



ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

تخمین حرکت سریع بر اساس ویژگی محتوا برای کدگذار

H.265 / HEVC با پیچیدگی کم

چکیده

فیلم‌های با وضوح بالا (HD^۱) و فوق HD می‌توانند به طور گسترده در کاربردهای پخش همگانی مورد استفاده قرار گیرند. در ضمن با افزایش وضوح ویدئو حجم داده‌های اطلاعات HD بصری خام به میزان قابل توجهی افزایش می‌یابد و این مسأله تبدیل به چالشی برای ذخیره سازی، پردازش و انتقال داده‌های بصری HD شده است. استاندارد پیشرفته فشرده سازی ویدئو H.265 / HEVC داده‌های تصویری خام HD را به صورت موثر فشرده می‌کند اما نرخ فشرده‌سازی بالا به قیمت بار محاسباتی سنگین به دست می‌آید. از این رو کاهش پیچیدگی کدگذاری برای کدگذار H.265 / HEVC به منظور استفاده در برنامه‌های پخش همگانی حیاتی است. ما در این مقاله، با توجه به بهترین همبستگی انتخاب بردار حرکت در میان حالت‌های مختلف پیش‌بینی اندازه، یک روش تخمین حرکت (ME^۲) را برای کاهش پیچیدگی کدگذاری H.265 / HEVC پیشنهاد کردیم. ابتدا با توجه به نوع پارتیشن واحد پیش‌بینی (PU^۳)، همه PU ها به ترتیب به دو دسته PU والد و PU فرزند طبقه‌بندی می‌شوند. سپس بر اساس بهترین همبستگی انتخاب بردار حرکت بین PU والد و PU های فرزند، فرآیند جست و جوی انطباق بلوک PU های فرزند رها می‌شود، اگر PU والدشان نقطه جست و جوی اولیه را به عنوان بردار نهایی مطلوب حرکت خود در فرآیند ME انتخاب کند. نتایج آزمایشگاهی نشان می‌دهد که روش پیشنهادی ما در مقایسه با جست و جوی HM-TZ به طور متوسط ۲۰٪ صرفه‌جویی در زمان ME صرفه‌جویی می‌کند. در ضمن افت کارایی نرخ تخریب قابل چشم پوشی است.

کلمات کلیدی : فشرده‌سازی ویدئو، H.265 / HEVC تخمین حرکت سریع، واحد پیش‌بینی.

۱- معرفی

^۱ high definition

^۲ motion estimation

^۳ prediction unit

ویدئوهای با وضوح بالا (HD) و فوق HD در پاسخ به پیشرفت‌های تکنولوژی‌های ضبط و نمایش پدیدار شده‌اند که به‌طور گسترده در نظارت امنیتی، سیستم تلویزیون فوق HD و غیره مورد استفاده قرار می‌گیرند. با افزایش وضوح ویدیو، حجم داده‌های اطلاعات HD بصری خام به میزان قابل توجهی افزایش می‌یابد. از آنجایی که هنوز ظرفیت ذخیره‌سازی، محاسبات و انتقال محدود است این مسئله به یک چالش برای ذخیره‌سازی، پردازش و انتقال داده‌های بصری تبدیل شده است. از این رو سرعت فشرده‌سازی بالا و پیچیدگی کم ملزومات کلیدی برای فیلم‌های HD هستند که به‌طور گسترده در برنامه‌های پخش همگانی استفاده می‌شوند. اخیراً یک تیم مشترک در زمینه کدگذاری ویدیو (JCT-VC) تحت گروه کارشناسان کدگذاری ویدئو ITU-T (VCEG) و گروه کارشناسان فیلم متحرک (MPEG), ISO/IEC یک استاندارد فشرده‌سازی ویدیویی به نام H.265 / (HEVC) را تأسیس کرده‌اند [۲]، [۳]. در مقایسه با H.264/ (AVC) که استاندارد فشرده‌سازی ویدیویی نسل قبلی است [۴]، H.265 / HEVC دارای حدود ۵۰ درصد صرفه جویی در نرخ بیت با همان کیفیت بصری است. البته نرخ فشرده‌سازی بالاتر به قیمت پیچیدگی محاسباتی سنگین مجموعه‌ای از ابزارهای پیشرفته کدگذاری مانند ساختار چاردرخت^۴ مبتنی بر واحد کدگذاری، واحد پیش‌بینی نامتقارن و بزرگ میان فریم / درون فریم و غیره که در H.265 / HEVC استفاده می‌شود، به دست می‌آید. بنابراین کاهش پیچیدگی کدگذاری برای کدگذار H.265 / HEVC که به‌طور گسترده در برنامه‌های پخش همگانی مورد استفاده قرار می‌گیرد حیاتی است.

به منظور کاهش پیچیدگی کدگذاری H.265 / HEVC بسیاری از محققان بر روی بهینه‌سازی فرآیند پیش‌بینی بین فریمی H.265 / HEVC تمرکز کرده‌اند [۵] - [۱۰]. پان و دیگران [۵] روش تصمیم‌گیری سریع اندازه CU با استفاده از همبستگی انتخاب عمق چاردرخت CU بین CU فعلی و CU همسایه مکانی و زمانی آن پیشنهاد کرده‌اند. از آنجا که اندازه CU به محتوای CU بستگی زیادی دارد، شنت آل [۶] با استفاده از همبستگی بین اندازه CU و محتوای CU یک روش تصمیم‌گیری سریع CU پیشنهاد کرده است. ژانگ و همکاران [۷] پیشنهاد یک روش تصمیم‌گیری اندازه سریع CU مبتنی بر یادگیری ماشین ارائه داده‌اند که پیچیدگی کدگذاری را با افت کارایی نرخ

^۴ quadtree

تخریب قابل قبولی بهینه می‌کند. پانتا [۸] بر اساس فعالیت حرکت CU و همبستگی عمق سلسله مراتبی، تصمیم‌گیری زودهنگام ادغام / رها کردن را برای کاهش پیچیدگی فرآیند کدگذاری PU پیشنهاد کرده است. شنت [۹] با استفاده از همبستگی بین سطح CU و CU همسایه فضایی-زمانی، یک تصمیم‌گیری درون حالت انطباقی برای کاهش پیچیدگی H.265 / HEVC پیشنهاد کرده است. آنه و همکاران [۱۰] بر اساس پارامترهای کدگذاری زمانی - مکانی از کدکننده H.265 / HEVC یک روش کدگذاری سریع CU برای کدگذاری H.265 / HEVC پیشنهاد کرده‌اند. این روش‌ها می‌توانند به طور موثر پیچیدگی محاسباتی فرآیند پیش‌بینی بین فریم را کاهش دهند، با این حال پیچیدگی کدگذاری پیش‌بینی بین فریم عمدتاً از فرآیند تخمین حرکت (ME) حاصل می‌شود، پیچیدگی کدگذاری H.265 / HEVC را می‌توان با بهینه‌سازی فرآیند ME کاهش داد.

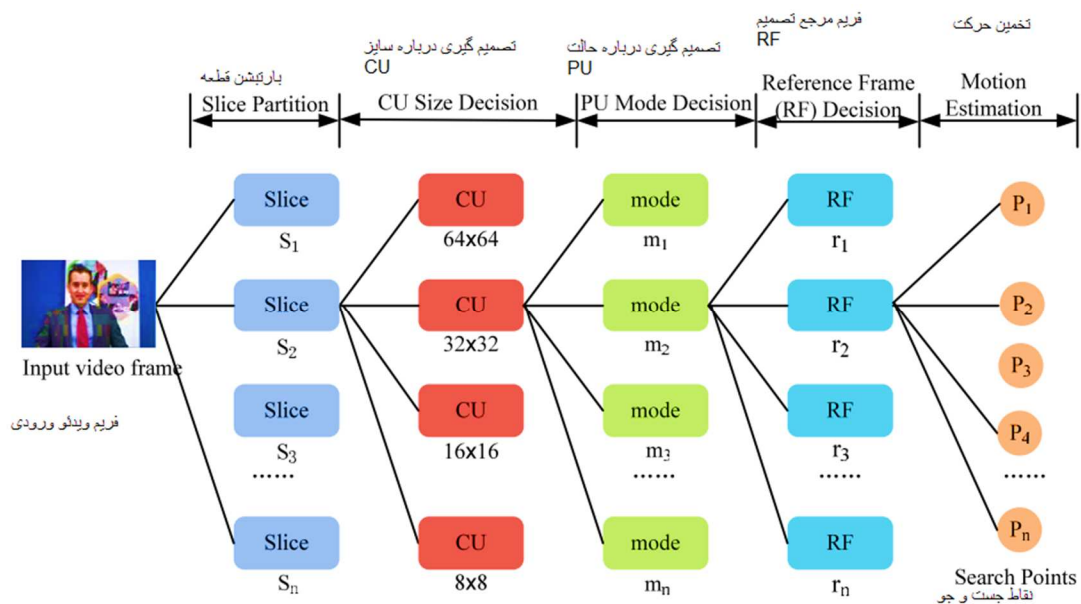
برای کاهش بیشتر پیچیدگی محاسباتی فرآیند ME، الگوریتم‌های سریع ME زیادی ارائه شده‌اند. مانند الگوریتم‌های سریع ME کلاسیک [۱۴] - [۱۱] که شامل جست و جوی سه مرحله‌ای [۱۱]، جست و جوی چهار مرحله‌ای [۱۲]، جست و جوی لوزی [۱۳]، جست و جوی شش گوشه [۱۴] و غیره است. بوست در [۱۵] الگوریتم جست و جوی چندگانه شش ضلعی شبکه‌های ترکیبی نامتقارن را پیشنهاد کرده است که مسئله کمینه محلی را به طور مؤثر در پردازش ME حل می‌کند. در [۱۶]، پان و دیگران [۱۶] با توجه به بهترین توزیع بردارهای حرکت یک الگوریتم سریع ME برای کاهش پیچیدگی محاسباتی فرآیند ME در کدگذار H.264 / AVC پیشنهاد کرده‌اند. در [۱۷] یک معیار برآورد کارایی و نابرابری مبتنی بر کدگذاری ویدئو چند رسانه‌ای با توجه به تصمیم‌گیری زودهنگام حالت، خاتمه زود هنگام ME/DE و کاهش محدوده جست و جوی انطباقی برای H.264 / AVC پیشنهاد شده است. با این حال، این الگوریتم‌های سریع ME به دلیل استفاده از ویدئوهای با وضوح بالا و تکنیک‌های کدگذاری ویدئویی جدید مانند چاردرخت مبتنی بر CU و PU های بزرگتر و نامتقارن، برای استفاده مستقیم در کدگذار H.265 / HEVC مناسب نیستند.

به منظور بررسی محدودیت‌ها و استفاده کامل از ویژگی‌های تکنیک‌های جدید کدگذاری، محققان تلاش‌های خود را بر طراحی الگوریتم‌های سریع ME برای H.265 / HEVC به کار گرفته‌اند [۱۸] - [۲۴]. در [۱۸] یک الگوریتم

سریع ME با تصحیح الگوی جست و جو و اتخاذ یک خاتمه زود هنگام برای جست و جوی TZ در نرم افزار مرجع H.265 / HEVC پیشنهاد شده است. با استفاده از جست و جوی شش ضلعی به جای TZ، روش ME سریع برای H.265 / HEVC توسعه داده شده است [۱۹]. بر اساس ویژگی محتوی CU های با اندازه های مختلف و بهترین توزیع MV ها، خاتمه زود هنگام برای جست و جوی TZ در نرم افزار مرجع H.265 / HEVC پیشنهاد شده است [۲۰]. در [۲۱] یک الگوریتم سریع ME با استفاده از چگالی حرکت بلوک پیشنهاد شده است که توسط MV و تفاوت MV بلوک های نزدیک ارزیابی شده است، که در صورت وجود چگالی حرکت بزرگ، جست و جوی TZ اصلی استفاده می شود و در غیر این صورت، جست و جوی شش گوشه پذیرفته می شود. یانگ و همکاران [۲۲] با در نظر گرفتن حرکات با احتمال بالا در جهت افقی و عمودی، ME سریع مبتنی بر جست و جوی جهت دار پیشنهاد کرده اند. در [۲۳] یک روش انتخاب جست و جوی پنجره ای برای تخمین حرکت سریع HEVC طراحی شده است که در آن پنجره جست و جو بر اساس پیش بینی کننده MV با اندازه 64×64 CU تعیین می شود. در [۲۴]، یک روش تصمیم گیری انطباقی برای ME پیشنهاد شده است، که در آن در ابتدا روابط بین اختلاف PU، طول MV پیشگویی شده و بهترین محدوده جست و جو آموزش داده می شود؛ سپس الگوریتم k-نزدیکترین همسایگی برای تعیین دامنه جست و جوی بهینه بر اساس نتایج یادگیری به کار گرفته می شود.

ضمن اینکه پیچیدگی کدگذاری ME می تواند بعدها با توجه به هزینه جست و جوی انطباق بلوک و استراتژی خاتمه زود هنگام و پیچیدگی محاسباتی مازاد، بهبود داده شود. علاوه بر این، بهترین همبستگی تصمیم گیری MV در حالت های مختلف PU با این الگوریتم ها در نظر گرفته نمی شود.

در این مقاله، ما یک روش ME سریع به منظور کاهش پیچیدگی H.265 / HEVC پیشنهاد می دهیم. ادامه این مقاله به شرح زیر است: بخش دوم، بررسی مختصری در مورد فرایند ME در H.265 / HEVC را ارائه می دهد. بخش سوم جزئیات روش پیشنهادی - روش ME سریع مبتنی بر شباهت محتوا - را ارائه می دهد. نتایج تجربی در بخش چهارم ارائه شده است. در نهایت بخش پنجم این مقاله را جمع بندی می کند.



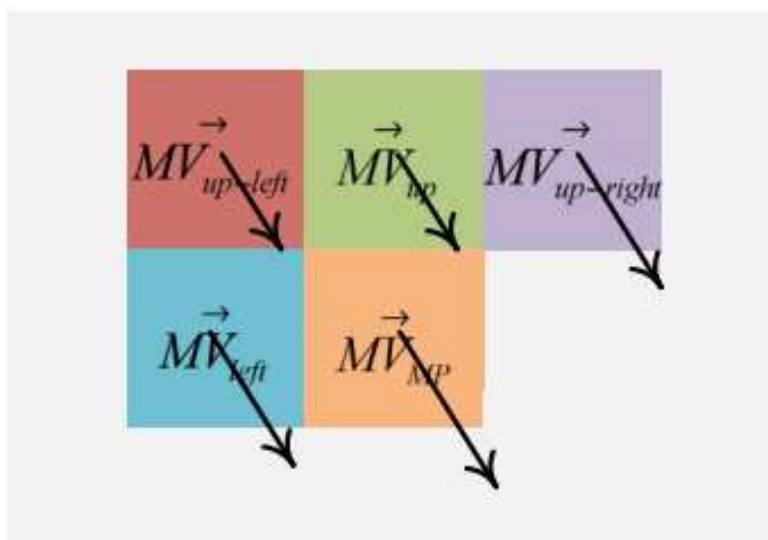
شکل ۱ - مراحل کد گذاری پیش بینی بین فریمی H.265 / HEVC

۲- مروری بر فرایند کد گذاری ME در H.265/HEVC

در فرایند کد گذاری H.265 / HEVC تمام تصاویر / فریم‌های ویدئویی به طور پیوسته کد گذاری می‌شوند. همانطور که در شکل ۱ نشان داده شده است، هر فریم به یک یا چند قطعه تقسیم می‌شود (S_1, S_2, \dots, S_n) در ادامه قطعات به دنباله‌ای از واحدهای درخت کد گذاری (CTUs) که واحد پردازش پایه کد کننده H.265 / HEVC هستند، تقسیم می‌شوند. CTU ها بر اساس یک چاردرخت به دنباله ای از CU ها تقسیم می‌شوند، که از CU های با سایز 64×64 تا 8×8 را پشتیبانی می‌کند. پس از آن بر اساس نوع پیش‌بینی، CU را می‌توان به یک، دو یا چهار PU تقسیم کرد که واحد پردازش پایه‌ای پیش بینی درون فریمی و بین فریمی است. برای H.265 / HEVC هشت حالت بین PU پشتیبانی می‌شوند. به عنوان مثال، $Inter_N \times 2N$ ، $Inter_2N \times N$ ، $Inter_2N \times 2N$ ، $Inter_N \times N$ ، $Inter_2N \times nU$ ، $Inter_2N \times nD$ ، $Inter_nL \times 2N$ و $Inter_nR \times 2N$. این حالت‌های بین PU به منظور حذف افزونگی‌های زمانی، فرایند ME را در فریم‌های مرجع انجام می‌دهند (به عنوان مثال، r_1, r_2, \dots, r_n) تا بهترین تطبیق بلوک را با توجه به کمینه‌سازی تابع هزینه لاگرانژی RD پیدا کنند [۲۵].

فرایند ME پیکسل صحیح سریع در H.265 / HEVC متشکل از دو مرحله، یعنی تصمیم‌گیری نقطه اولیه جست و جو (ISP)، و جست و جوی تطبیق بلوک است. برای پروسه تصمیم‌گیری ISP، پیش‌بینی کننده میانگین [۱۵] به عنوان یکی از نامزدهای ISP برای پیش‌بینی ISP استفاده می‌شود، که توسط پیش‌بینی کننده میانگین، $M\vec{V}_{MP}$ ، به دست آمده است، به صورت زیر نشان داده می‌شود.

$$M\vec{V}_{MP} = \text{Median}(M\vec{V}_{left}, M\vec{V}_{up}, M\vec{V}_{up-right}), \quad (1)$$



شکل ۲ - یک مثال از بلوک‌های نزدیک به بلوک فعلی

این از مقدار متوسط MV ی سمت چپ، بالا و بلوک سمت بالای بلوک فعلی به عنوان ISP استفاده می‌کند. شکل ۲ بلوک‌های نزدیک در سمت چپ، بالا و سمت راست بالای CU ی فعلی را نشان می‌دهد و MV های آنها به ترتیب به صورت $M\vec{V}_{left}$ ، $M\vec{V}_{up}$ و $M\vec{V}_{up-right}$ نشان داده می‌شود. نامزد دیگر ISP $(0,0)$ است. در نهایت، بهترین ISP یعنی $M\vec{V}_{ISP}$ با توجه به کمینه سازی تابع هزینه عملکرد لاگرانژی تعیین می‌شود که به صورت زیر است.

$$\begin{cases} \vec{MV}_{ISP} = \arg \min_{\vec{MV}_n} J(\vec{MV}_n, \lambda_{MOTION}), s.t. \vec{MV}_n \in S, \\ J(\vec{MV}_n, \lambda_{MOTION}) = SAD(o, c(\vec{MV}_n)) \\ \quad + \lambda_{MOTION} \cdot R(\vec{MV}_n - \vec{PMV}), \\ SAD(o, c(\vec{MV})) \\ \quad = \sum_{x=1}^{p_x} \sum_{y=1}^{p_y} |o(x, y) - c(x - MV_x, y - MV_y)|, \end{cases} \quad (2)$$

که \vec{MV}_n ، \vec{MV} ی ISP ی نامزد است.

$$S = \{\vec{MV}_{MP}, (0, 0)\}$$

SAD مجموع اختلاف مطلق بین PU اصلی و PU پیش بینی شده آن در موقعیت مشخص شده توسط \vec{MV}_n در فریم مرجع است. λ_{MOTION} ضرب کننده لاگرانژ است. \vec{PMV} بردار پیش بینی شده است. $R(\vec{MV}_n - \vec{PMV})$ تعداد بیتها برای کدگذاری اطلاعات حرکت را نشان می دهد که توسط یک جدول جست و جو به دست آمده است. p_x, p_y سایز PU و مساوی با ۶۴، ۳۲، ۱۶ یا ۸ است.

پس از تعیین بهترین ISP ، یک پنجره جست و جو با محدوده جست و جوی \bar{FW} در فریم مرجع ساخته می شود. سپس، یک جست و جوی تطبیق بلوک در پنجره جست و جو انجام می شود تا بهترین نقطه جست و جو با MV ، MV_{best} ، را با توجه به کمینه سازی تابع هزینه RD لانگرانژی، که در رابطه ۲ نشان داده شده است، تعیین کند.

$$\begin{aligned} \vec{MV}_{best} \\ = \arg \min_{\vec{MV}_n} J(\vec{MV}_c, \lambda_{MOTION}), s.t. \vec{MV}_c \in (\vec{MV}_m \cup \vec{MV}_{ISP}), \end{aligned} \quad (3)$$

که \vec{MV}_c نشان دهنده MV های همه نقاط جست و جوی نامزد است. \vec{MV}_m به معنای MV های نامزدهای نقاط جست و جو در پنجره جست و جوست. \vec{MV}_{ISP} نشان دهنده MV ی ISP است. از لحاظ تئوری، پیچیدگی محاسباتی جست و جوی تطبیق بلوک مبتنی بر ME به اندازه پنجره جست و جو و استراتژی جست و جو بستگی دارد. تعداد نقاط جست و جوی نامزد در جست و جوی تطبیق بلوک با افزایش اندازه پنجره جست و جو و پیچیدگی

استراتژی جست و جو افزایش می یابد. بنابراین، اگر جست و جوی انطباق ساده سازی یا نادیده گرفته شود، تعداد نامزدهای نقاط جست و جو به طور قابل توجهی کاهش می یابد، که باعث صرفه جویی بیشتر در پیچیدگی محاسباتی می شود.

جدول ۱ - شرایط آزمایش

ماکزیمم سایز CU	64 × 64
ماکزیمم عمق CU	4
پارامتر کوانتیزاسیون (QP)	22,27,32,37
روش ME	TZSearch
بازه جست و جو	[-64,64] کم-تاخیر - اصلی (LDM)
پروفایل کدگذاری	دسترسی تصادفی اصلی (RAM)

۳- ME ی با پیچیدگی کم مبتنی بر ویژگی محتوا برای H.265/HEVC

A- تجزیه و تحلیل پیچیدگی کدگذاری در ME ی H.265 / HEVC

پیچیدگی کدگذاری ME در H.265 / HEVC بستگی به تعداد کل نقاط جست و جو در فرآیند تطبیق بلوک دارد، فرض کنید یک CU با اندازه $2N \times 2N$ پیکسل است و محدوده جست و جو ME در هر دو جهت افقی و عمودی FW است، $(2W + 1)^2$ نقطه نامزد جست و جو در داخل پنجره جست و جو وجود دارد. از این رو، اگر تعداد نقاط جست و جو کاهش یابد، می توان در پیچیدگی محاسباتی ME می تواند صرفه جویی کرد. جست و جو ی TZ یک روش ME سریع در نرم افزار مرجع H.265 / HEVC است، و این باعث می شود که تعداد نقاط جست و جو ME با ساده تر شدن فرآیند جست و جوی تطبیق بلوک کاهش یابد. با این حال، با استفاده از ابزارهای پیشرفته کدگذاری مانند CU مبتنی بر ساختار چاردرخت، حالت‌های بزرگتر و نامتقارن PU، پیچیدگی کدگذاری ME سریع بر اساس جست و جو ی TZ هنوز هم بالا است. برای تجزیه و تحلیل پیچیدگی کدگذاری فرایند ME در H.265 / HEVC / نرم افزار مرجع H.265 / HEVC یعنی HM12.0 استفاده می شود. پنج توالی ویدیوی استاندارد

HEVC (Traffic), "Cactus", "Johnny", "BasketballDrill", "BQSquare") تست شده است، ویژگی این پنج دنباله تست و تنظیمات HM12.0 در [۲۶] تعریف شده است، وضوح این پنج توالی به ترتیب 416×240 ، 832×480 ، 1280×720 ، 1920×1080 و 1600×2560 است. تعداد فریم‌های کد شده برای این پنج دنباله به ترتیب ۱۶۱، ۱۲۹، ۹۷، ۶۵ و ۳۳ است. در میان این توالی‌های ویدئویی، "BQSquare" دارای حرکت متوسط است؛ "BasketballDrill" سریع حرکت می‌کند؛ "Johnny" دارای حرکت آهسته و پس زمینه ساده است؛ "Cactus" دارای پس زمینه پیچیده است و اجسام سریع حرکت می‌کنند؛ "Traffic" محتوای پیچیده‌ای دارد و اجسام دارای حرکت متوسط هستند. شرایط آزمایش در جدول ۱ آورده شده است. سایر پارامترهای کدگذاری، تنظیمات پیش فرض در HM12.0 و پروفایل‌های کدگذاری را پذیرفته‌اند. در این مقاله، پیچیدگی کدگذاری $ME_{complexity}$ ، به صورت زیر محاسبه می‌شود.

$$ME_{complexity} = (T_{ME}/T_{encoder}) \times 100\%, \quad (4)$$

که T_{ME} زمان کدگذاری مصرف شده توسط فرآیند ME را نشان می‌دهد؛ $T_{encoder}$ به معنی زمان کل کدگذاری HM12.0 است. نتایج آماری در جدول ۲ ذکر شده است.

از جدول ۲ می‌توان مشاهده کرد که پیچیدگی کدگذاری ME با سایزهای متغیر برای CU و PU بسیار زمانبر است، برای پروفایل کدگذاری (RAM)، $ME_{complexity}$ از ۵۱٫۳۹٪ تا ۷۵٫۳۳٪، به طور میانگین ۶۷٫۸۲٪ و برای LDM از ۵۵٫۰۵٪ تا ۸۳٫۱۵٪ و به طور میانگین ۷۳٫۷۸٪ است. علاوه بر این، می‌توانیم ببینیم که با افزایش مقدار پارامتر رقمی سازی (QP)، مقدار $ME_{complexity}$ بزرگ می‌شود. این به این دلیل است که وقتی QP افزایش می‌یابد، زمان کل کدگذاری کاهش می‌یابد، در حالی که تغییر در زمان ME نسبت به این مقدار کم است.

جدول ۲ - نتایج آماری از پیچیدگی کدگذاری ME در H.265/HEVC

پروفایل	دنباله	QP=22	QP=27	QP=32	QP37	میانگین
RAM	BQSquare	51.39	62.77	69.74	73.72	64.34
	BasketballDrill	59.97	65.97	71.54	74.61	68.02

	Johnny	n/a	n/a	n/a	n/a	n/a
	Cactus	58.80	68.45	72.70	74.75	68.68
	Traffic	62.55	69.44	73.61	75.33	70.23
	میانگین	58.18	66.66	71.83	74.60	67.82
LDM	BQSquare	55.05	66.50	73.70	78.84	68.52
	BasketballDrill	67.77	71.30	76.55	79.57	73.05
	Johnny	74.23	80.03	82.44	83.15	79.96
	Cactus	63.04	72.94	77.97	80.42	73.59
	Traffic	n/a	n/a	n/a	n/a	n/a
	میانگین	64.27	72.69	77.67	80.50	73.78

همچنین می‌توانیم ببینیم که برای توالی‌های ویدیویی مختلف، مقادیر $ME_{complexity}$ متفاوت هستند، این به این دلیل است که فعالیت حرکت و بافت محتوای ویدئویی متفاوت است. معمولاً محتوای ویدئویی با بافت پیچیده و فعالیت حرکتی سریع نیاز به زمان بیشتری برای رمزگذاری ME دارد، در عوض، مقدار $ME_{complexity}$ کوچک است. از این مقادیر، می‌توانیم دریابیم که پیچیدگی کدگذاری ME در کدکننده H.265 / HEVC بسیار بالا است، که جای زیادی برای بهینه‌سازی دارد. بنابراین، اگر بتوان فرآیند تطبیق بلوک ME را رها کرد، زمان زیادی برای رمزگذاری ذخیره می‌شود.

B. روش پیشنهادی ME سریع بر پایه شباهت محتوا

(۱) تصمیم‌گیری نقطه جست و جوی اولیه: در فرآیند کدگذاری ویدیو، به دلیل ثبات شی، همبستگی MV بزرگ بین CU فعلی و CU های همسایه فضایی آن، از قبیل CU های بالا، سمت چپ و بالا سمت راست، وجود دارد. روشهای ME جست و جوی انطباق بلوک بر اساس این فرض هستند که خطای تطبیق بلوک در هنگام نزدیک شدن به نقطه جست و جوی بهینه جهانی یکنواخت می‌شود. از این رو، یک ISP دقیق‌تر منجر به پیچیدگی کدگذاری کم در پیدا کردن نقطه جست و جوی بهینه جهانی خواهد شد. در این مقاله پیش‌بینی‌کننده میانگین \vec{MV}_{MP} و $(0,0)$ نیز به عنوان نامزدهای ISP پذیرفته شده‌اند. در نهایت \vec{MV}_{ISP} برای ISP ، مقدار زیر است.

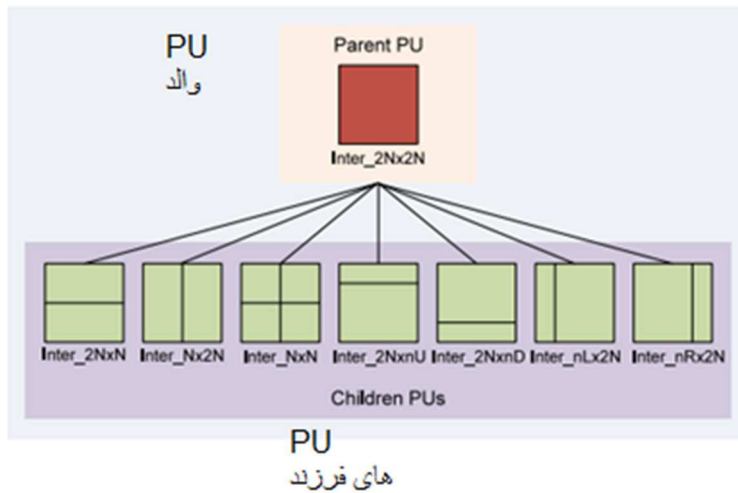
$$\begin{aligned}
 & \bar{M}\bar{V}_{ISP} \\
 & = \arg \min_{\bar{M}\bar{V}_n} J(\bar{M}\bar{V}_n, \lambda_{MOTION}), s.t. \bar{M}\bar{V}_n \in \{ \bar{M}\bar{V}_{MP}, (0,0) \}.
 \end{aligned}
 \tag{5}$$

به منظور ارزیابی کارایی ISP، فرض کنیم رویداد A نشان دهنده این است که نقطه نهایی جست و جوی بهینه در فرآیند ME با ISP یکسان است. احتمال P (A) با پنج توالی ویدئویی آزمون HEVC مورد تجزیه و تحلیل قرار گرفته ، شرایط آزمون در جدول ۱ ذکر شده است. نتایج آماری P (A) در جدول ۳ لیست شده است. از جدول ۳، می توان دید که بیشتر PU ها ISP را به عنوان نقطه نهایی جست و جوی مطلوب جهانی خود در فرآیند ME انتخاب می کنند. برای مشخصات کدگذاری RAM، از ۶۱،۰۶٪ تا ۷۵،۲۶٪، به طور متوسط ۶۹،۲۷٪، PU ها ISP را به عنوان نقطه نهایی جست و جوی جهانی خود انتخاب می کنند. برای مشخصات کدگذاری LDM، از ۵۵،۹۰٪ تا ۸۵،۳۰٪، ۶۸،۰۱٪ به طور متوسط، PU ها ISP را به عنوان نهایی خود انتخاب می کنند. علاوه بر این، می توانیم ببینیم که احتمال P (A) برای توالی ویدیویی «Johnny» برای پروفایل کدگذاری LDM به طور متوسط بیش از ۸۰،۵۰٪ است. این به این دلیل است که «Johnny» بیشتر دارای زمینه استاتیک است و این مناطق همبستگی فضایی قوی دارند. از این مقادیر می توان نتیجه گرفت که روش تصمیم گیری پیشنهاد شده ISP برای یافتن نقطه نهایی مطلوب جهانی با پیچیدگی کم در فرآیند ME به طور موثر کار می کند

جدول ۳- نتایج آماری P (A) (%)

پروفایل	دنباله	QP=22	QP=27	QP=32	QP37	میانگین
RAM	BQSquare	61.06	62.51	63.82	65.28	63.17
	BasketballIDrill	67.97	73.46	73.46	75.43	71.91
	Johnny	n/a	n/a	n/a	n/a	n/a
	Cactus	63.73	69.46	72.85	75.26	70.33

	Traffic	69.7 3	71.1 0	72.62	73.2 3	71.67
	میانگین	65.6 2	68.4 7	70.69	72.3 0	69.27
LDM	BQSquare	55.9 0	57.5 2	59.09	60.2 3	58.19
	Basketball rill	61.4 6	63.7 2	66.84	70.6 2	65.66
	Johnny	75.5 0	79.2 3	81.97	85.3 0	80.50
	Cactus	59.9 6	66.2 1	70.15	74.5 2	67.71
	Traffic	n/a	n/a	n/a	n/a	n/a
	میانگین	63.2 1	66.6 7	6.51	72.6 7	68.01



شکل ۳- نمایش PU والد و PU های فرزند

۲) استراتژی مبتنی بر شباهت محتوا با دور زدن جست و جوی انطباق بلوک: در فرایند رمزگذاری PU در H.265 / HEVC، CU ها در ادامه با توجه به نوع پیش بینی PU به PU های کوچکتر تقسیم می شوند. به عنوان مثال یک CU را می توان به یک PU با اندازه $N \times 2N^2$ تقسیم کرد، دو PU با اندازه $N \times N^2$ ، $N \times 2N$ ، $2N \times nD$ ، $2N \times nU$ ، $nL \times 2N$ یا $nR \times 2N$ یا چهار PUs با اندازه $N \times N$ در فرآیند ME لازم است همه ی inter PU ها ($Inter_2N \times 2N$ ، $Inter_2N \times N$ ، $Inter_N \times 2N$ ، $Inter_N \times N$ ، $Inter_2N \times nU$ ، $Inter_2N \times nD$ ، $Inter_2N \times nL$ و $Inter_nR \times 2N$) جست و جوی تطبیق بلوک در فریم مرجع را برای

پیدا کردن تطبیق بلوک مطلوب انجام دهند. بنابراین، با توجه به اندازه PU، حالت $2N \times 2N$ Inter می‌تواند به عنوان والد سایر حالت های PU در نظر گرفته شود به دلیل اینکه محتوای PU فرزند بخشی از محتوی PU والدشان شان است. نمونه ای از PU والد و PU فرزند در شکل ۳ نشان داده شده است. در فرآیند ME، اگر PU والد با ISP را به عنوان نقطه جست و جو بهینه نهایی خود انتخاب کنند، نشان دهنده آن است که محتوای PU والد با اندازه $2N \times 2N$ یک ناحیه استاتیک که یا دارای بافت ساده است و یا یکنواخت حرکت می‌کند. در این شرایط، PU های فرزند نیز ممکن است احتمال زیادی را برای انتخاب ISP به عنوان بهترین نقطه جست و جو خود داشته باشند. به دلیل همبستگی فضایی بالا و ویژگی های مشابه پیکسل های درون یک CU.

جدول ۴ - نتایج آماری احتمال شرطی (C | B) (%)

پروفایل	دنباله	QP=22	QP=27	QP=32	QP37	میانگین
RAM	BQSquare	93.97	94.22	94.47	94.84	94.38
	BasketballDrill	91.90	92.16	92.95	93.92	92.73
	Johnny	n/a	n/a	n/a	n/a	n/a
	Cactus	89.67	91.34	92.69	93.88	91.90
	Traffic	93.59	94.08	94.60	94.84	94.28
	میانگین	92.28	92.95	93.68	94.37	93.32
LDM	BQSquare	97.16	97.12	97.01	97.04	97.08
	BasketballDrill	91.33	91.71	92.77	94.26	92.52
	Johnny	94.42	95.45	96.44	97.27	95.90
	Cactus	88.32	90.19	91.53	93.27	90.83
	Traffic	n/a	n/a	n/a	n/a	n/a
	میانگین	92.81	93.62	94.44	95.46	94.08

به منظور تجزیه و تحلیل همبستگی تصمیم‌گیری نقطه مطلوب نهایی بین PU والد و PU های فرزند آن، رویداد B نشان دهنده این است که PU والد ISP را به عنوان بهترین نقطه نهایی خود انتخاب می‌کند و رویداد C نشان می‌دهد که PU های فرزند ISP را انتخاب می‌کنند. پنج دنباله ویدئویی تحت شرایط مندرج در جدول تست شده‌اند. نتایج آماری احتمال شرطی $P(C | B)$ در جدول ۴ آورده شده‌اند.

از جدول ۴، می‌توان دید که برای کدگذاری RAM، احتمال $P(C|B)$ از $۸۹,۶۷\%$ تا $۹۴,۸۴\%$ ، به طور متوسط $۹۳,۳۲\%$ است. برای کدگذاری LDM، احتمال $P(C|B)$ از $۸۸,۳۲\%$ تا $۹۷,۲۷\%$ و به طور میانگین $۹۴,۰۸\%$ است. علاوه بر این، می‌توانیم ببینیم که برای دنباله ویدئویی «Cactus» احتمال $P(C|B)$ کمی کاهش یافته است. به ترتیب $۹۱/۹۱$ و $۸۳/۹۰$ درصد برای کدگذاری RAM و LDM. این به این دلیل است که محتوای "Cactus" بافت پیچیده دارد که منجر به کاهش همبستگی مکانی MV فضایی می‌شود. علاوه بر این احتمال $P(C|B)$ با افزایش QP افزایش می‌یابد، دلیل این امر آن است که وقتی QP افزایش می‌یابد، اعوجاج‌ها تبدیل شده و به مقادیر کوچک کوانتیزه می‌شوند که منجر به همبستگی MV بزرگ فضایی میان PU والد و PU های فرزند می‌شود. از این مقادیر می‌توان نتیجه گرفت که اگر PU والد ISP را به عنوان بهترین نقطه انتخاب کنند، PU های فرزند نیز احتمال بالایی برای انتخاب ISP به عنوان نقطه نهایی مطلوب خود دارند. در نتیجه، جست و جوی تطبیق بلوک برای PU های فرزند می‌تواند رها شود و صرفه جویی در قابل توجهی در زمان کدگذاری به دست آید.

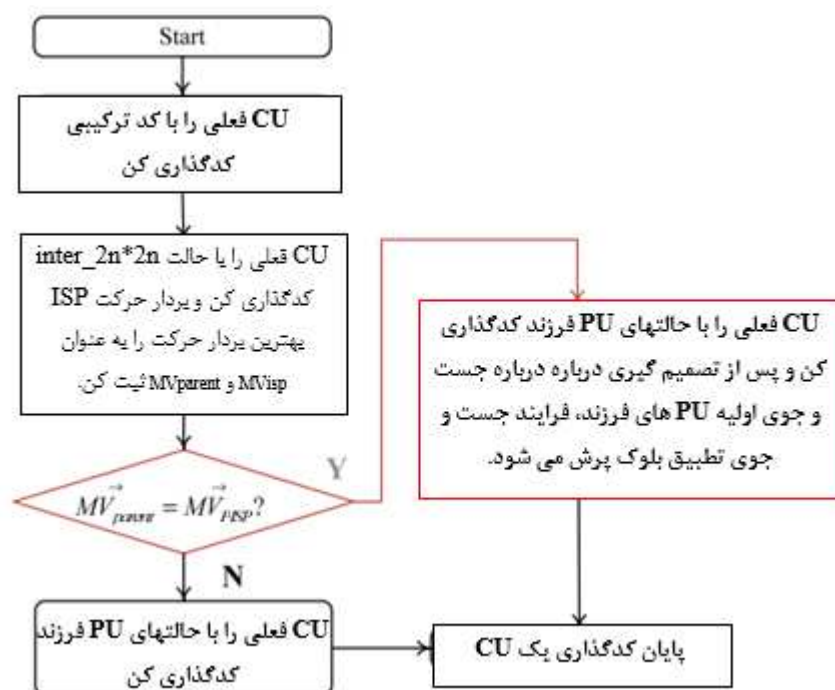
بر اساس تجزیه و تحلیل بالا، نقطه نهایی مطلوب جست و جوی PU های فرزند $\vec{MV}_{children}$ توسط معادله (۶) تعیین می‌شود. همانطور که در پایین این صفحه نشان داده شده است جایی که \vec{MV}_{CISP} MV ISP ی PU های فرزند است. \vec{MV}_{parent} نشان دهنده MV نقطه نهایی مطلوب جست و جوی PU والد است. \vec{MV}_{PISP} نشان دهنده MV ی ISP ی PU والد است. Ω مجموعه ای از MV های نامزد نقاط جست و جوی تطبیق بلوک برای PU های فرزند است.

۳) روش کلی: بر اساس تجزیه و تحلیل های فوق، روش سریع ME پیشنهادی به صورت گام به گام به شرح زیر خلاصه شده است، همچنین در شکل ۴ نشان داده شده است.

مرحله ۱: CU فعلی را با حالت Merge / Skip، حالت PU والد، حالت $2N \times 2N$ Inter_2N کد گذاری کنید و

سپس ISP و بهترین نقطه جست و جوی نهایی حالت PU والد را به ترتیب به صورتهای \vec{MV}_{PISP} و

\vec{MV}_{parent} ثبت کنید. اگر $\vec{MV}_{parent} = \vec{MV}_{PISP}$ ، به مرحله ۲ بروید. در غیر این صورت، به مرحله ۳ بروید.



شکل ۴ - فلوچارت روش ME سریع ارائه شده

مرحله ۲: CU فعلی را با حالت های PU فرزند، از جمله $Inter_N \times N$, $Inter_N \times 2N$, $Inter_2N \times N$, $Inter_2N \times nD$, $Inter_2N \times nU$ و $Inter_nL \times 2N$ کدگذاری کنید. پس از تصمیم گیری نقطه اولیه جست و جو برای این حالت های PU فرزند، فرآیند جست و جوی مربوط به بلوک برای این حالت ها را می شود و نقطه جست و جوی اولیه به عنوان نقطه جست و جوی مطلوب نهایی تعیین می شود. به مرحله ۴ بروید

مرحله ۳: CU فعلی را با حالت های PU، از جمله $Inter_N \times N$, $Inter_N \times 2N$, $Inter_2N \times N$, $Inter_2N \times nD$, $Inter_2N \times nU$ و $Inter_nL \times 2N$ رمزگذاری کنید. به مرحله ۴ بروید.

مرحله ۴: حالتها بین PU برای کدگذاری CU فعلی استفاده می شود. به مرحله ۵ بروید؛

مرحله ۵: اطلاعات کدگذاری را ذخیره کنید و جریان بیت کد گذاری شده را بنویسید. برای پردازش CU بعدی به مرحله ۱ بروید.

۴- نتایج آزمایشگاهی

A. تجزیه و تحلیل پیچیدگی رمزگذاری بر روی روش پیشنهادی

از روش کلی، می توانیم ببینیم که صرفه جویی در پیچیدگی محاسباتی روش پیشنهادی سریع ME به تعداد PU والد که ISP را به عنوان نقطه جست و جو مطلوب نهایی انتخاب می کنند بستگی دارد. اگر بیشتر PU ها ISP را به عنوان بهترین نقطه نهایی انتخاب کنند، صرفه جویی در پیچیدگی محاسباتی روش پیشنهادی سریع ME بزرگ خواهد بود. در غیر این صورت، صرفه جویی در پیچیدگی محاسباتی ME کوچک خواهد بود.

$$MV_{children} = \begin{cases} MV_{CISP}, & \text{if } MV_{parent} = MV_{PISP}, \\ \arg \min_{MV_n} J(MV_n, \lambda_{MOTION}), & \text{s.t. } MV_n \in (\Omega \cup MV_{CISP}), \text{ otherwise,} \end{cases}$$

جدول ۵ درصد ISP انتخاب شده به عنوان نقاط بهینه نهایی برای PU والد (٪).

پروفایل	دنباله	QP=22	QP=27	QP=32	QP37	میانگین
RAM	BQSquare	88.97	89.81	90.51	91.16	90.11
	BasketballDrill	84.89	86.86	88.95	90.56	87.82
	Johnny	n/a	n/a	n/a	n/a	n/a
	Cactus	82.42	86.11	88.76	90.46	89.95
	Traffic	89.09	89.81	90.25	90.64	89.95
	میانگین	86.34	88.15	89.62	90.75	88.72
LDM	BQSquare	96.09	96.20	96.29	96.17	96.18
	BasketballDrill	85.63	87.56	89.61	91.70	88.63
	Johnny	80.58	93.48	94.66	96.06	94.05
	Cactus	88.58	84.39	87.03	89.06	85.27
	Traffic	n/a	n/a	n/a	n/a	n/a
	میانگین	88.58	90.42	91.89	93.25	91.03

برای تجزیه و تحلیل مقادیر نظری تعداد PU های فرزند که جست و جوی انطباق بلوک را رها می کنند ، پنج توالی ویدیو در کدگذاری واقعی ویدئو مورد آزمایش قرار می گیرند. شرایط آزمون در جدول ۱ ذکر شده است و درصد PU های والد که ISP را به عنوان نقطه جست و جوی مطلوب نهایی خود انتخاب می کنند در جدول ۵ دیده می شود.

از جدول ۵ می توان دید که بیشتر PU های والد ISP را به عنوان نقطه نهایی مطلوب جست و جو در فرآیند ME انتخاب می کنند. برای مشخصات کدگذاری RAM، از ۸۲٫۴۲٪ تا ۹۱٫۱۶٪، به طور متوسط ۸۸٫۷۲٪، PU های والد هستند که ISP را به عنوان بهترین نقطه نهایی جست و جو ی خود انتخاب می کنند. برای مشخصات کدگذاری LDM، حدود ۸۰٫۵۸٪ تا ۹۶٫۲۶٪، به طور متوسط ۹۱٫۰۳٪، PU والد ISP را انتخاب به عنوان نقطه نهایی مطلوب خود انتخاب کردند. به عبارتی، حدود ۸۸٫۷۲٪ و ۹۱٫۰۳٪ از PU های فرزند به ترتیب فرایند جست و جوی تطبیق بلوک در پروسه ME برای کد گذاری RAM و LDM را رها کرده اند. از این مقادیر می توان نتیجه گرفت که روش پیشنهادی سریع ME می تواند به طور موثر پیچیدگی محاسباتی پردازش ME کد کننده H.265 / HEVC را کاهش دهد.

B. عملکرد رمزگذاری در PSNR، BR، صرفه جویی در زمان و صرفه جویی در کل کدگذاری

برای اثبات کارایی روش پیشنهادی سریع ME، محیط آزمایشی به شرح زیر تعریف شده است: پلت فرم نرم افزار مرجع HEVC HM12.0؛ پلت فرم سخت افزاری اینتل Xeon CPU E5-1620 v2 3.70GHz RAM 16.0GB با سیستم عامل ویندوز هفت ۶۴ بیتی است. شرایط آزمون در جدول ارائه شده است. پارامترهای دیگر کدگذاری تنظیمات پیش فرض در پروفایل HM12.0 و کدگذاری را می پذیرد.

ما روش پیشنهادی سریع الگوریتم ME را با دو که روش اخیرا منتشر شده، ETZSearch [۲۰] و ISOCC [۲۸] از لحاظ حداکثر نسبت سیگنال به نویز (PSNR)، نرخ بیت (BR)، صرفه جویی در زمان کل کدگذاری (TETS) و صرفه جویی در زمان ME (METS) مقایسه کرده ایم. معیار HM12.0 است، و نتایج مقایسه در جدول ۴ آمده است که $BDPSNR^{\text{H}}$ و $BDBR^{\text{H}}$ به ترتیب میانگین تفاوت PSNR به dB برای BR یکسان و میانگین تفاوت BR به درصد برای PSNR یکسان است و بر اساس [۲۹] محاسبه می شوند؛ TETS و METS به صورت زیر محاسبه می شوند :

^H jontegaard delta PSNR

^F Bjontegaard delta BR

$$\begin{cases} TETS = \frac{1}{4} \sum_{i=1}^4 \frac{T_{\psi}(QP_i) - T_{HM}(QP_i)}{T_{HM}(QP_i)} \times 100\%, \\ METS = \frac{1}{4} \sum_{i=1}^4 \frac{TE_{\psi}(QP_i) - TE_{HM}(QP_i)}{TE_{HM}(QP_i)} \times 100\%, \end{cases} \quad (7)$$

که $T_{\psi}(QP_i)$ کل زمان رمزگذاری HM12.0 را با روش ME سریع ψ تحت مقدار QP برای QP_i مشخص می کند $T_{HM}(QP_i)$ {شده پیشنهاد، [۲۸] ISOCC، [20] ETZSearch} $\psi \in$ زمان کل رمزگذاری HM12.0 با روش اصلی ME است، TZSearch، با مقدار QP برای QP_i است.

$T_{E_{\psi}}(QP_i)$ نشان دهنده کل ME زمان HM12.0 با روش ME سریع ψ با مقدار QP برای QP_i است.

$$\psi \in \{ETZSearch [20], ISOCC [28], \text{پیشنهاد}\}$$

$TE_{HM}(QP_i)$ نشان دهنده کل ME زمان HM12.0 با روش اصلی ME سریع، TZSearch، با مقادیر QP برای $QP_i = \{22, 27, 32, 37\}$ ، الگوریتم اصلی سریع ME، TZSearch، که در HM12.0 استفاده می شود، با HM-TZSearch مشخص می شود.

از جدول ۴، می توان مشاهده کرد که برای پروفایل کد گذاری LDM، ETZSearch کل زمان رمزگذاری را از ۱،۱۸٪ تا ۵،۴۸٪ و به طور متوسط ۲،۹۵٪ کاهش می یابد و موجب صرفه جویی در زمان ME از ۱،۴۰٪ تا ۸،۱۵٪، به طور متوسط ۴،۳۰٪ می شود. در همین حال، BDPSNR بین ETZSearch و HMTZSearch اصلی از -۰،۰۰۹ dB تا -۰،۰۰۱ dB و به طور متوسط -0.004 dB است و BDBR بین ETZSearch و HM-TZSearch اصلی از ۰،۰۱٪ تا ۰،۲۹٪ و به طور متوسط ۰،۱۲٪ است. برای مشخصات کدگذاری RAM، ETZSearch کل زمان کدگذاری از ۱،۴۳٪ تا ۴،۸۶٪، به طور متوسط ۲،۷۹٪ صرفه جویی می کند؛ و زمان ME را از ۲،۱۹٪ تا ۷،۰۲٪ و به طور متوسط ۴،۲۵٪ کاهش می دهد.

در حالی که BDPSNR بین ETZSearch و اصلی HM-TZSearch از -۰،۰۱۴ dB تا -۰،۰۰۱ dB به طور متوسط، -0.005 dB و BDBR بین ETZSearch و اصلی HM-TZSearch از ۰،۰۰۶٪ تا ۰،۳۹٪، به طور متوسط ۰،۱۴٪ است.

از این مقدار، می توانیم ببینیم که ETZSearch می تواند عملکرد اصلی RD HM12.0 را به طور موثر حفظ کند، در حالی که صرفه جویی در کل کدگذاری و زمان صرفه جویی در زمان ETZSearch بسیار محدود است.

Algorithm	Class	Resolution	Sequence	Low-delay-main coding profile				Random-access-main coding profile			
				BDPSNR (dB)	BDBR (%)	TETS (%)	METS (%)	BDPSNR (dB)	BDBR (%)	TETS (%)	METS (%)
ETZSearch [20] v.s. HM-TZSearch	A	2560×1600	Traffic	n/a ²	n/a	n/a	n/a	-0.014	0.39	-3.98	-5.53
			PeopleOnStreet	n/a	n/a	n/a	n/a	-0.009	0.20	-4.11	-6.19
	B	1920×1080	ParkScene	-0.003	0.09	-2.36	-3.35	-0.002	0.07	-1.84	-3.03
			Cactus	-0.004	0.13	-4.31	-5.45	-0.001	0.06	-2.96	-4.24
			BQTerrace	-0.002	0.09	-2.24	-3.08	-0.002	0.10	-1.57	-2.19
			BasketballDrive	-0.001	0.01	-6.44	-8.15	-0.002	0.10	-4.86	-7.02
	C	832×480	BQMall	-0.002	0.04	-4.52	-6.49	-0.008	0.20	-3.14	-4.79
			PartyScene	-0.001	0.03	-2.36	-3.35	-0.005	0.10	-1.84	-3.03
			BasketballDrill	-0.007	0.19	-3.30	-4.68	-0.001	0.03	-2.72	-4.34
	D	416×240	BQSquare	-0.001	0.03	-2.18	-4.23	-0.003	0.07	-1.87	-3.06
			BlowingBubbles	-0.002	0.07	-2.32	-3.81	-0.007	0.17	-1.43	-2.45
			BasketballPass	-0.002	0.04	-5.48	-6.87	-0.008	0.14	-3.14	-5.17
	E	1280×720	FourPeople	-0.002	0.05	-1.65	-3.92	n/a	n/a	n/a	n/a
			Johnny	-0.008	0.15	-1.99	-4.39	n/a	n/a	n/a	n/a
			KristenAndSara	-0.006	0.27	-2.68	-4.10	n/a	n/a	n/a	n/a
			Vidyo1	-0.007	0.21	-1.18	-1.40	n/a	n/a	n/a	n/a
			Vidyo3	-0.009	0.29	-1.36	-1.84	n/a	n/a	n/a	n/a
Vidyo4	-0.007	0.17	-2.77	-3.66	n/a	n/a	n/a	n/a			
Average			-0.004	0.12	-2.95	-4.30	-0.005	0.14	-2.79	-4.25	
ISOCC [28] v.s. HM-TZSearch	A	2560×1600	Traffic	n/a	n/a	n/a	n/a	-0.002	0.05	-0.89	-1.19
			PeopleOnStreet	n/a	n/a	n/a	n/a	-0.017	0.39	-0.77	-0.69
	B	1920×1080	ParkScene	-0.007	0.22	-0.47	-0.82	-0.004	0.12	-1.14	-1.20
			Cactus	-0.002	0.08	-0.92	-1.20	-0.005	0.19	-1.14	-1.20
			BQTerrace	-0.003	0.14	-3.48	-4.25	-0.006	0.31	-3.52	-5.15
			BasketballDrive	-0.004	0.16	-1.75	-2.24	-0.002	0.17	-1.45	-1.93
	C	832×480	BQMall	-0.008	0.19	-1.38	-1.41	-0.001	0.03	-1.20	-1.83
			PartyScene	-0.003	0.08	-0.98	-1.45	-0.009	0.20	-0.82	-1.47
			BasketballDrill	-0.002	0.04	-0.98	-1.63	-0.004	0.11	-0.90	-1.56
	D	416×240	BQSquare	-0.005	0.11	-0.81	-1.61	-0.010	0.21	-1.06	-1.75
			BlowingBubbles	-0.004	0.11	-1.14	-1.87	-0.006	0.14	-1.07	-1.33
			BasketballPass	-0.029	0.62	-0.87	-1.00	-0.008	0.16	-1.06	-0.90
	E	1280×720	FourPeople	-0.005	0.03	-0.73	-1.01	n/a	n/a	n/a	n/a
			Johnny	-0.005	0.22	-1.09	-1.17	n/a	n/a	n/a	n/a
			KristenAndSara	-0.006	0.23	-1.06	-1.17	n/a	n/a	n/a	n/a
			Vidyo1	-0.003	0.16	-1.01	-1.17	n/a	n/a	n/a	n/a
			Vidyo3	-0.007	0.18	-0.92	-1.28	n/a	n/a	n/a	n/a
Vidyo4	-0.005	0.30	-0.87	-1.14	n/a	n/a	n/a	n/a			
Average			-0.006	0.18	-1.15	-1.53	-0.006	0.17	-1.25	-1.68	
Proposed v.s. HM-TZSearch	A	2560×1600	Traffic	n/a	n/a	n/a	n/a	-0.015	0.41	-8.23	-11.52
			PeopleOnStreet	n/a	n/a	n/a	n/a	-0.071	1.58	-16.44	-25.69
	B	1920×1080	ParkScene	-0.020	0.63	-13.54	-20.45	-0.022	0.68	-9.64	-14.23
			Cactus	-0.012	0.49	-17.82	-22.35	-0.017	0.73	-12.90	-19.45
			BQTerrace	-0.006	0.29	-11.29	-17.95	-0.008	0.38	-10.99	-14.85
			BasketballDrive	-0.010	0.43	-26.14	-33.68	-0.015	0.72	-21.08	-27.62
	C	832×480	BQMall	-0.038	0.92	-23.36	-28.81	-0.043	1.05	-13.07	-20.74
			PartyScene	-0.030	0.67	-14.16	-22.35	-0.039	0.84	-11.01	-18.36
			BasketballDrill	-0.038	0.96	-19.98	-27.53	-0.031	0.76	-14.71	-22.94
	D	416×240	BQSquare	-0.017	0.40	-8.54	-13.35	-0.036	0.80	-7.28	-11.06
			BlowingBubbles	-0.023	0.6	-13.99	-17.65	-0.037	0.91	-8.90	-14.65
			BasketballPass	-0.053	1.12	-21.36	-28.91	-0.070	1.46	-13.21	-21.17
	E	1280×720	FourPeople	-0.009	0.20	-10.12	-13.69	n/a	n/a	n/a	n/a
			Johnny	-0.008	0.21	-10.57	-14.12	n/a	n/a	n/a	n/a
			KristenAndSara	-0.018	0.50	-11.77	-14.71	n/a	n/a	n/a	n/a
			Vidyo1	-0.008	0.28	-11.21	-13.56	n/a	n/a	n/a	n/a
			Vidyo3	-0.017	0.57	-12.48	-14.53	n/a	n/a	n/a	n/a
Vidyo4	-0.011	0.54	-14.32	-18.25	n/a	n/a	n/a	n/a			
Average			-0.020	0.55	-15.04	-20.12	-0.034	0.86	-12.29	-18.52	

جدول ۶ - خلاصه ای از نتایج

دلایل صرفه جویی در زمان کم اینهاست : (۱) ETZSearch برای خاتمه زود هنگام فرایند ME از جست و جوی لوزی چهار نقطه استفاده می کند و زمان کدگذاری بیشتری برای چک کردن این چهار نقطه استفاده می کند؛ (۲) HM-TZSearch اصلی برخی از استراتژی های خاتمه زود هنگام را برای فرآیند جست و جوی تطبیق بلوک پذیرفته است که باعث می شود ETZSearch در HM-TZSearch کار نکند.

برای مشخصات کدگذاری LDM، ISOCC به طور میانگین فقط ۱،۱۵٪ و ۱،۵۳٪ صرفه جویی در زمان کلی رمزگذاری و صرفه جویی در زمان ME دارد. میانگین BDPSNR و BDBR بین ISOCC و HM-TZSearch اصلی به ترتیب -۰،۰۰۶ dB و ۰،۱۸٪ می باشد. برای مشخصات کدگذاری RAM، ISOCC زمان کل کدگذاری و زمان کدگذاری ME را به ترتیب به ترتیب ۱،۲۵٪ و ۱،۶۸٪ کاهش می دهد. متوسط BDPSNR و BDBR بین ISOCC و HM-TZSearch اصلی به ترتیب -۰،۰۰۶ dB و ۰،۱۷٪ است. از این مقادیر می توان دید که صرفه جویی در پیچیدگی ISOCC بسیار محدود است. این به این دلیل است که ISOCC فقط الگویی جست و جوی تطبیق بلوک HM-TZSearch را اصلاح می کند، در حالی که اکثر فرآیند های جست و جوی بلوک در فرایند رمزگذاری HM-TZSearch رها می شوند.

همچنین می توان دید که روش پیشنهاد شده به طور موثری پیچیدگی کدبندی ME را از رمزگذار HEVC حذف می کند، در حالی که عملکرد RD قابل مقایسه ای دارد. برای مشخصات کدگذاری LDM، روش پیشنهاد شده، کل زمان رمزگذاری را از ۱۰،۱۲٪ تا ۲۶،۱۴٪ و به طور متوسط ۱۵،۰۴٪ کاهش می دهد و موجب صرفه جویی در زمان ME از ۱۳،۳۵٪ تا ۳۳،۶۸٪، به طور متوسط ۲۰،۱۲٪ شده است. در همین حال، BDPSNR بین روش پیشنهادی و HM-TZSearch اصلی از -۰،۰۵۳ dB تا -۰،۰۰۶ dB به طور متوسط 0.020 dB است. BDBR بین روش پیشنهاد شده و HM-TZSearch اصلی از ۰،۲۱٪ تا ۱،۱۲٪، به طور متوسط ۰،۵۵٪ است. برای مشخصات کدگذاری RAM، روش پیشنهادی زمان کدگذاری را از ۷،۲۸٪ تا ۲۱،۰۸٪ و به طور متوسط ۱۲،۲۹٪ کاهش می دهد؛ و زمان صرفه جویی در وقت ME از ۱۱،۰۶٪ تا ۲۷،۶۲٪، به طور متوسط ۱۸،۵۲٪ است. در عین حال، BDPSNR بین روش پیشنهادی و HM-TZSearch اصلی از -۰،۰۷۰ dB تا -۰،۰۰۸ dB به طور متوسط -0.034 dB

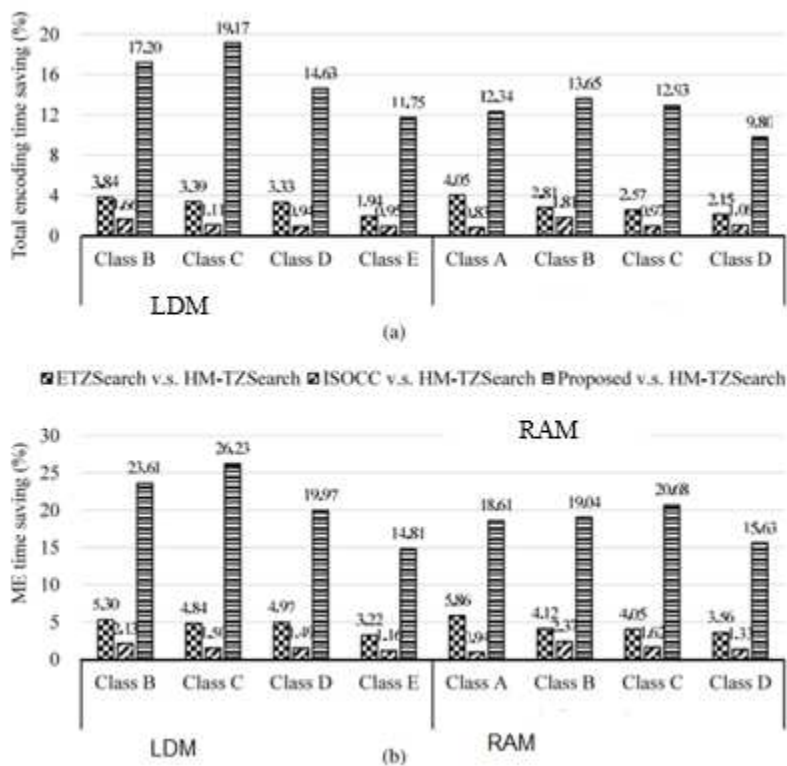
است. و BDBR بین روش پیشنهاد شده و HM-TZSearch اصلی از ۰,۳۸٪ تا ۱,۵۸٪، به طور متوسط ۰,۸۶٪ است.

از این مقادیر می توان دید که TETS و TMES الگوریتم پیشنهادی با مشخصات کدگذاری LDM بزرگتر از پروفایل کدگذاری RAM هستند، به همین دلیل است که پیچیدگی کدگذاری LDM پیچیده تر از RAM است. علاوه بر این، روش پیشنهادی، پیچیدگی رمزگذاری بیشتر برای توالی های ویدئویی، BasketballDrill، BasketballPass و PeopleOnStreet را حذف می کند، این به این دلیل است که این فیلم ها دارای محتوای فعالیت حرکتی سریع هستند که باعث می شود زمان تعیین نهایی بهترین نقطه در جست و جوی تطبیق بلوک را بسیار بیشتر کند. بنابراین، فرآیند جست و جوی تطبیق بلوک برای این توالی های ویدئویی رها می شود و در نتیجه صرفه جویی در زمان کدگذاری قابل توجهی به دست می آید.

برای نشان دادن کارایی صرفه جویی در زمان روش پیشنهادی، مقایسه صرفه جویی در زمان کدگذاری بین ETZSearch، ISOCC و روش پیشنهادی در شکل ۵ ارائه شده است. دیده می شود که روش پیشنهادی می تواند به طور موثر زمان گذار H.265 / HEVC و زمان کدگذاری پردازش ME را کاهش دهد. برای مشخصات کدگذاری LDM، میانگین صرفه جویی در زمان کلی رمزگذاری ETZSearch، ISOCC و روش پیشنهادی به ترتیب ۲,۹۵٪، ۱,۱۵٪ و ۱۵,۰۴٪ است. و میانگین صرفه جویی ME در ETZSearch، ISOCC، و روش پیشنهادی به ترتیب ۴,۳۰٪، ۱,۵۳٪ و ۲۰,۱۲٪ می باشد.

برای مشخصات کدگذاری RAM، متوسط صرفه جویی در کل زمان کدگذاری ETZSearch، ISOCC و روش پیشنهادی به ترتیب ۲,۷۹٪، ۱,۲۵٪ و ۱۲,۲۹٪ است. و میانگین صرفه جویی ME در ETZSearch، ISOCC، و روش پیشنهاد شده به ترتیب ۴,۲۵٪، ۱,۶۸٪ و ۲۸,۵۲٪ است. در مقایسه با ETZSearch، روش پیشنهادی به ترتیب ۱۲,۴۶٪ و ۹,۷۷٪ در زمان کل کدگذاری برای پرونده های کدگذاری LDM و RAM صرفه جویی می کند. روش پیشنهاد شده به ترتیب حدود ۱۶,۵۳٪ و ۱۴,۹۰٪ زمان کدگذاری برای فرایند ME در کد کننده H.265 / HEVC با پرونده های کدگذاری RAM و LDM را کاهش می دهد. برای مشخصات کدگذاری LDM، روش

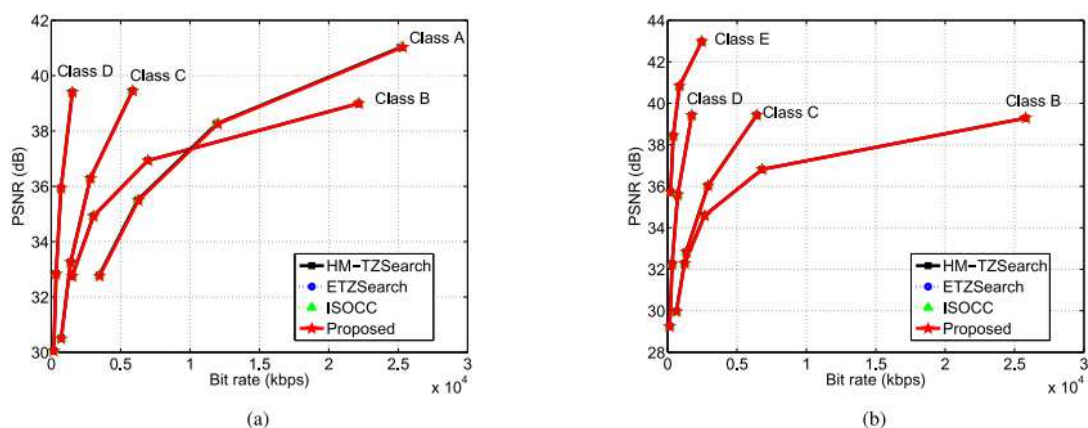
پیشنهادی می تواند زمان کدگذاری کل را ۱۴,۰۵٪ و زمان کدگذاری ME را ۱۸,۸۸٪ در مقایسه با ISOCC کاهش دهد؛ روش پیشنهادی موجب صرفه جویی ۱۱,۱۸٪ در کل زمان کدگذاری و ۱۶,۸۵٪ در زمان کدگذاری ME برای رمزگذاری RAM در مقایسه با ISOCC می شود.



شکل ۵ - خلاصه صرفه جویی در زمان کدگذاری (a) صرفه جویی در زمان کدگذاری کد کننده HEVC. (ب) صرفه جویی در زمان کل رمزگذاری فرایند HEVC در ME

از این مقادیر می توان نتیجه گیری کرد که روش پیشنهاد شده پیچیدگی کدگذاری پروسه ME کد کننده H.265 / HEVC را کاهش می دهد که کارآمدتر از ETZSearch و ISOCC است. برای نشان دادن مستقیم عملکرد RD، منحنی های RD کلاس A، B، C و D، با پروفایل های برنامه ریزی RAM و منحنی های RD کلاس B، C، D و E به ترتیب در شکل ۶ (a) و شکل ۶ (b) آمده است. از شکل ۶ می توان دید که روش پیشنهادی سریع ME تقریباً همان عملکرد RD را با ETZSearch، HM-TZSearch و ISOCC به دست می آورد. از داده های

فوق، نتیجه می گیریم که روش پیشنهاد شده به طور موثر پیچیدگی کدگذاری را از ME پیکسل صحیح در H.265 / HEVC حذف می کند، در عین حال، میزان تخریب عملکرد RD معقول است.



شکل ۶ - منحنی های (a) RD مشخصات کدگذاری (b) RAM مشخصات کدگذاری LDM

جدول ۷- نرخ تصمیم گیری غلط بهترین MV (%)

Profile	Sequence	ETZSearch	ISOCC	Proposed
RAM	BQSquare	4.97	5.08	5.63
	BasketballDrill	5.46	6.39	7.27
	Johnny	n/a	n/a	n/a
	Cactus	7.05	7.62	8.11
	Traffic	5.41	4.16	5.72
	Average	5.72	5.81	6.68
LDM	BQSquare	2.01	2.35	2.92
	BasketballDrill	6.68	6.23	7.48
	Johnny	3.54	4.03	4.11
	Cactus	8.47	7.68	9.17
	Traffic	n/a	n/a	n/a
	Average	5.18	5.07	5.92

C. نرخ تصمیم اشتباه بهترین بردار حرکت

برای ارزیابی بهترین کارایی تصمیم گیری MV در روش پیشنهادی ME سریع ، نرخ تصمیم گیری غلط بهترین MV ی ETZSearch، ISOCC و روش پیشنهادی ما با پنج توالی استاندارد HEVC، BQSquare، BasketballDrill، "Johny"، "Cactus" و "Traffic"، تحت پروفایل های کدگذاری RAM و LDM ارزیابی شده است. جدول ۷ نتایج آزمون میانگین بهترین نرخ تصمیم گیری اشتباه MV را نشان می دهد.

از جدول ۷ می توان دید که برای کدگذاری RAM میانگین نرخ تصمیم گیری غلط بهترین MV برای ETZSearch، ISOCC و روش پیشنهاد ما به ترتیب ۵,۷۲٪، ۵,۸۱٪ و ۶,۶۸٪ است. برای مشخصات کدگذاری LDM میانگین نرخ تصمیم گیری غلط بهترین MV برای ETZSearch، ISOCC و روش پیشنهادی ما به ترتیب ۵,۱۸٪، ۵۷,۰۷٪ و ۵,۹۲٪ است. از این نتایج می توان دید که نرخ تصمیم گیری غلط بهترین MV برای روش پیشنهاد ما بسیار نزدیک به ETZSearch و ISOCC است، ضمن اینکه مقدار آن در محدوده قابل قبول است.

۵. نتیجه گیری

ما برای اینکه H.265 / HEVC را وادار کنیم تا نیازهای کاربردهای پخش همگانی فعلی، مانند پیچیدگی کم، فشرده سازی بالا را برآورده کند، روش سریع ME را برای کاهش پیچیدگی محاسباتی H.265 / HEVC در این مقاله پیشنهاد دادیم. در مرحله اول، براساس نوع پیش بینی، تمام PU ها به دو دسته تقسیم می شوند، PU والد و PU فرزند. سپس، بر اساس بهترین همبستگی انتخاب MV بین PU والد و PU فرزند، فرایند جست و جوی تطبیق بلوک PU فرزند در فرایند ME به صورت انطباقی رها می شود ، اگر PU والد ISP را به عنوان بهترین MV نهایی خود انتخاب کند. نتایج آزمایشگاهی نشان می دهند که روش پیشنهادی سریع ME به طور موثر پیچیدگی محاسباتی فرایند ME را از کد کننده H.265 / HEVC حذف می کند در حالی که عملکرد قابل مقایسه RD را حفظ می کند.

REFERENCES

- [1] B. Chen *et al.*, "Color image analysis by quaternion-type moments," *J. Math. Imag. Vis.*, vol. 51, no. 1, pp. 124–144, Jan. 2015.
- [2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
- [3] "Information technology-high efficiency coding and media delivery in heterogeneous environments—Part 2: High efficiency video coding," ISO/IEC 23008-2:2013 ITU-T Rec. H.265, 2013.
- [4] J. Lei, K. Feng, M. Wu, S. Li, and C. Hou, "Rate control of hierarchical B prediction structure for multi-view video coding," *Multimedia Tools Appl.*, vol. 72, no. 1, pp. 825–842, Sep. 2014.
- [5] Z. Pan, S. Kwong, Y. Zhang, J. Lei, and H. Yuan, "Fast coding tree unit depth decision for high efficiency video coding," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, Oct. 2014, pp. 3214–3218.
- [6] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective CU size decision method for HEVC encoders," *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.
- [7] Y. Zhang *et al.*, "Machine learning-based coding unit depth decisions for flexible complexity allocation in high efficiency video coding," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2225–2238, Jul. 2015.
- [8] Z. Pan, S. Kwong, M.-T. Sun, and J. Lei, "Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC," *IEEE Trans. Broadcast.*, vol. 60, no. 2, pp. 405–412, Jun. 2014.
- [9] L. Shen, Z. Zhang, and Z. Liu, "Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatiotemporal correlations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 10, pp. 1709–1722, Oct. 2014.
- [10] S. Ahn, B. Lee, and M. Kim, "A novel fast CU encoding scheme based on spatiotemporal encoding parameters for HEVC inter coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 422–435, Mar. 2015.
- [11] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Lshiguro, "Motioncompensated interframe coding for video conferencing," in *Proc. Nat. Telecommun. Conf.*, New Orleans, LA, USA, 1981, pp. G5.3.1–G5.3.5.
- [12] L.-M. Po and W.-C. Ma, "A novel four-step search algorithm for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 313–317, Jun. 1996.
- [13] S. Zhu and K.-K. Ma, "A new diamond search algorithm for fast block-matching motion estimation," *IEEE Trans. Image Process.*, vol. 9, no. 2, pp. 287–290, Feb. 2000.
- [14] C. Zhu, X. Lin, and L.-P. Chau, "Hexagon-based search pattern for fast block motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 5, pp. 349–355, May 2002.
- [15] F. Bossen, D. Flynn, and K. Suehring, *JCTVC AHG Report: HEVC HM Software Development and Software Technical Evaluation (AHG3)*, document JCTVC-O0003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Geneva, Switzerland, Oct. 2013.
- [16] Z. Pan, S. Kwong, L. Xu, Y. Zhang, and T. Zhao, "Predictive and distribution-oriented fast motion estimation for H.264/AVC," *J. Real Time Image Process.*, vol. 9, no. 4, pp. 597–607, Dec. 2014.
- [17] Z. Pan, Y. Zhang, and S. Kwong, "Efficient motion and disparity estimation optimization for low complexity multiview video coding," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 166–176, Jun. 2015.
- [18] N. Parnachand, L. N. Alves, and A. Navarro, "Fast motion estimation algorithm for HEVC," in *Proc. Int. Conf. Consum. Electron. Berlin (ICCE-Berlin)*, Berlin, Germany, Sep. 2012, pp. 34–37.
- [19] X. Li, R. Wang, W. Wang, Z. Wang, and S. Dong, "Fast motion estimation methods for HEVC," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, Beijing, China, Jun. 2014, pp. 1–4.
- [20] Z. Pan, Y. Zhang, S. Kwong, X. Wang, and L. Xu, "Early termination for TZSearch in HEVC motion estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013, pp. 1389–1393.
- [21] X. Li, R. Wang, X. Cui, and W. Wang, "Context-adaptive fast motion estimation of HEVC," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Lisbon, Portugal, May 2015, pp. 2784–2787.
- [22] S.-H. Yang, J.-Z. Jiang, and H.-J. Yang, "Fast motion estimation for HEVC with directional search," *Electron. Lett.*, vol. 50, no. 9, pp. 673–675, Apr. 2014.
- [23] Z.-T. Liao and C.-A. Shen, "A novel search window selection scheme for the motion estimation of HEVC systems," in *Proc. Int. SoC Design Conf. (ISOCC)*, Gyeongju, South Korea, Nov. 2015, pp. 267–268.
- [24] Y. Li, Y. Liu, H. Yang, and D. Yang, "An adaptive search range method for HEVC with the k-nearest neighbor algorithm," in *Proc. Vis. Commun. Image Process. (VCIP)*, Singapore, Dec. 2015, pp. 1–4.
- [25] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—Including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.
- [26] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-J1100, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Geneva, Switzerland, Mar. 2012.
- [27] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-D600, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Geneva, Switzerland, Jan. 2011.
- [28] N. Parmar and M. H. Sunwoo, "Enhanced test zone search motion estimation algorithm for HEVC," in *Proc. Int. SoC Design Conf. (ISOCC)*, Jeju, South Korea, Nov. 2014, pp. 260–261.
- [29] G. Bjontegaard, *Calculation of Average PSNR Differences Between RD Curves*, document VCEG-M33, ITU-T VCEG, Geneva, Switzerland, Apr. 2001.

این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

لیست مقالات ترجمه شده ✓

لیست مقالات ترجمه شده رایگان ✓

لیست جدیدترین مقالات انگلیسی ISI ✓

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی