیشنهادی

به منظور ارزیابی روش پیشنهادی و مقایسه با روش‌های اخیر از پارامتر-های زیر استفاده می‌گردد.

- BD-Rate: از آنجایی که در روش پیشنهادی دو پارامتر نرخ بیت و PSNR دچار تغییر می‌گردند مقایسه با روش‌های دیگر مشکل است. به همین دلیل باید یکی از پارامترها یکسان گردد و تغییرات پارامتر دیگر مقایسه شود. بدین منظور دو پارامتر BD-Rate و BD-PSNR تعریف می‌شوند. این پارامتر مخفف Bjontegard Delta rate نام دارد که میزان اختلاف نرخ بیت بین دو روش را تحت کیفیت ویدئو یکسان بیان می‌کند به‌طوری‌که مقادیر منفی BD-Rate نشان می‌دهد که نرخ بیت در کیفیت یکسان ویدئو کاهش یافته است و مقدار مثبت BD-Rate نشان می‌دهد که میزان نرخ بیت مورد نیاز افزایش یافته است. کیفیت ویدئو در ویدئوهای رنگی با پارامتری به نام PSNR که از مؤلفه‌های لوما و کروما مطابق با رابطه ۸ محاسبه می‌شود [۱۷].

$$PSNR = ۱۰ \cdot Log\left(\frac{۲۵۵^{۲}}{MSE}\right) \qquad (۸)$$

در حالی که MSE میانگین مربعات خطا است و به صورت زیر از مؤلفه‌های لوما و کروما به دست می‌آید که در آن MSE_Y، MSE_U، MSE_V به ترتیب میانگین مربعات خطای مؤلفه‌های Y، U، V می‌باشد. مقدار نرخ بیت و PSNR روش پیشنهادی و نسخه مرجع HEVC را به ازای چهار پارامتر کوانتیزاسیون ۲۲، ۲۷، ۳۲، ۳۷ به عنوان ورودی تابع BD-Rate داده می‌شود تا مقدار BD-Rate را به عنوان خروجی برگرداند واحد اندازه‌گیری آن کیلو بیت بر ثانیه می‌باشد [۱۸].

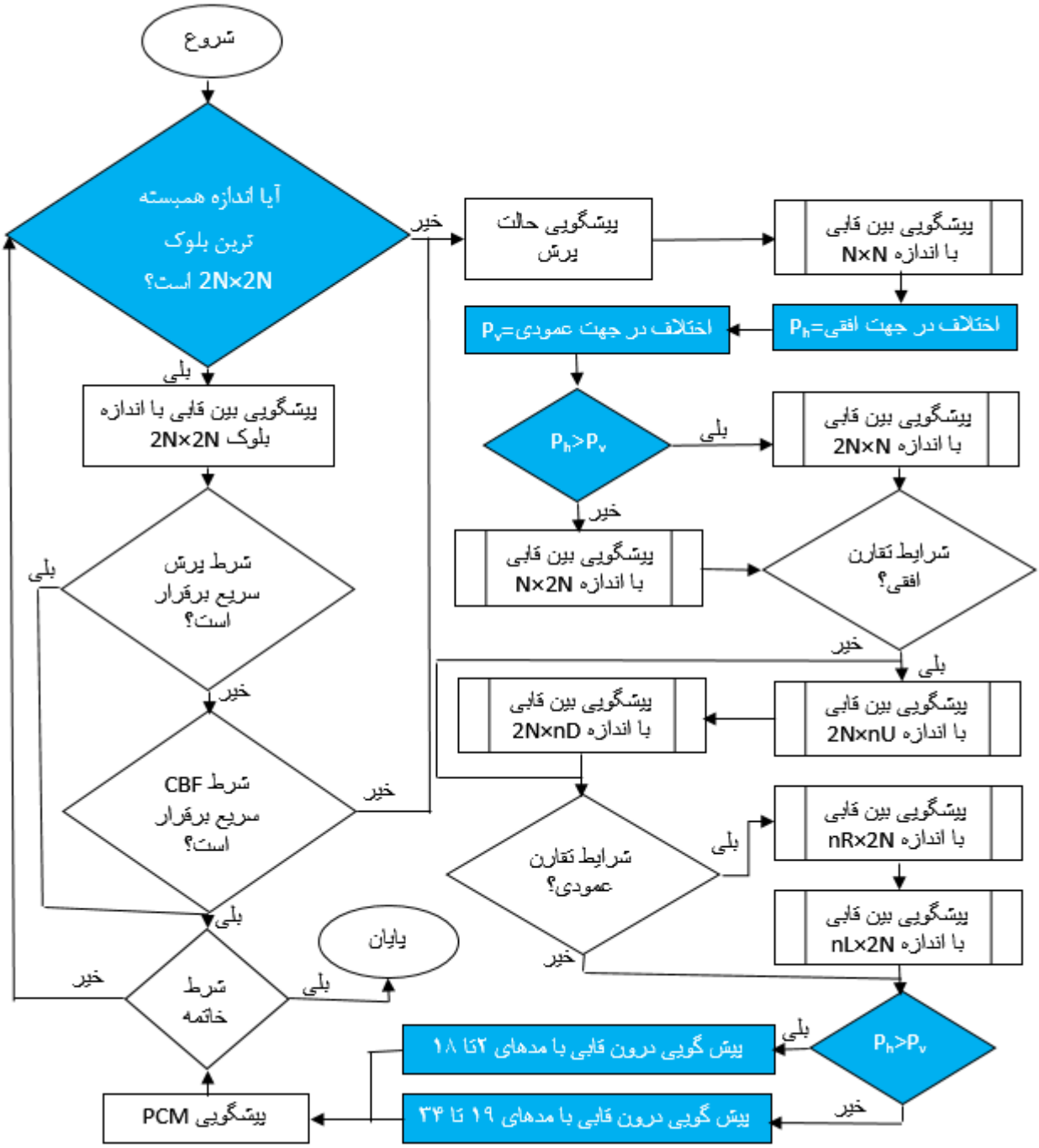
$$MSE = \frac{۶ \times MSE_y + MSE_u + MSE_v}{۸} \qquad (۹)$$

- BD-PSNR: مخفف کلمه Bjontegard Delta PSNR است که میزان کاهش یا افزایش کیفیت تصویر بین دو روش تحت نرخ بیت یکسان گزارش می‌دهد و واحد اندازه‌گیری آن دسی‌بل می-باشد. مقدار منفی BD-PSNR نشان می‌دهد که کیفیت روش دوم نسبت به اول در نرخ بیت ثابت کاهش یافته و مقدار مثبت آن نشان می‌دهد که در نرخ بیت ثابت روش دوم کیفیت بیش‌تری نسبت به روش اول دارد.

- کاهش زمان کد کردن: از پارامتر کاهش زمان کد کردن برای بررسی کاهش پیچیدگی استفاده می‌نماییم که از رابطه ۱۰ محاسبه می‌شود که در آن $Time_{proposed(QPi)}$ زمان کد کردن با روش

پیشنهادی به ازای پارامتر کوانتیزاسیون مختلف ۲۲، ۲۷، ۳۲و ۳۷ می باشد و $Time_{anchor(QPi)}$ زمان کد کردن با نسخه اصلی HEVC می‌باشد.

$$TS = \frac{1}{4}\left(\sum_{i=1}^{4}\frac{Time\,propposed(QP_i) - Time\,anchor(QP_i)}{Time\,anchor(QP_i)}\right)\times 100 \quad (10)$$



شکل ۸: طرح کلی روش پیشنهادی، مستطیل ها با دو خط عمودی برای اندازه بلوک پیشگویی ۶۴×۶۴ اجرا نمی شوند. اشکال آبی رنگ به فلوچارت اصلی استاندارد HEVC اضافه گردیده اند تا پیچیدگی محاسباتی آن کاهش یابد.

۵-۳- تجزیه و تحلیل نتایج شبیه‌سازی

در قسمت ابتدایی جدول ۴، تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی که مربوط به پیش‌گویی بین‌قابی است را در مقایسه با نرم‌افزار مرجع HM-۱۶.۸ نشان می‌دهد. در این جدول اعدادی که به صورت مثبت و منفی درج شده‌اند به ترتیب بیان‌گر افزایش و کاهش پارامترهای اندازه‌گیری شده هستند. این جدول نشان می‌دهد که روش پیشنهادی توانسته در مرحله‌ی اول زمان کدکردن را به صورت

میانگین ۲۴ درصد کاهش دهد. در حالی که نرخ بیت ۱/۲ درصد افزایش داشته است. بیش‌ترین کاهش زمان کد کردن مربوط به ویدئو Basketballpass می‌باشد که ۳۱/۱۵ درصد زمان کد کردن را کاهش می‌دهد و کم‌ترین کاهش زمان کد کردن مربوط به ویدئو Peopleonstreet می‌باشد که حدود ۲۳/۸ درصد زمان کد کردن را کاهش داده است. از آن‌جایی که هر چه ویدئو دارای جزییات زیادی نباشد روش پیشنهادی تعداد حالت‌های بیش‌تری را کاهش می‌دهد، روش پیشنهادی توانسته است در ویدئوهایی با جزییات کم‌تر پیچیدگی محاسباتی را بیش‌تر کاهش دهد. این کاهش به دست آمده در زمان کد کردن به دلیل این واقعیت است که روش پیشنهادی حالت‌های پیش‌گویی را که شانس کم‌تری برای انتخاب شدن به عنوان بهترین حالت دارند را حذف می‌نماید. در قسمت بعدی، تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی که مربوط به پیش‌گویی درون-قابی است را در مقایسه با نرم‌افزار مرجع HM-۱۶.۸ نشان داده شده است. این جدول نشان می‌دهد که روش پیشنهادی توانسته در مرحله‌ی دوم زمان کد کردن را به صورت میانگین ۳/۰۱ درصد کاهش دهد. در حالی که نرخ بیت تنها ۰/۲ درصد افزایش داشته است. بیش-ترین کاهش زمان کد کردن مربوط به ویدئو Blowingbubbles می-باشد که ۳/۹ درصد زمان کد کردن را کاهش می‌دهد و کم‌ترین کاهش زمان کد کردن مربوط به ویدئو PeopleOnStreet می‌باشد که حدود ۲/۱ درصد زمان کد کردن را کاهش داده است.

در انتهای جدول ۴، مقایسه‌ی تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی را نسبت به نرم‌افزار مرجع HM-۱۶.۸ نشان می-دهد. همان‌طور که در این جدول مشاهده می‌شود، بیش‌ترین زمان کد کردن مربوط به ویدئو Blowingbubbles می‌باشد که حدود ۳۲ درصد زمان کد کردن را کاهش داده است و کم‌ترین زمان کد کردن مربوط به ویدئو Peopleonstreet می‌باشد که حدود ۲۷ درصد زمان کد کردن را کاهش داده است. روش پیشنهادی در نهایت توانسته ۲۹/۵۲ درصد زمان کد کردن را کاهش دهد، در حالی که نرخ بیت ۰/۹ درصد افزایش پیدا کرده است و پارامتر BD-PSNR مقدار ۰/۰۹ دسی‌بل کاهش را نشان می‌دهد. با توجه به نتایج حاصله می‌توان بیان کرد روش پیشنهادی توانسته زمان کد کردن را به طور قابل ملاحظه-ای کاهش دهد. در حالی که نرخ بیت افزایش ناچیزی داشته است.

جدول ۴ : مقایسه‌ی تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی در مقایسه با نرم‌افزار مرجع HM-۱۶.۸

ویدئو	رزولوشن	مرحله اول روش پیشنهادي			مرحله دوم روش پیشنهادي			نتیجه نهایي		
		BD-Rate [%]	BD-PSNR	TS [%]	BD-Rate [%]	BD-PSNR	TS (%)	BD-Rate [%]	BD-PSNR	TS [%]
Basketballpass	۴۱۶×۲۴۰	۰/۹	-۰/۰۸	-۳۱/۱۵	۰/۳	-۰/۰۱	-۳/۶	۰/۷	-۰/۰۹	-۲۹/۰۱
RaceHorses	۴۱۶×۲۴۰	۱/۴	-۰/۱	-۲۹/۰۵	۰/۳	-۰/۰۳	-۳/۸	۱/۱	-۰/۱۲	-۳۰/۰۵
Blowingbubbles	۸۳۲×۴۸۰	۱/۳	-۰/۱۳	-۲۷/۵۵	۰/۲	-۰/۰۲	-۳/۹	۰/۹	-۰/۱۰	-۳۱/۲۳
Basketballdrill	۸۳۲× ۴۸۰	۱/۵	-۰/۰۵	-۲۸/۹۳	۰/۲	-۰/۰۱	-۳/۱	۱/۱	-۰/۰۸	-۲۹/۲
Fourpeople	۱۲۸۰×۷۲۰	۱/۱	-۰/۰۶	-۲۶/۸	۰/۱	-۰/۰۱	-۲/۹	۰/۵	-۰/۰۸	-۲۹/۲۵
Parkscene	۱۹۲×۱۰۸۰	۱/۲	-۰/۰۷	-۲۵/۸۴	۰/۳	-۰/۰۲	-۲/۶	۰/۷	-۰/۰۵	-۳۰/۸۵
Partyscene	۱۹۲×۱۰۸۰	۱/۷	-۰/۰۵	-۲۶/۰۵	۰/۴	-۰/۰۲	-۳/۰۱	۱/۵	-۰/۰۷	-۲۹/۵۵
Kimono1	۱۹۲×۱۰۸۰	۱/۱	-۰/۰۹	-۲۶/۴۶	۰/۲	-۰/۰۱	-۲/۴	۱/۲	-۰/۰۶	-۳۰/۵۲
Peopleonstreet	۲۵۶×۱۶۰۰	۰/۸	-۰/۰۷	-۲۳/۸۲	۰/۱	-۰/۰۱	-۲/۱	۱/۱	-۰/۰۷	-۲۵/۶۶
Traffic	۲۵۶×۱۶۰۰	۱/۳	-۰/۰۴	-۲۴/۶	۰/۲	-۰/۰۳	-۲/۶	۰/۸	-۰/۰۸	-۲۶/۵
میانگین		۱/۲	-۰/۰۸	-۲۷/۰۲	۰/۲	-۰/۰۱	-۳/۰۱	۰/۹	-۰/۰۹	-۲۸/۸۱

۵-۴- مقایسه زمان کد کردن و نرخ بیت به ازاء کیفیت یکسان تصویر در روش پیشنهادی نسبت به سایر روش‌ها

به دلیل این‌که روش‌های دیگر هر یک تحت نسخه‌های مختلف HEVC پیاده‌سازی شده‌اند یک مقایسه‌ی عادلانه اغلب مشکل است و این باعث می‌شود اختلاف کمی در گزارش زمان کد کردن و نرخ بیت پیش بیاید، جدول ۵، مقایسه‌ی تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی را نسبت به روش‌های مقایسه شده را نشان می‌دهد. از نظر پارامتر BD-Rate روش‌های پیشنهادی مراجع [۱۲، ۱۳] نتایج بدتری نسبت به روش پیشنهادی گزارش کرده‌اند در حالی- که مرجع [۱۴] مقدار BD-Rate بهتری نسبت به روش پیشنهادی به دست آورده است. از نظر پارامتر کاهش زمان کد کردن روش پیشنهادی مرجع [۱۴] کم‌تر از روش‌های دیگر می‌باشد. همان‌طور که مشاهده می‌شود بعضی از روش‌ها سرعت بیش‌تری را نشان داده‌اند اما پارامتر BD-Rate بیش‌تری دارند و بالعکس. برای محاسبه‌ی میزان بهبود روش پیشنهادی نسبت به روش‌های دیگر می‌بایستی از معیار شایستگی استفاده نمود. رابطه‌ی ۱۱، فرمول معیار شایستگی را نشان می‌دهد. همان‌طور که مشاهده می‌شود طبق این رابطه معیار شایستگی متوسط افزایش نرخ بیت به کاهش زمان کد کردن را گویند. پارامتر FOM یک شاخص خوبی برای مصالحه بین زمان کد کردن و نرخ بیت می باشد. به طوری‌که هر چه مقدار FOM کم‌تر باشد تعادل بهتری بین کاهش زمان کد کردن و افزایش نرخ بیت فراهم می‌کند. بنابراین هر روشی که میزان FOM کم‌تری داشته باشد بهتر است. جدول ۶، معیار شایستگی روش پیشنهادی و مقایسه با سه روش بیان شده را نشان می‌دهد. همان‌طور که در جدول ۶ مشاهده می‌شود، روش پیشنهادی دارای FOM کم‌تری است که این به این معناست که کارایی الگوریتم پیشنهادی بهتر از سه روش دیگر است.

$$FOM = \left|\frac{BD-Rate}{Ts}\right| \times ۱۰۰ \qquad (۱۱)$$

برای ارزیابی نسبت سیگنال به نویز از نمودار نرخ اعوجاج استفاده می‌شود که در آن محور عمودی نسبت سیگنال به نویز و محور افقی نرخ بیت را نشان می‌دهد. شکل ۹، نمودار نرخ اعوجاج ویدئوی FourPeople را نشان می‌دهد. همان‌طور که مشاهده می‌شود در این نمودار محور افقی، نرخ بیت را نشان می‌دهد و محور عمودی، PSNR را نشان می‌دهد. این نمودار بر اساس مقادیر مختلف پارامتر کوانتیزاسیون ۲۲، ۲۷، ۳۲و ۳۷ رسم شده است، همان‌طور که در نمودار مشاهده می‌شود نقطه‌ی ابتدایی منحنی مربوط به پارامتر کوانتیزاسیون ۲۲ می‌باشد و نقطه‌ی انتهایی در منحنی مربوط به پارامتر کوانتیزاسیون ۳۷ می‌باشد. همان‌طور که مشاهده می‌شود، نمودارهای نرخ-اعوجاج روش پیشنهادی بسیار مشابه با نرم‌افزار مرجع HM-۱۶.۸ شده است که نشان‌دهنده‌ی این است که کارایی نرخ تغییر زیادی نداشته است در حالی که زمان کد کردن در روش پیشنهادی حدود ۲۹ درصد کاهش پیدا کرده است.

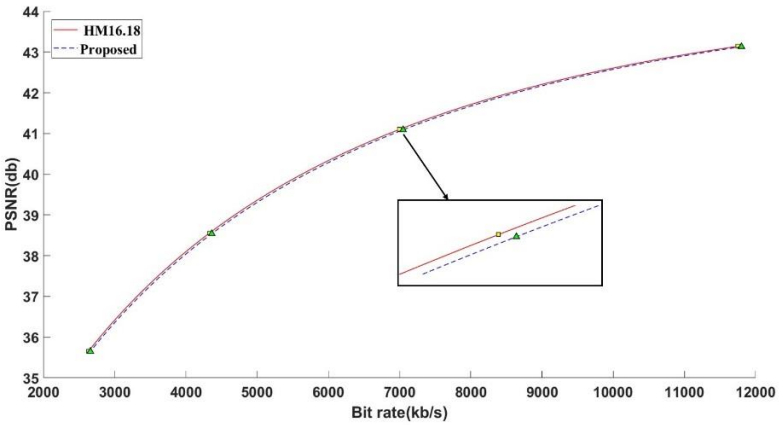
۶- نتیجه‌گیری

در این مقاله به منظور کاهش پیچیدگی محاسباتی در استاندارد کد کننده‌ی ویدئویی پربازده، یک روش مبتنی بر ویژگی‌های مکانی و زمانی معرفی شده است. این روش برای کاهش پیچیدگی محاسباتی از محدود کردن تعداد بلوک‌ها و حالت‌های پیش‌گویی استفاده می‌کند که این امر منجر به کاهش سربار محاسباتی روش پیشنهادی نسبت به استاندارد HEVC می‌گردد و کاهش سربار محاسباتی خود نیز باعث کاهش پیچیدگی محاسباتی و در نتیجه کاهش زمان کد کردن می- شود. روش پیشنهادی با استفاده از HM-۱۶.۸ پیاده‌سازی شده است. نتایج شبیه سازی‌ها روی ویدئوهای مختلف با بافت و رزولوشن‌های متفاوت نشان می‌دهد که روش پیشنهادی می‌تواند زمان کد کردن را حدود ۲۹ درصد کاهش دهد در حالی که نرخ بیت افزایش ناچیزی حدود ۰/۹ درصد دارد. مقایسه با کارهای مشابه نشان می‌دهد که روش پیشنهادی توانسته کارایی بهتری را نسبت به سایر روش‌های مورد مقایسه ارائه دهد.

نتایج شبیه‌سازی روش پیشنهادی نشان می‌دهد که در ویدئوهایی که دارای بافت ساده‌تری هستند نتایج بهتری به‌دست می‌آید. در صورتی- که بافت تصویر پیچیده باشد اختلاف شدت پیکسل‌ها در جهت افقی یا عمودی به هم نزدیک خواهد بود در نتیجه انتخاب حالت‌های افقی یا عمودی برای پیش‌گویی درون‌قابی سخت خواهد بود و ممکن است بهترین حالت انتخاب نشود. به عنوان کار آینده می‌توان اختلاف شدت پیکسل‌های تصویر را در جهت‌های مختلف ارزیابی کرد تا جهت دقیق‌تری برای حالت پیش‌گویی درون‌قابی ارائه داد و تعداد حالت‌های پیش‌گویی بیشتری کاهش یابد.

جدول ۵: مقایسه‌ی تفاوت زمان کد کردن و نرخ بیت روش پیشنهادی در مقایسه با سایر روش‌ها

نام ویدئو	الگوریتم پیشنهادی			Huang و همکاران [۱۲]			Tariq و همکاران [۱۳]			Fernandez و همکاران [۱۴]		
	BD-PSNR [%]	BD-Rate [%]	TS [%]	BD-PSNR [%]	BD-Rate [%]	TS [%]	BD-PSNR [%]	BD-Rate [%]	TS [%]	BD-PSNR [%]	BD-Rate [%]	TS [%]
Basketballpass	-۰/۰۹	۰/۷	-۲۹/۰۱	-۰/۱۱	۲/۱۰	-۳۹/۲	-۰/۰۸	۲/۱۵	-۳۳/۸۹	-۰/۱۲	۰/۱	-۱۰/۲۱
RaceHorses	-۰/۱۲	۱/۱	-۳۰/۰۵	-۰/۰۹	۱/۹۷	-۳۹/۱	-۰/۰۸	۲/۴۸	-۲۲/۷۷	-۰/۰۴	۰/۵	-۲۳/۸۰
Blowingbubbles	-۰/۱۰	۰/۹	-۳۱/۲۳	-۰/۰۸	۱/۷۸	-۳۹/۸	-۰/۱۴	۱./۱۸	-۳۴/۹۴	-۰/۰۲	۱/۶	-۳۲/۰۸
Parkscene	-۰/۰۸	۱/۱	-۲۹/۲	-۰/۱۲	۲/۳	-۳۸/۷	-۰/۱۷	۲/۶۵	-۳۸/۵۳	-۰/۰۹	۱	-۳۴/۶۶
Partyscene	-۰/۰۸	۰/۵	-۲۹/۲۵	-۰/۱۶	۱/۸۷	-۳۷/۹	-۰/۱۳	۱/۵۷	-۲۹/۵۷	-۰/۰۳	۱/۱	-۲۲/۱۶
Basketballdrill	-۰/۰۵	۰/۷	-۳۰/۸۵	-./۲۱	۱/۶۳	-۳۷/۶	-۰/۱۶	۱/۵۳	-۳۸/۱۲	-۰/۱۶	۰/۸۸	-۲۵/۳۶
Fourpeople	-۰/۰۷	۱/۵	-۲۹/۵۵	-./۱۷	۱/۴	-۳۸/۷	-۰/۰۷	۲/۳۴	-۳۹/۰۹	-۰/۰۸	۰/۷۵	-۲۷/۰۹
Kimono1	-۰/۰۶	۱/۲	-۳۰/۵۲	-۰/۱۴	۱/۸۵	-۳۹/۴	-۰/۰۵	۲/۱۳	-۳۹/۳	-۰/۰۹	۰/۲	-۲۴/۰۲
Peopleonstreet	-۰/۰۷	۱/۱	-۲۵/۶۶	-./۱۹	۱/۷	-۳۹/۲	-۰/۱	۲/۳	-۴۰/۰۳	-۰/۰۹	۰/۶	-۲۲/۳۶
Traffic	-۰/۰۸	۰/۸	-۲۶/۵	-./۱۴	۱/۹	-۳۸/۳	-۰/۰۸	۲/۰۶	-۳۹/۶	-۰/۰۶	۰/۹	-۲۵/۰۲
میانگین	-۰/۰۹	۰/۹	-۲۸/۸۱	-۰/۱۵	۱/۸	-۳۸/۹	-۰/۱۱	۲/۰۹	-۳۸/۲۲	-۰/۰۷	۰/۸۶	-۲۴/۵



شکل۹: مقایسه‌ی نرخ اعوجاج روش پیشنهادی و HM-۱۶.۸ برای ویدئو FourPeople

جدول ۶: مقایسه‌ی پارامتر معیار شایستگی روش‌های مختلف

پارامتر	الگوریتم پیشنهادی	Huang و همکاران [۱۲]	Tariq و همکاران [۱۳]	Fernandez و همکاران [۱۴]
FoM	۳/۱	۴/۶	۵/۴	۳/۵

مراجع

[1] J. Wang, L. Li, G. Zhi, Z. Zhang and H. Zhang, "Efficient algorithms for HEVC bitrate transcoding," *Multimedia Tools and Applications*, vol. 76, no. 24, pp. 26581–26601, 2017.

[2] I.-K. Kim, K. McCann, K. Sugimoto, B. Bross, W.-J. Han, and G. Sullivan, "High efficiency video coding (HEVC) test model 15 (HM15) encoder description," *In: Joint Collaborative Team on Video Coding (JCT-VC), Document of JCTVC-Q1002, 17th Meeting Val*encia, 2014.

[3] X. Wang, and Y. Xue , "Fast HEVC inter prediction algorithm based on spatio-temporal block information", *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pp. 1-5, 2017.

[4] C.E. Rhee, K. Lee, T.S. Kim, and H.J. Lee, "A Survey of Fast Mode Decision Algorithms for Inter-Prediction and Their Applications to High Efficiency Video Coding," *IEEE Transaction on Consumer Electronics*, vol. 58, no. 4, pp. 1375-1383, 2013.

[5] X. Li, X. He, X. Peng and S. Xiong, "An image feature-based method to efficiently determine inter-coding depth in HEVC," *AEU - International Journal of Electronics and Communications*, vol. 71, pp. 96-104, 2017.

[6] J. Xiong, H. Li , F. Meng, Q. Wu and K. N. Ngan , "Fast HEVC Inter CU Decision Based on Latent SAD Estimation," *IEEE Transactions on Multimedia*, vol. 17, no. 12, pp. 2147 - 2159, 2015.

[7] S.h. Jung and H. W. Park, "A fast mode decision method in HEVC using adaptive ordering of modes," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 10, pp. 1846 - 1858, 2016.

[8] M.J. Chen, Y.D. Wu, C.H. Yeh, K.M. Lin, and S.D. Lin, "Efficient CU and PU Decision Based on Motion Information for Inter-Prediction of HEVC," *IEEE Transactions on Industrial Informatics*, 2018.

[9] Z. Liu, T.-L. Lin and C.-C. Chou, "Efficient prediction of CU depth and PU mode for fast HEVC encoding using statistical analysis," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 474-486, 2016.

[10] H. Li, K. Fan, R. Wang, G. Li, and W. Wang, "A motion aided merge mode for HEVC", *IEEE Transaction on Circuits and Systems for Video Technology*, pp.28-50, 2018.

[11] K. Saurty, P.C. Catherine, and K.M. Soyjaudah, "Inter Prediction Complexity Reduction for HEVC Base on Residuals Characteristics", *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, pp. 1649-1668, 2016.

[12] X. Huang, Q. Zhang, X. Zhao, W. Zhang, Y. Zhang and Y. Gan, "Fast inter-prediction mode decision algorithm for HEVC," *Signal, Image and Video Processing*, vol. 11, no. 1, pp. 33-40, 2017.

[13] J. Tariq, S. Kwong and H. Yuan, "Spatial/temporal motion consistency based merge mode early decision for HEVC," *Journal of Visual Communication and Image Representation*, vol. 44, pp. 198-213, 2018.

[14] D.G. Fernandez, A.A. Del Barrio, G. Botteal, and R. Hermida, "Complexity Reduction in HEVC Standard Based on Smooth Region Classification," *Digital Signal Processing*, vol. 73, pp. 24-39, 2018.

[15] J. Vanne, M. Viitanen, D. Timmo, " Efficient Mode Decision Schemes for HEVC Inter Prediction". *IEEE Transaction on Circuits and Systems for video Technology*, vol. 24, no. 9, pp. 1579-1593, 2014.

[16] H. Nan, E. Yang, "Fast inter mode decision for HEVC based on transparent Composite model", *IEEE International Conference on image Processing*, 2015.

[۱۷] هادی شکری، محمدحسین کهایی، «حسگری فشرده تصاویر ابرطیفی با دسته‌بندی طیفی و بازسازی با تنظیم کننده تغییرات کلی طیفی-مکانی»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۷، شماره ۴، صفحه ۱۵۱۳-۱۵۲۱، ۱۳۹۶۰

[۱۸] سید محمد علوی، محسن بیات، «ارسال تصویر از طریق سیستم‌های چند ورودی-چند خروجی با تخصیص توان ارسالی نامتقارن»، *مجله مهندسی برق دانشگاه تبریز*، دوره ۴۵، شماره ۴، صفحه ۱۳۹-۱۵۲، ۱۳۹۴.

زیرنویس‌ها

[1] Advanced Video Coding (AVC)

[2] International Telecommunication Union Telecommunication sector

[3] Internation Organization for Standardization and the International Electrotechnical Commission

[4] High Definition

[5] High Efficiency Video Coding

[6] Merge Mode

[7] Skip Mode

[8] Coding Tree Unit

[9] Sum of Absolute difference