

# چند معیار جدید انطباق، مناسب پیاده‌سازی سخت‌افزاری برای کاربرد فشرده‌سازی ویدیو

محمد رحمتی  
استادیار  
دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

حمیدرضا پوررضا  
دانشجوی دکتری  
دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر

فرید به‌آذین  
استادیار  
مجتمع تحقیقاتی یا مهدی، سازمان هوافضا

## چکیده

از مهمترین بخش‌های یک فشرده‌ساز ویدیو، بخش تخمین حرکت آن است. یکی از تکنیک‌های رایج بکار رفته برای این بخش، استفاده از تکنیک انطباق بلوک است. در این تکنیک با استفاده از یک معیار انطباق و یک روش جستجو، حرکت انتقالی برای بلوک‌های تصویر تخمین زده می‌شود. مقاله حاضر به ارائه سه معیار جدید انطباق که مناسب پیاده‌سازی سخت‌افزاری است می‌پردازد. این معیارها بر اساس دو معیار موجود پایه‌گذاری شده و با اصلاحاتی کارایی آنها بهبود داده شده است. دو معیار از این معیارها، معیارهایی مبتنی بر انطباق صفحه بیت (Bit-Plane Matching) است. در این تکنیک‌ها، از چهار صفحه بیت که کاهش نمونه داده شده‌اند استفاده شده است. معیار مبنا برای این دو معیار، معیاری بوده که برای کاربرد تثبیت (Stabilization) رشته تصاویر دیجیتالی ارائه شده است. مقاله حاضر با تغییر این معیار برای کاربرد فشرده‌سازی، ضمن بیان ایرادات آن به بررسی دلایل بروز و رفع آن پرداخته است. آزمایش‌های انجام شده نمایانگر بالاتر بودن کارایی معیارهای پیشنهادی این مقاله نسبت به معیار پایه‌ای است که برای کاربرد فشرده‌سازی اصلاح شده است، بطوریکه روش اول بطور متوسط ۶٪ و روش دوم بطور متوسط ۹٪ افزایش کارایی را ارائه کرده است. معیار سوم، یک شکل وقتی از معیار انطباق کلاس بندی تفاضل پیکسلها (PDC) است. در روش ارائه شده، دو اصلاح نسبت به PDC پایه انجام شده است که یکی از این اصلاحات به منظور کاهش بار محاسباتی آن و دیگری به منظور افزایش کارایی بوده است. آزمایش‌ها نشان می‌دهند که این روش قادر است کارایی بیش از ۹۹٪ PDC را در حالتیکه از بهترین مقدار آستانه استفاده می‌شود را ارائه نماید. عملیات اضافه شده به PDC پایه برای وقتی کرده آن تنها ۰/۰۰۵۶ عمل بر پیکسل است. که البته با اصلاح ارائه شده برای کاهش بار محاسبات، خود پیاده‌سازی PDC نیز ساده‌تر شده است.

## کلمات کلیدی

تثبیت کننده دیجیتالی ویدئو، تخمین حرکت، جبران حرکت، بنیان ماشین، پردازش تصاویر ویدئویی، فشرده سازی ویدئو

## New Matching Criteria for Video Compression Application: Appropriate for Hardware Implementation

H.R. Pourreza  
Graduate Student  
Computer Engineering Department,  
Amirkabir University of Technology

M. Rahmati  
Assistant Professor  
Computer Engineering Department,  
Amirkabir University of Technology

F. Behazin  
Assistant Professor  
Ya Mahdi Research Center

## Abstract

*Estimation of motion is an important step in video compression applications. Block matching motion estimation technique is one of the approaches that are widely used for this purpose. In this technique, by incorporating a matching criterion and employing a search strategy, the translational motion is estimated on the image blocks. In this paper, we present three new criteria, which are suitable for hardware implementation. In two of these introduced criteria incorporated an existing criterion and modified it for a better performance. These criteria are based on bit-plane matching where four sub-sampled bit-planes were used. Basically, these two are based on a criterion which is presented for digital image stabilization application. In this paper, while this were modified criterion for video compression we investigated the cause of problems and suggested the possible solution. Our experimental results indicated that the two proposed criteria yield better performance than the original method, where in the first case the performance increased by %6 and for the other case the growth is %9. Our third criterion is the adaptive form of pixel difference classification (PDC). In our proposed method, two modifications were proposed on the basic PDC, in which for one of them we reduced computational burden and in the other the performance we increased. Our experiments indicate that our proposed adaptive approach can yield %99 of PDC performance when used the best threshold. only need about 0.0056 additional operations per pixel for obtaining an adaptive form of PDC. But the modification that is presented for reducing the computational burden can compensate these additional computations.*

## Keywords

*Digital Video Stabilizer, Motion Destimation, Motion Compensation, Machine Vision, Digital Video Image Processing, Vedeo compression*

## مقدمه

ذخیره‌سازی و ارسال تصاویر ویدیو، بدلیل حجم بالای اطلاعات، همیشه کاربران را با مشکل مواجه می‌سازد. استفاده از تکنیکهای فشرده‌سازی برای غلبه بر این مشکل همواره مورد توجه بوده است. روشهای کدکردن پیوندی<sup>۱</sup> از جمله روش‌هایی هستند که در خیلی از استانداردهای فشرده‌سازی ویدیو مانند ISO/IEC MPEG1/2/4، CCITT H.261 و ITU-T H.263 مورد توجه قرار گرفته‌اند [1,2]. این روش‌ها از پیش‌بینی و تبدیلات بطور همزمان استفاده کرده و نرخ فشرده‌سازی بالایی را می‌توانند ارائه نمایند. شکل (۱) بلوک دیاگرام عمومی چنین کدکننده‌ای را نشان می‌دهد. جبران‌سازی حرکت، به عنوان تکنیکی مبتنی بر پیش‌بینی، افزونگی<sup>۲</sup> زمانی بین فریم‌های متوالی را برای افزایش نرخ فشرده‌سازی بکار می‌گیرد. بدین ترتیب در چنین کدکننده‌ای موضوع تخمین حرکت جایگاه ویژه‌ای پیدا می‌کند. دلایل زیر موجب توجه محققین به موضوع تخمین حرکت در فشرده‌سازی ویدیو شده است:

بخش زیادی از نیاز محاسباتی در یک کدکننده ویدیو مربوط به بخش تخمین حرکت است (حدود ۶۰ تا ۸۰٪ زمان محاسبه) که این مساله کارایی کدکننده را از نظر سرعت بشدت متاثر می‌کند [1].

با یک نرخ بیت خاص، دقت الگوریتم تخمین حرکت اثر بسزایی بر کیفیت تصویر کد شده می‌گذارد [1].

در استاندارد مشخص نشده که از چه روشی برای تخمین بردارهای حرکت استفاده شود، و لذا این موضوع زمینه‌ای مناسب را برای رقابت در ساخت کدکننده‌های کارا تر فراهم کرده است [1]. هر چند که روش پیدا کردن بردارهای حرکت می‌تواند مستقیماً نرخ بیت خروجی کدکننده ویدیو را متاثر کند، با این حال برای یک گیرنده تصاویر کد شده مهم نیست که از چه روشی برای پیدا کردن این بردارها استفاده شده است.

تخمین حرکت علاوه بر فشرده‌سازی ویدیو دارای کاربردهای دیگری از جمله تثبیت الکترونیکی رشته تصاویر دیجیتالی<sup>۳</sup> [7-10] و ردیابی هدف [11,12] نیز هست. البته بدلیل تفاوت در معیار ارزیابی نتیجه تخمین حرکت در کاربردهای مختلف،

یک نتیجه خوب از کاربردی به کاربرد دیگر ممکن است متفاوت باشد. به عنوان مثال معیار ارزیابی نتیجه در فشرده‌سازی ویدیو، کم بودن خطای پیش‌بینی فریم<sup>۴</sup> است در حالیکه این معیار برای ردیابی هدف و یا تثبیت الکترونیکی رشته تصاویر دیجیتال، دقت تخمین حرکت است.

تکنیک انطباق بلوک یکی از تکنیک‌های ساده و کارا است که بصورت گسترده‌ای در بخش تخمین حرکت کدکننده‌های پیوندی بکار گرفته شده است [2]. در این تکنیک که عمدتاً بر مبنای دو فریم است و حرکت را بین این دو فریم تخمین می‌زند، یکی از فریمها (فریم مبنا) به مجموعه‌ای از بلوکهای غیرهمپوشان به ابعاد  $N \times N$  پیکسل تقسیم شده و با مراجعه به فریم دیگر (فریم مرجع) و با استفاده از الگوریتم‌های انطباق بلوک<sup>۵</sup> (BMA)، بردار حرکت برای هر یک از این بلوکها تخمین زده می‌شود. فریم مرجع از نظر زمانی ممکن است قبل (تخمین حرکت مستقیم) و یا بعد از فریم مبنا (تخمین حرکت معکوس) باشد. بدین ترتیب با داشتن فریم مرجع و بردارهای حرکت، فریم مبنا قابل پیش‌بینی خواهد بود. برای تخمین بردار حرکت هر بلوک، جستجو در فریم مرجع، در مکانی متناظر با مکان بلوک در فریم مبنا، ولی در فضایی بزرگتر از اندازه آن و به اندازه  $(N+2w+1) \times (N+2w+1)$  صورت می‌پذیرد. به این فضا محدوده جستجو<sup>۶</sup> (SA) گفته می‌شود و اندازه آن وابسته به بیشترین مقدار حرکتی است که بین دو فریم مبنا و مرجع تخمین زده می‌شود. در اینجا این مقدار  $w$  فرض شده است.

برای تخمین بردار حرکت بلوک مبنا، با این دو فرض که اولاً حرکت بلوک از نوع انتقالی است و ثانیاً، کلیه پیکسل‌های بلوک دارای حرکت یکسانی بوده‌اند، جستجو برای یافتن بلوکی از SA که بیشترین شباهت را به بلوک مبنا دارد انجام می‌شود. بدین ترتیب فرایند تخمین بردار حرکت با دو مساله مواجه است، یکی نحوه جستجو در SA و دیگری معیاری که تحت آن، شباهت بلوک مبنا با بلوکهای انتخابی از SA مورد ارزیابی قرار می‌گیرد، که از این معیار تحت عنوان معیار انطباق نام برده خواهد شد. قاعدتاً بهترین روش جستجو آن است که تمام بلوکهای نامزد انطباق در SA، از نظر انطباق مورد ارزیابی قرار گیرند. به این روش اصطلاحاً جستجوی کامل<sup>۷</sup> (FS) گفته می‌شود [2]. در صورتیکه تابع توزیع خطای پیکسل‌ها در محل بهترین انطباق نرمال باشد، بهترین معیاری که می‌توان برای انطباق استفاده کرد، معیار مجموع مربعات خطا<sup>۸</sup> (SSD) است [26]، که البته برای کاهش بار محاسبات در خیلی از موارد بجای استفاده از SSD از مجموع قدرمطلق‌های خطا<sup>۹</sup> (SAD) استفاده میشود [26]. نکته قابل ذکر اینکه معیارهای SSD و SAD در واقع معیار عدم انطباق هستند و نه انطباق، چرا که هر چه انطباق بین دو بلوک بیشتر باشد، این دو معیار عدد کوچکتری را ارائه می‌کنند، ولی با این وجود چون در خیلی از متون بدون توجه به این امر، از این دو معیار هم با نام معیار انطباق نام می‌برند، در اینجا نیز چنین شده است.

همانطور که ذکر گردید، یکی از خصوصیات برجسته تخمین حرکت در فشرده‌سازی ویدیو، بار محاسباتی بالای آن است. همین امر موجب شده است که عمده کارهای صورت گرفته توسط محققین، به منظور کاهش بار محاسباتی باشد. در روش‌هایی که هدف کاهش بار محاسباتی بوده است، برای رسیدن به این هدف از سه منظر به این موضوع پرداخته شده است [2]. روشهای دسته اول سعی کرده‌اند که تعداد بلوکهایی که در فرایند جستجو مورد ارزیابی انطباق قرار می‌گیرند را کاهش دهند. تکنیک‌های مختلفی برای کاهش بلوکهای نامزد انطباق و برای یافتن محل بهترین انطباق ارائه شده‌اند. برخی از این تکنیک‌ها سعی کرده‌اند با این فرض که مقادیر معیار انطباق در محدوده جستجو تنها دارای یک ماکزیمم و یا مینیمم است، تعداد بلوکهای مورد ارزیابی انطباق را کاهش دهند، و برخی دیگر سعی کرده‌اند تا با پیش‌بینی بردار حرکت به این هدف نایل آیند [20-2,13]. دسته دوم سعی کرده‌اند تا نقاط بکار رفته در ارزیابی انطباق را کاهش دهند [2]. و بالاخره دسته سوم سعی کرده‌اند تا رابطه ارزیابی انطباق را ساده کنند تا بدین ترتیب با بار محاسبات کمتری بتوان این ارزیابی را انجام داد [21,7,2,5].

روابط ارزیابی انطباق ممکن است با ملاحظات برای پیاده‌سازی سخت‌افزاری و یا نرم‌افزاری طراحی شده باشند. معیاری که برای پیاده‌سازی توسط سخت‌افزاری خاص بسیار خوب باشد، ممکن است در پیاده‌سازی نرم‌افزاری کارایی مورد نظر را ارائه نکند. معیار PDC<sup>۱۰</sup> نمونه‌ای از این نوع معیارها است که در ادامه مقاله شرح داده خواهد شد. ساده‌سازی معیار انطباق برای پیاده‌سازی سخت‌افزاری، امکان طراحی تراشه‌هایی با حجم و مصرف پایین را فراهم می‌سازد که این امر در کاربردهایی چون تلفن همراه تصویری<sup>۱۱</sup> می‌تواند از اهمیت ویژه‌ای برخوردار باشد. بر این اساس، مقاله حاضر به ارائه چند معیار جدید انطباق مناسب برای پیاده‌سازی سخت‌افزاری می‌پردازد.

در ادامه مقاله در بخش دوم، الگوریتم انطباق صفحه بیت<sup>۱۲</sup> (BPM) در حالت کلی بیان و بر اساس آن دو معیار جدید از این خانواده ارائه می‌شود. بخش سوم نیز ضمن معرفی معیار انطباق کلاس‌بندی تفاضل پیکسل‌ها (PDC)، به بیان اشکالات آن و ارائه یک روش جدید وقفی می‌پردازد. در بخش چهارم نیز مقایسه‌ای روی سخت‌افزارهای مورد نیاز برای پیاده‌سازی الگوریتم‌های پیشنهادی انجام می‌شود. در بخش‌های پنجم و ششم نیز نتایج آزمایش‌های انجام شده و نتیجه‌گیری آمده‌اند.

## ۱- انطباق صفحه بیت

یکی از راه‌های کاهش پیچیدگی سخت‌افزار تخمین حرکت، استفاده از ارائه روشنایی پیکسل‌ها با سطوح کمتر است. هر چه تعداد بیت‌های بکار رفته برای بیان روشنایی پیکسل کمتر باشد، بالطبع سخت‌افزاری که برای کار محاسبات بر روی مقدار این پیکسل‌ها طراحی می‌شود، پیچیدگی کمتری را دارا خواهد بود. در این راستا بهترین وضعیت آن است که روشنایی پیکسل‌های تصویر تنها با یک بیت بیان شوند. در این حالت معیارهای SSD و SAD به معیار BPM تبدیل می‌شوند. بدین ترتیب معیار BPM در موقعیت مکانی  $(i,j)$  از SA بصورت زیر تعریف می‌شود:

$$BPM(i, j) = \sum_k \sum_l XOR(\tilde{C}(x+k, y+l), \tilde{R}(x+i+k, y+j+l)) \quad (1)$$

که در آن XOR بیانگر عملگر منطقی یای انحصاری<sup>۱۳</sup> است،  $(l,k) \in \{\text{block area}\}$  است،  $C(x,y)$  و  $R(x,y)$  بترتیب سطوح خاکستری فریم مبنا و مرجع در مکان  $(x,y)$  است و  $\tilde{C}(x,y)$  و  $\tilde{R}(x,y)$  تبدیل یافته‌های این دو بشکل تک‌بیتی هستند. بردار حرکت تخمینی برای بلوک، متناظر با  $(i,j)$  از ناحیه جستجو است که در آن  $BPM(i,j)$  کمینه<sup>۱۴</sup> شود. تبدیل تک‌بیتی کننده تصویر به شکل زیر می‌تواند بیان شود:

$$\tilde{C}(i, j) = T[C(i, j)] \quad , \quad \tilde{R}(i, j) = T[R(i, j)] \quad (2)$$

تاکنون توابع مختلفی برای عملگر T پیشنهاد شده است. برخی از این توابع با بکارگیری یک مقدار آستانه، عمل تک‌بیتی کردن سطوح خاکستری را انجام می‌دهند [27,28].

کو<sup>۱۵</sup> و همکارانش [7] برای این تبدیل، استفاده از یکی از بیت‌های بیان باینری را پیشنهاد کردند. به این ترتیب که برای پیکسل‌های هشت بیتی، بیت وزن چهارم به عنوان بیان تک‌بیتی پیکسل استفاده شود. یعنی اگر سطح خاکستری در مکان  $(i, j)$  ام را بصورت زیر بنویسیم:

$$f(i, j) = b_7 2^7 + b_6 2^6 + \dots + b_0 2^0 \quad (3)$$

در این صورت تابع تبدیل تک‌بیتی کننده بصورت زیر خواهد بود:

$$T_1[f(i, j)] = b_4 \quad (4)$$

پیشنهاد آقای کو و همکارانش بدلیل عدم نیاز به انجام محاسبات برای تبدیل تصویر به شکل تک‌بیتی می‌تواند جاذبه بسیاری داشته باشد. البته کاربرد مورد نظر مقاله حاضر تخمین حرکت برای فشرده‌سازی ویدیو است در حالیکه پیشنهاد آقای کو در استفاده از بیت چهارم، برای تخمین حرکت در کاربرد تثبیت بوده است. آزمایش‌هایی که بر روی ۲۰ رشته تصویر تست انجام شد نشان می‌دهند که در کاربرد تخمین حرکت برای فشرده‌سازی، استفاده از بیت ششم، بطور متوسط نتیجه بهتری را ارائه می‌کند (جدول ۱). در این جدول با بکارگیری PSNR که میانگین مربعات خطای پیش‌بینی را به عنوان نوز استفاده می‌کند، معیارهای انطباق با هم مقایسه شده‌اند. PSNR (بر حسب dB) برای تصویری با سطوح خاکستری هشت بیتی و ابعاد  $Y \times X$  بصورت زیر تعریف می‌شود:

$$PSNR = 10 \log_{10} \left( \frac{255^2}{\frac{1}{X.Y} \sum_{x=1}^X \sum_{y=1}^Y (C(x,y) - \hat{C}(x,y))^2} \right) \quad (5)$$

که در آن  $C(x,y)$  بیانگر تخمین  $C(x,y)$  به کمک پیش‌بینی است. با توجه به این رابطه، هرچه پیش‌بینی فریم دقیق‌تر انجام شده باشد، مخرج رابطه (۵) کوچک‌تر شده و PSNR افزایش می‌یابد. ستون‌های دوم تا پنجم جدول (۱)، مقدار PSNR را برای روش آقای کو با بکارگیری بیت‌های مختلف چهارم تا هفتم نشان می‌دهد. برای هر رشته تصویر، مناسب‌ترین بیت، که بیشترین PSNR را ارائه می‌کند، با رنگ خاکستری مشخص شده است. ایراد اساسی که در روش آقای کو مشاهده می‌شود این است که بیت مناسب برای تخمین حرکت از رشته تصویری به رشته تصویر دیگر تغییر می‌کند. بر این اساس دو موضوع تعیین عامل این تغییر و ارائه راه حلی برای برطرف کردن این ایراد مورد توجه قرار گرفت. قسمت (۲-۱) به بررسی عوامل ایجادکننده مشکل پرداخته و قسمت (۲-۲) و (۲-۳) نیز به ارائه راه‌حلهایی برای رفع آن پرداخته است.

### ۱-۱- بررسی عوامل ایجاد مشکل در روش کو

با این سؤال که چرا در رشته تصاویر مختلف، بیت‌های مختلفی به عنوان بیت مناسب، بهترین PSNR را ارائه می‌کنند (جدول ۱)، بررسی‌ها ابتدا با رشته تصاویر مختلف انجام داده شد. برای پرداختن به جواب می‌توان از دو منظر موضوع را مورد توجه قرار داد، یکی از منظر اطلاعات درون یک فریم و دیگری از منظر تغییرات بین فریم‌ها. با توجه به پیکسل‌های یک فریم، قطعاً توزیع سطوح خاکستری یک تصویر از عوامل مهم تأثیرگذار بر پاسخ سؤال فوق است. مثلاً اگر تصویری با سطوح خاکستری ۸ بیتی وجود داشته باشد که تمام پیکسل‌های آن کمتر و یا بیشتر از ۱۲۸ باشد، در این صورت بیت هفتم هیچگونه اطلاعاتی نخواهد داشت که بتوان از آن برای ارزیابی انطباق استفاده کرد. از منظر بین فریمی فرض کنید که بین این دو فریم هیچگونه حرکتی وجود نداشته. در این حالت عامل تأثیرگذار بر انتخاب بیت مناسب، نویز بین دو فریم است. بدلیل کوچک بودن دامنه نویز، عمدتاً بیت‌های وزن پایین از آن تأثیر می‌پذیرد و باعث می‌شود که این بیت‌ها در ارزیابی انطباق قابل استفاده نباشند. از سوی دیگر، از آنجا که دقت روش تخمین حرکت با انطباق بلوک در حد پیکسل است، همیشه خطای اندازه‌گیری حداکثر نیم پیکسل وجود دارد. بدین ترتیب اگر بین دو فریم حرکتی کمتر از نیم پیکسل وجود داشته باشد، این حرکت صفر تخمین زده می‌شود. این امر خود موجب می‌شود که در محل بیشترین انطباق خطایی ناشی از همین امر بوجود آید. در اثبات این امر، گیرد<sup>۱۶</sup> [4] نشان داد که با کاهش دقت تخمین حرکت، واریانس خطا در محل انطباق افزایش می‌یابد. از عوامل دیگر خطا از منظر بین فریمی، می‌توان به تغییر روشنایی صحنه، خطای منظر دید<sup>۱۷</sup> و عوامل دیگری از این دست نیز اشاره کرد.

آنچه که در اینجا مورد ارزیابی قرار گرفته خطاهایی است که از منظر بین فریمی ایجاد می‌شوند. جهت بررسی این خطاها بصورت یک نویز جمع شونده مدل شده و در نسبت سیگنال به نویزهای مختلف، قابلیت انطباق مورد ارزیابی قرار گرفت. برای این منظور فریم اول هر یک از رشته تصویر ۲۰ گانه تست انتخاب و در سیگنال به نویزهای مختلف، شکل تابع خودهمبستگی<sup>۱۸</sup> مورد ارزیابی قرار گرفت. مطابق مرجع [7] برای این منظور تابع زیر برای اندازه خودهمبستگی بکار گرفته شد:

$$AC(m,n) = \frac{1}{X.Y} \sum XOR(\tilde{C}_k(x,y), \tilde{C}_k(x+m,y+n)) \quad (6)$$

که در آن  $X$  و  $Y$  ابعاد تصویر،  $\tilde{C}_k$  باینری شده یک فریم نویزی به روش کو (بیت  $k$  ام) و  $AC$  نیز معرف مقدار خودهمبستگی است.

در یک سیگنال به نویز خاص، بهترین بیت در معیار کو آن است که بیشترین گرادیان تابع خودهمبستگی را در اطراف محل بهترین انطباق (که البته در اینجا چون خود همبستگی محاسبه می‌شود، محل  $(0,0) = (m,n)$  خواهد بود) ارائه کند [7, 32]، تا بدین ترتیب با قوت بیشتری بتوان بر مناسب بودن بیت مربوطه برای تشخیص انطباق تأکید کرد. شکل (۲) گرادیان

تابع خودهمبستگی را در سیگنال به نویزهای مختلف برای اولین فریم از رشته تصویر Foreman نشان میدهد، لازم به ذکر است که برای دیگر رشته تصاویر نیز کم و بیش شکل مشابه شکل (۲) است. همانطور که از شکل برمی آید، در سیگنال به نویزهای بالا، بیتهای پایین تر، انتخاب مناسب تری برای معیار کو می توانند باشند. با افزایش نویز، قاعدتاً بیت های وزن پایین ارزش خود را از دست می دهند و بنابراین بیت های وزن بالاتر در فرایند تطبیق مؤثرتر عمل خواهند کرد. از آنجاییکه وجود نویز ناشی از عوامل مختلف، بالاخص نویز تصویربرداری و نویز ناشی از حرکت، اجتناب ناپذیر است، لذا بیت های وزن بالاتر عموماً نتیجه بهتری خواهند داشت. جدول (۱) نشان می دهد که در رشته تصاویر بیست گانه تست، بسته به میزان حرکت موجود در صحنه، بیت های پنجم تا هفتم، بهترین نتیجه را ارائه می کنند (ستون های دوم تا پنجم جدول را ببینید).

توجه به حرکات در صحنه های تصاویر تست و جدول (۱) ما را به این نتیجه می رساند که دلیل عمده در مناسب تر بودن بیت های وزن بالاتر، وجود حرکت بیشتر در صحنه و یا به عبارت دیگر پویا تر بودن صحنه است. یعنی هرچه میزان پویایی صحنه بیشتر بوده، بیت های بالاتر نتیجه بهتری را ارائه می کنند. نکته مهم این است که در صحنه های عمومی همیشه میزان پویایی صحنه ثابت نیست. در مواقعی ممکن است صحنه ساکن و در مواقع دیگر حرکت زیاد باشد. این امر موجب می شود که استفاده از یک بیت ثابت در تکبیتی کردن تصاویر همیشه نتیجه خوبی ندهد. برای رفع مشکل، دو راه حل ممکن است استفاده شود، یکی وفقی کردن انتخاب بیت با میزان حرکت در صحنه، بطوریکه با افزایش پویایی صحنه، بیت بالاتری برای تبدیل تصویر به شکل تکبیتی استفاده شود و دیگری اصلاح روش ارائه شده توسط آقای کو، بطوریکه با داشتن یک الگوریتم غیروفقی بتوان نتیجه بهتری در صحنه های با پویایی متفاوت کسب کرد. آنچه که در اینجا ارائه خواهد شد، روشهایی غیروفقی است که در ادامه به شرح آنها پرداخته خواهد شد.

### ۱-۲- انطباق چند صفحه بیت<sup>۱۹</sup> (MBPM)

همانطور که قبلاً ذکر شد، بسته به میزان پویایی صحنه، بیت های مختلف نتیجه بهتری را در الگوریتم کو ارائه می کنند. در صورت استفاده همزمان از چند بیت مختلف، این امکان فراهم می آید که چنین معیاری در صحنه های مختلف نتیجه خوبی ارائه کند. از طرفی چنین کاری موجب افزایش بار محاسباتی در مقام مقایسه با روش کو می شود. در روش MBPM پیشنهاد می شود که از چهار صفحه بیت ۴، ۵، ۶ و ۷ بطور همزمان استفاده شود. استفاده همزمان از این صفحات بیت کمک می کند تا در صحنه های با میزان پویایی مختلف، این معیار خوب عمل کند. برای آنکه از نظر محاسباتی این معیار در رده معیار کو قرار گیرد، می توان در هر یک از صفحات بیت، نمونه ها را به نسبت ۴ به یک کاهش داده و صفحات بیت را در هم ادغام کرد. حاصل کار بدین ترتیب است که برای چهار پیکسل مجاور هم، برای یک پیکسل بیت چهارم، برای پیکسل بعد بیت پنجم و به همین ترتیب برای پیکسل های بعد بیت های ششم و هفتم بیان باینری آن به عنوان تبدیل شده آن به تکبیتی انتخاب شوند. این امر را می توان بصورت زیر بیان کرد:

$$T_2[f(i, j)] = \begin{cases} b_7, & \text{if } [\text{mod}(i, 2) = 1 \text{ and } \text{mod}(j, 2) = 1] \\ b_6, & \text{if } [\text{mod}(i, 2) = 0 \text{ and } \text{mod}(j, 2) = 1] \\ b_5, & \text{if } [\text{mod}(i, 2) = 1 \text{ and } \text{mod}(j, 2) = 0] \\ b_4, & \text{if } [\text{mod}(i, 2) = 0 \text{ and } \text{mod}(j, 2) = 0] \end{cases} \quad (7)$$

که در آن  $1 \leq i \leq X$  و  $1 \leq j \leq Y$  است و مختصات پیکسل  $f$  را بیان می کند. شکل (۳) نتیجه چنین تبدیلی را نشان می دهد. همانطور که ملاحظه می شود، نسبت به روش کو هیچ بار محاسباتی در فرایند تبدیل تصاویر به شکل تکبیتی ایجاد نشده، بلکه صرفاً یک انتخاب بیت وجود دارد. همانطور که در شکل نشان داده شده است، در روش کو برای تکبیتی کردن تصویر، مثلاً با بیت چهارم، فقط بیت چهارم استفاده می شود، در حالیکه در MBPM بسته به موقعیت پیکسل، بیتهای مختلف چهارم تا هفتم استفاده شده است.

### ۱-۳- انطباق چند صفحه بیت وزندار<sup>۲۰</sup> (WMBPM)

از آنجاییکه تفاوت در بیتهای وزن بالاتر دو مقدار، در اکثر موارد گویای تفاوت بیشتر بین آن دو است، بنابراین در صورت

در نظر گرفتن وزن مناسب برای بیت‌های مختلف در رابطه (۷)، کارایی معیار انطباق افزایش خواهد یافت، چرا که تفاوت در بیت‌های وزن بالاتر تاثیر بیشتری بر مقدار نتیجه ارزیابی انطباق می‌گذارند. در این حال با توجه به اینکه وزن‌ها ضربی از ۲ هستند، پیاده‌سازی آنها براحتی و به کمک استفاده از شیفت به چپ قابل پیاده‌سازی است. در این حالت چهار تابع تبدیل سطوح خاکستری هشت‌بیتی به تک‌بیتی را می‌توان به صورت زیر تعریف کرد:

$$T_4[f(m,n)] = \begin{cases} b_4, & \text{if } [\text{mod}(m,2) = 0 \text{ and } \text{mod}(n,2) = 0] \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$T_5[f(m,n)] = \begin{cases} b_5, & \text{if } [\text{mod}(m,2) = 1 \text{ and } \text{mod}(n,2) = 0] \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

$$T_6[f(m,n)] = \begin{cases} b_6, & \text{if } [\text{mod}(m,2) = 0 \text{ and } \text{mod}(n,2) = 1] \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$$T_7[f(m,n)] = \begin{cases} b_7, & \text{if } [\text{mod}(m,2) = 1 \text{ and } \text{mod}(n,2) = 1] \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

با این توابع، معیار انطباق WMBPM بصورت زیر خواهد بود:

$$\begin{aligned} WMBPM(i,j) &= BPM_{\tilde{C}_4, \tilde{R}_4}(i,j) + 2BPM_{\tilde{C}_5, \tilde{R}_5}(i,j) \\ &\quad + 4BPM_{\tilde{C}_6, \tilde{R}_6}(i,j) + 8BPM_{\tilde{C}_7, \tilde{R}_7}(i,j) \\ &= BPM_{\tilde{C}_4, \tilde{R}_4}(i,j) + 2(BPM_{\tilde{C}_5, \tilde{R}_5}(i,j) \\ &\quad + 2(BPM_{\tilde{C}_6, \tilde{R}_6}(i,j) + 2(BPM_{\tilde{C}_7, \tilde{R}_7}(i,j)))) \end{aligned} \quad (12)$$

که در آن  $\tilde{C}_k$  و  $\tilde{R}_k$  بترتیب تبدیل شده‌های C و R تحت تابع  $T_k$  است.

## ۲- کلاس‌بندی تفاضل پیکسل‌ها

کلاس‌بندی تفاضل پیکسل‌ها از الگوریتم‌های شناخته‌شده‌ای است که برای کاهش بار محاسباتی طراحی شده است [5]. در معیار انطباق PDC، تفاضل پیکسل‌ها بر اساس مقدارشان به نقاط منطبق و غیرمنطبق تقسیم شده و در هر محلی که نقاط منطبق بیشتر باشند، مفهوم انطباق بیشتر را برای آن می‌توان قائل شد. در الگوریتم پایه ارائه شده برای PDC که توسط آقای غروی و همکارش [5] ارائه شده است، با استفاده از یک مقدار آستانه، مجموعه نقاط به نقاط منطبق و غیرمنطبق کلاس‌بندی می‌شوند. معیار PDC در محل  $(i, j)$  از ناحیه جستجو بصورت زیر تعریف می‌شود:

$$PDC(i,j) = \sum_k \sum_l G_{i,j}(k,l) \quad (13)$$

که در آن با یک مقدار آستانه تعریف شده  $t_g$ ، مقدار  $G_{i,j}$  نیز بصورت زیر تعریف شده است:

$$G_{i,j}(k,l) = \begin{cases} 1, & \text{if } |C(x+k,y+l) - R(x+i+k,y+j+l)| \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

که در آن  $(l,k) \in \{block\ area\}$  و  $C(x,y)$  و  $R(x,y)$  نیز بترتیب سطوح خاکستری فریم مبنا و مرجع در مکان  $(x,y)$  است. بردار حرکت تخمینی برای بلوک، متناظر با  $(i,j)$  از ناحیه جستجو است که در آن  $PDC(i,j)$  بیشینه<sup>۲۱</sup> شود. معیار PDC

نسبت به معیار SAD عملیات انباشت را در رابطه (۱۳) ساده می‌کند، بطوریکه برای آن می‌توان از یک شمارنده استفاده کرد. البته در عوض این ساده‌سازی فرایند دوسطحی کردن (اعمال رابطه ۱۴) اضافه شده است. اشکال اساسی که در PDC وجود دارد این است که مقدار بهینه برای  $t_g$  وابسته به تصویر است [3]. نلسون<sup>۲۲</sup> [6] پس از انجام بررسی‌هایی دریافت که مقادیر آستانه ۵، ۱۵ و ۲۵ بسته به تصاویر می‌توانند مقادیر مناسبی باشند. چن<sup>۲۳</sup> و همکارانش [3] پیشنهاد کردند که در صحنه‌های با پویایی کم از آستانه کم و در صحنه‌های با پویایی زیاد از آستانه زیاد استفاده شود. شکل (۴) تاثیر تغییر مقدار آستانه بر میزان PSNR را برای رشتع تصویر Foreman نشان می‌دهد. قاعدتاً مقدار بهینه  $t_g$  معادل مقداری است که PSNR را ماکزیمم کند. ستون دوم جدول (۲) مقدار آستانه بهینه را برای ۲۰ رشته تصویر متفاوت نشان می‌دهد. همانطور که مشاهده می‌شود، بسته به میزان پویایی صحنه در رشته تصویر ورودی، مقدار بهینه  $t_g$  نیز تغییر می‌کند. بنابراین در صورت تغییر خودکار مقدار آستانه، متناسب با حرکت موجود در صحنه، می‌توان نتایج بدست آمده در الگوریتم تخمین حرکت را بهبود بخشید. بخش بعد برای این منظور روشی را ارائه می‌دهد که در آن مقدار آستانه بصورت وقتی تعیین می‌شود.

## ۲-۱- کلاس‌بندی وقتی تفاضل پیکسل‌ها<sup>۲۴</sup> (APDC)

همانطور که در شرح PDC آمد، این الگوریتم، کلاس‌بندی تفاضل پیکسل‌ها را مورد توجه قرار داده است. الگوریتم پیشنهادی غروی از یک مقدار آستانه ثابت برای کلاس‌بندی تفاضل پیکسل‌ها استفاده می‌کند. آنچه در اینجا پیشنهاد می‌شود استفاده از کلاس‌بندی پارامتریک برای این منظور است، بطوریکه بسته به رشته تصویر ورودی مقدار آستانه بکار رفته در کلاس‌بندی تغییر کند [31]. در روش ارائه شده، ابتدا مطابق روش‌های کلاس‌بندی آماری، برای مقدار تفاضل پیکسل‌ها در نقاط منطبق و غیرمنطبق دو تابع توزیع احتمال نرمال با میانگین صفر مانند شکل (۵) در نظر گرفته شده است. سپس با توجه به این دو تابع توزیع، مقدار آستانه مناسب برای کلاس‌بندی فوق تعیین شده است. برای یک محدوده جستجوی  $(N+2w+1) \times (N+2w+1)$ ، احتمال وقوع کلاس مقادیر منطبق به غیرمنطبق برابر  $\frac{1}{(2w+1)^2 - 1}$  است. از طرفی به دلیل کم بودن نقاط منطبق لازم است که هزینه بالایی برای کلاس‌بندی غلط<sup>۲۵</sup> نقاط منطبق به عنوان نقاط غیر منطبق در نظر گرفته شود. هر چه محدوده جستجو بزرگتر شود، بایستی این هزینه را نیز افزایش داد چرا که نقاط منطبق کسر کمتری از نقاط مورد ارزیابی را تشکیل خواهند داد. در اینجا نسبت هزینه کلاس‌بندی غلط کلاس‌ها برابر عکس احتمال وقوع کلاس‌ها در نظر گرفته شده است. بر این اساس، برای به حداقل رساندن هزینه کلاس‌بندی غلط، بایستی مقدار آستانه  $t_g$  مطابق شکل (۵) در محل تقاطع این دو توزیع انتخاب شود [30]، بدین ترتیب:

$$\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{t_g^2}{2\sigma_1^2}} = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{t_g^2}{2\sigma_2^2}} \quad (15)$$

$$\Rightarrow t_g = \left( 2 \log \left( \frac{\sigma_1}{\sigma_2} \right) / \left( \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) \right)^{\frac{1}{2}} \quad (16)$$

که در آن مقادیر  $\sigma_1$  و  $\sigma_2$  بترتیب انحراف استاندارد کلاس نقاط منطبق و غیرمنطبق است. در صورتیکه  $\sigma_1^2 \ll \sigma_2^2$  باشد، رابطه (۱۶) می‌تواند به شکل زیر ساده شود:

$$t_g \approx \sigma_1 \left( 2 \log \left( \frac{\sigma_1}{\sigma_2} \right) \right)^{\frac{1}{2}} \quad (17)$$

در الگوریتم پیشنهادی، در هر فریم مقادیر واریانس‌ها تخمین زده شده و مقدار آستانه بدست آمده از این مقادیر در فریم بعدی بکار گرفته می‌شود. البته آزمایش‌ها نشان می‌دهند که مادامیکه صحنه عوض نشود عموماً مقدار آستانه تغییر زیادی نکرده و لذا می‌تواند مادامیکه صحنه ثابت است مقدار آستانه با نرخ کمتری تخمین زده شود تا بار محاسبات باز هم کاهش



یابد. این مساله را در بخش (۵-۲) و در شرح شکل (۷) خواهید دید.

واریانس تابع توزیع احتمال تفاضل نقاط غیرمنطبق قاعدتاً تابعی از میزان همبستگی هر پیکسل به پیکسل‌های مجاور است. از آنجاییکه محدوده جستجو در اطراف محل بیشترین انطباق قرار دارد، هر چه همبستگی پیکسل‌ها به پیکسل‌های مجاور بیشتر باشد، مقدار تفاضل پیکسل‌های مبنا با پیکسل‌های نامزد انطباق کمتر خواهد بود. از طرفی همانطور که قبلاً ذکر شد، عواملی چون نویز تصویربرداری، حرکت، تغییر روشنایی، تغییر در منظر دید و ... می‌توانند موجب شوند تا در الگوریتم انطباق بلوک و در محل بیشترین انطباق، خطا وجود داشته باشد. از آنجاییکه بعضی از این عوامل با زمان تغییر می‌کنند، بنابراین پارامترهای تابع توزیع احتمال برای نقاط منطبق نیز با زمان تغییر خواهند کرد. لذا لازم خواهد بود که در کلاس‌بندی پارامتریک، مرتباً واریانس کلاس‌ها بصورت پویا تخمین زده شود. از طرفی، فرایند تخمین واریانس، خود ایجاد بار محاسباتی می‌کند، که این امر با ساده‌سازی معیار انطباق در تضاد است. بنابراین لازم است که این تخمین با بار محاسبات کم صورت گیرد. برای این منظور تنها یک پیکسل از هر بلوک برای تخمین این دو پارامتر بکار گرفته شد. بدین ترتیب، در ناحیه جستجو به ازای هر بلوک نامزد انطباق، یک نقطه برای تخمین واریانس نقاط غیرمنطبق بکار گرفته شد و تنها یک نقطه نیز از محل بهترین انطباق برای تخمین واریانس نقاط منطبق استفاده شد. در این حال به ازاء هر ارزیابی انطباق، سه عمل یعنی یک محاسبه مربع (برای مقدار تفاضل)، یک انباشت<sup>۲۶</sup> مربعات تفاضل (برای محاسبه مجموع مربعات تفاضل نقاط غیرمنطبق) و یک انباشت شرطی مربع تفاضل در صورتیکه مربوط به محل بیشترین انطباق باشد، بایستی اضافه بر PDC انجام شود. این مساله به معنای افزایش  $\frac{3}{(N+2 \times w+1)^2}$  عمل بر پیکسل در APDC نسبت به PDC است.

درست است که وقتی کردن PDC موجب افزایش جزئی در محاسبات شده اما می‌توان با روشهایی و از جهتی دیگر محاسبات را کاهش داد. همانطور که قبلاً ذکر شد، PDC نسبت به SAD دارای عملیات اضافی اعمال آستانه بر روی تفاضل پیکسل‌ها است که خود نیاز به مقایسه‌کننده دارد. در صورتیکه مقدار آستانه توانی از ۲ انتخاب شود، عمل مقایسه می‌تواند به شکل ساده‌ای انجام شود. بدین ترتیب که اگر از  $2^i$  به عنوان مقدار آستانه استفاده شود، یک بودن هر کدام از بیت‌های وزن  $i$  و بالاتر، بیانگر بزرگتر یا مساوی بودن قدرمطلق تفاضل نسبت به مقدار آستانه است. بنابراین در این حالت سخت‌افزار مقایسه‌کننده رابطه (۱۴) صرفاً یک دریچه منطقی NOR روی بیت‌های وزن  $i$  و بالاتر خواهد بود. بدین ترتیب پیاده‌سازی مقایسه‌کننده به شکلی بسیار ساده خواهد بود. بطوریکه اگر  $D_{i,j}(k,l) = |C(x+k, y+l) - R(x+i+k, y+j+l)|$  فرض شود، در این صورت می‌توان رابطه (۱۴) را بصورت زیر ساده کرد:

$$G_{i,j}(k,l) = \text{NOR}_i(D_{i,j}(k,l)) \quad (18)$$

که در آن  $\text{NOR}_i(D)$  بیانگر عملگر NOR منطقی بر روی بیت‌های وزن  $t$  و بالاتر مقدار  $D$  است. بنابراین برای این اصلاح پیشنهادی کافی است پس از بدست آوردن مقدار  $t_g$ ، این مقدار را به نزدیکترین مقداری که توانی از ۲ باشد گرد کرد. حاصل کار پس از این اصلاح، از این پس  $\text{APDC}^+$  نامیده می‌شود.

### ۳- مقایسه سخت‌افزارهای مورد نیاز

جهت روشن شدن میزان توانایی الگوریتمهای فوق در ساده‌سازی سخت‌افزار، بخش حاضر با مبنا قرار دادن معماری آرایه خطی مبتنی بر بردار حرکت<sup>۲۷</sup> [2]، مقایسه‌ای اجمالی بین سخت‌افزار دو دسته روش فوق یعنی روشهای انطباق صفحه بیت و کلاس‌بندی تفاضل پیکسل‌ها انجام می‌دهد. شکل (۶) شمایی از این معماری را برای پیاده‌سازی معیار انطباق SAD نشان می‌دهد. این سخت‌افزار برای بلوکهای  $16 \times 16$ ، قادر است که ۱۶ مقدار ارزیابی انطباق را بصورت موازی محاسبه کند [2]. همانطور که مشاهده می‌شود هر یک از عناصر پردازشی<sup>۲۸</sup> در این معماری تشکیل شده است از یک رجیستر (D)، یک مالتی‌پلکسر (mux)، یک عنصر محاسبه‌کننده قدرمطلق تفاضل و نهایتاً یک عنصر محاسبه انباشت (Acc).

با تغییر معیار انطباق از SAD به PDC، یک عنصر آستانه‌گیر بین محاسبه قدرمطلق تفاضل و عنصر انباشت قرار می‌گیرد و عنصر انباشت نیز به صورت یک شمارنده ساده می‌شود. استفاده از مقدار آستانه‌ای برابر با توانی از ۲ موجب می‌شود که

عنصر آستانه‌گیر به یک دریچه منطقی NOR ساده تبدیل شود. بدین ترتیب نسبت به استفاده از معیار SAD، عنصر انباشت با یک دریچه منطقی NOR و یک شمانده جایگزین شده است. در این حالت یک عنصر پردازشی ساده نیز بایستی برای وفقی کردن PDC به آن اضافه شود.

در مورد استفاده از معیارهای انطباق صفحه بیت، در عنصر پردازشی شکل (۶)، رجیستر D و مالتی‌پلکسر mux به حالت تک بیتی ساده تبدیل شده و عنصر محاسبه قدرمطلق تفاضل نیز با یک دریچه منطقی XOR جایگزین خواهد شد. بعلاوه Acc نیز با یک شمارنده جایگزین می‌شود. استفاده از معیار MBPM ایجاب می‌کند که دریچه منطقی XOR با چهار XOR و یک مالتی‌پلکسر ساده ۴ به ۱ جایگزین شود. در معیار WMBPM نیز نسبت به MBPM، بایستی شمارنده جایگزین شده بجای Acc شمارنده‌ای با قابلیت افزایش ۱، ۲، ۴ و ۸ واحد باشد تا بدین ترتیب بسته به بیت‌های با وزن‌های مختلف میزان افزایش آن تغییر کند.

لازم به ذکر است که سخت‌افزار ارائه شده در شکل (۶)، یکی از معماری‌های عمومی است که در پیاده‌سازی سخت‌افزارهای تخمین حرکت بکار می‌رود و در مرجع [2] نیز بطور کامل شرح داده شده است. این سخت‌افزار پیاده‌سازی عملی نشده است بلکه بدلیل عمومیتش از آن به عنوان معیاری برای نشان دادن کاهش پیچیدگی سخت‌افزار در استفاده از معیارهای ارائه شده مقاله استفاده شده است.

اطلاعات ارائه شده در اینجا به همراه نتایج آمده در بخش بعد، کمک می‌کند تا بتوان بسته به کارایی مورد انتظار و میزان پیچیدگی سخت‌افزار، بین روش‌های ارائه شده یکی را انتخاب نمود.

## ۴- نتیجه آزمایش‌ها

برای ارزیابی روش‌های ارائه شده، ۲۰ رشته تصویر شناخته شده که عموماً در کاربردهای فشرده‌سازی تصویر بکار گرفته می‌شوند، استفاده شده است. فریم اول هر یک از این رشته تصاویر در شکل (۷) نشان داده شده است. این رشته تصاویر دارای رزولوشن ۱۷۶ در ۱۴۴ پیکسل و نرخ فریم، ۱۰ فریم بر ثانیه هستند. در این آزمایشها از  $w=7$ ،  $N=8$  و جستجوی کامل استفاده شده و نتایج بدست‌آمده در اینجا ارائه می‌شوند.

### ۴-۱- نتایج آزمایش‌های MBPM و WMBPM

جدول (۱) نتایج بدست آمده برای ارزیابی معیارهای MBPM و WMBPM را ارائه می‌کند. جهت ارزیابی، همانطور که قبلاً نیز گفته شد، از PSNR استفاده شده است. بنابراین مقادیر داده شده در جدول، مقادیر مختلف PSNR برای رشته تصاویر مختلف و به ازای معیارهای مختلف انطباق است.

ابتدا برای ارزیابی روش کو، بیت‌های مختلف چهارم تا هفتم برای تک‌بیتی کردن تصویر بکار گرفته شد (نتایج ارائه شده در ستون‌های دوم تا پنجم). با توجه به این ستون‌ها مشاهده می‌شود که بسته به میزان پویایی صحنه، مناسب‌ترین بیت برای بکارگیری در این روش تغییر می‌کند، مثلاً در رشته تصویر Grandma، که جزئیات صحنه دارای حرکت کم است، بیت پنجم بهترین نتیجه را ارائه می‌کند، در حالیکه برای رشته تصویر Football که صحنه از پویایی بالایی برخوردار است، بیت هفتم بهترین نتیجه را ارائه می‌کند. در این جدول و در ستون‌های مربوط به روش کو، بهترین بیت برای هر رشته تصویر با رنگ خاکستری مشخص شده است. بر اساس سطر آخر این جدول مشاهده می‌شود که بطور متوسط بیت ششم نتیجه مناسب‌تری را ارائه می‌کند.

اما وقتی معیار MBPM را بکار می‌بریم، با توجه به مقادیر PSNR ارائه شده در ستون ششم، در عین اینکه می‌دانیم این روش نسبت به روش کو بار محاسباتی را افزایش نمی‌دهد، مشاهده می‌شود که برای کلیه رشته تصاویر، مقدار PSNR بدست آمده از این روش همیشه از روش کو با بهترین بیت (مشخص شده با رنگ خاکستری) بالاتر بوده است. بنابراین مستقل از اینکه میزان حرکت در صحنه چقدر بوده است، براحتی می‌توان MBPM را بکار گرفته و نتیجه‌ای بهتر از روش کو بدست آورد. با دادن وزن به بیت‌های مختلف، یعنی روش WMBPM، که براحتی و با اعمال چهار شیفت منطقی به چپ قابل پیاده‌سازی است، با مراجعه به نتایج ارائه شده در ستون هفتم، مشاهده می‌شود که باز هم مقدار PSNR به میزان خوبی افزایش

پیدا کرده است. بطور متوسط در صورت استفاده از روش MBPM، مقدار PSNR نسبت به روش کو حدود ۰/۶٪ افزایش می‌یابد. این مقدار برای روش WMBPM حدود ۰/۹٪ است. به منظور مقایسه PSNRهای بدست آمده از روش کو، MBPM و WMBPM، مقدار PSNR برای وقتی که از معیارهای انطباق SSD و SAD استفاده می‌شوند نیز در ستون‌های هشتم و نهم آمده‌اند. همانطور که مشاهده می‌شود، روش WMBPM نتایج نزدیکی را به SSD ارائه می‌کند.

## ۴-۲- نتایج آزمایش‌های APDC

جدول (۲) نتایج بدست آمده در ارزیابی APDC را ارائه می‌کند. برای ارزیابی، ابتدا به ازای مقادیر آستانه مختلف از ۲ تا ۴۰ مقدار PSNR برای هر رشته تصویر بدست آمد. سپس برای هر رشته تصویر بهترین مقدار آستانه که معادل بدست آوردن بیشترین متوسط PSNR بود تعیین شد، که این مقادیر آستانه بهینه و PSNR مربوطه بترتیب در ستون‌های دوم و سوم جدول (۲) ارائه شده است. سپس با اعمال الگوریتم APDC نیز مقدار متوسط PSNR برای هر یک از رشته تصاویر بدست آمد که مقادیر آن در ستون چهارم جدول (۲) آمده است. برای درک بهتر از نتیجه بدست آمده، نسبت PSNR بدست آمده از روش PDC و فقی به PSNR بدست آمده از بهترین مقدار آستانه بدست آمد، که بصورت درصد در ستون پنجم ارائه شده است. نتایج نشان می‌دهند که الگوریتم و فقی ارائه شده در این مقاله در عین اینکه بار محاسبات اضافی زیادی را تحمیل نکرده، توانسته است نتایج بسیار نزدیکی به بهترین مقدار آستانه را ارائه نماید. مقدار متوسط این درصدها نشان می‌دهد که بطور میانگین این الگوریتم کارایی بیش از ۹۹٪ بهینه را ارائه می‌نماید. این کارایی برای  $w=8$  تنها با افزایش ۰/۰۵۶ عمل بر پیکسل نسبت به PDC بدست آمده است. پس از اصلاح دوم، یعنی استفاده از مقادیر آستانه‌ای که توانی از دو باشند ( $APDC^+$ )، مقادیر ستون‌های چهارم و پنجم بصورت ستون‌های ششم و هفتم در خواهند آمد. ملاحظه می‌شود که این اصلاح در عین کاهش پیچیدگی سخت‌افزار مورد نیاز، تنها ۰/۰۶٪ کاهش در نسبت PSNR آن به PSNR آستانه بهینه ایجاد کرده است که مقدار بسیار ناچیزی است.

نکته مهمی که بایستی ذکر شود این است که در بسیاری از صحنه‌ها میزان پویایی صحنه مرتباً تغییر می‌کند، که در این حالت استفاده از مقدار آستانه ثابت می‌تواند باعث کاهش کارایی شود. مثلاً یک پخش خبر را در نظر بگیرید، در برخی از صحنه‌ها صرفاً گوینده در صحنه وجود داد که دارای حرکت بسیار کمی است، در این حالت قاعداً بایستی از یک مقدار آستانه کوچک استفاده شود. لحظه‌ای بعد ممکن است گزارشی با یک صحنه پویا پخش شود. قاعداً در این حالت بایستی از مقدار آستانه بزرگتری استفاده شود. از آنجاییکه الگوریتم و فقی ارائه شده قادر است مقدار آستانه را متناسب با صحنه تغییر دهد، در چنین وضعیتی خواهد توانست کارایی خوبی را ارائه نماید. برای ارزیابی این وضعیت، دو رشته تصویر Lady و Flower را پشت سر هم قرار داده و به عنوان یک رشته تصویر فرض شد. مقدار آستانه بهینه برای این دو رشته تصویر، بترتیب ۶ و ۳۲ است. الگوریتم APDC برای این رشته تصویر بکار گرفته شده و تغییرات مقدار آستانه بر حسب فریم رسم شد (شکل ۸). این دو رشته تصویر هر یک شامل ۱۶ فریم است. انتظار می‌رود که تا فریم شانزدهم مقدار آستانه حدود ۶ و بعد از آن تقریباً ۳۲ باشد. همانطور که در شکل مشاهده می‌شود، الگوریتم APDC براحتی توانسته است مطابق حرکت در صحنه مقدار مناسب آستانه را انتخاب نماید (الگوریتم  $APDC^+$  این مقادیر آستانه را بصورت ۸ و ۳۲ استفاده خواهد کرد). در این شکل O-Tr بیانگر مقدار آستانه بهینه برای این رشته تصویر دو قطعه‌ای و A-Tr نیز مقدار آستانه بدست آمده از الگوریتم APDC است. تنها مساله‌ای که وجود دارد، افزایش ناگهانی مقدار آستانه در فریم ۱۶ است که باعث افت زیادی در PSNR می‌شود. این امر ناشی از تغییر ناگهانی صحنه است. البته استفاده از یک مقدار ثابت برای آستانه نیز در چنین رشته تصویری باعث افت ناگهانی PSNR در لحظه تغییر صحنه می‌شود.

## ۵- نتیجه‌گیری

در این مقاله سه روش جدید برای ارزیابی انطباق بین دو تصویر معرفی شده است. هر سه این روش‌ها در پیاده‌سازی سخت‌افزاری می‌توانند به ساده‌سازی سخت‌افزار کمک زیادی بکنند. دو روش اولیه بنام‌های MBPM و WMBPM هستند که هر دو از روشهای انطباق صفحه بیت محسوب شده و بر مبنای روشی هستند که توسط آقای کو ارائه شده است. ابتدا این روش

روش که خود برای کاربرد تثبیت الکترونیکی رشته تصاویر دیجیتالی ارائه شده است، برای کاربرد فشرده‌سازی ویدیو تصحیح شده و سپس دو روش ارائه شده در مقاله با آن مقایسه شده‌اند. آزمایش‌ها نشان می‌دهد که این دو روش مستقل از میزان حرکت موجود در صحنه می‌توانند نتیجه‌ای بهتر از روش کو ارائه کنند. این دو روش می‌توانند مقدار PSNR را بترتیب حدود ۶٪ و ۹٪ افزایش دهند.

روش دیگر ارائه شده یک روش PDC وفقی است. این روش با خصوصیت وفقی بودن قادر است مقدار آستانه بکار رفته در روش PDC را متناسب با پویایی و یا به عبارتی میزان حرکت در صحنه اصلاح نماید. آزمایش‌ها نشان داده‌اند که با بکارگیری این روش می‌توان کارایی در حد بیش از ۹۹٪ آنچه از روش PDC با بهترین مقدار آستانه بدست می‌آید را اخذ کرد.

## زیر نویس‌ها

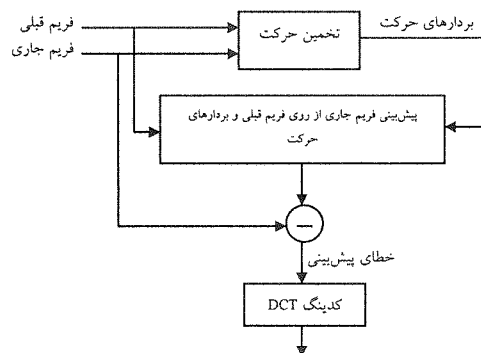
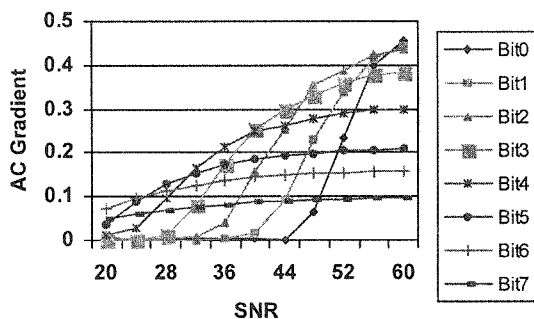
- |  |   |
|--|---|
| 1-Hybrid   | 15-Ko                                       |
| 2-Redundancy                                       | 16-Girod                                    |
| 3- Electronic Digital Image Stabilization          | 17-Parallax                                 |
| 4- تفاوت بین یک فریم با مقدار پیش‌بینی شده برای آن | 18-Autocorrelation                          |
| 5- Block Matching Algorithm                        | 19-Multiple Bit-Plane Matching              |
| 6-Search Area                                      | 20-Weighted Multiple Bit-Plane Matching     |
| 7- Full Search                                     | 21-Maximize                                 |
| 8-Sum of Square Difference                         | 22-Nelson                                   |
| 9-Sum of Absolute Difference                       | 23-Chan                                     |
| 10-Pel Difference Classification                   | 24-Adaptive Pixel Difference Classification |
| 11- Mobile Video Phone                             | 25- Misclassification                       |
| 12-Bit-Plane Matching                              | 26-Accumulation                             |
| 13-Exclusive OR                                    | 27-Motion-vector-based Linear Array         |
| 14- Minimize                                       | 28-Processing Element                       |

جدول (۱) مقدار PSNR پیش‌بینی فریم برای معیارهای انطباق مختلف، در ۲۰ رشته تصویر متفاوت. سلولهای خاکستری نشان‌دهنده بهترین جوابها در روش Ko است.

Sequence	Ko Method				MBP M	WMBP M	SSD	SAD
	Bit4	Bit5	Bit6	Bit7				
Akiyo	29.66	31.67	29.58	22.68	33.91	34.17	38.14	38.03
Bus	17.03	16.90	19.85	19.86	20.28	20.74	24.19	24.06
Carphone	22.74	24.81	24.97	23.89	27.77	28.72	32.74	32.56
Claire	25.37	28.28	29.63	23.92	32.14	32.94	37.07	36.96
Coastguard	20.76	22.72	25.22	21.77	25.33	25.95	28.96	28.86
Container	24.03	24.81	30.56	25.11	30.18	31.47	34.93	34.91
Flower	14.77	16.40	19.12	19.41	18.22	19.73	22.93	22.84
Football	17.86	18.50	19.51	20.04	20.63	21.30	23.64	23.47
Foreman	22.48	23.38	26.31	24.57	27.08	27.72	30.69	30.54
Grandma	28.10	32.16	28.75	23.66	32.92	33.29	36.60	36.51
Hall-Monitor	31.14	31.84	28.09	23.89	32.97	32.57	37.83	37.75
Lady	29.87	32.35	33.06	26.64	34.66	34.98	38.45	38.29
Mobilecalendar	15.69	17.77	20.14	21.79	19.90	21.20	24.72	24.69
New	24.50	26.36	27.44	21.57	28.37	29.30	32.71	32.53
Salesman	26.94	28.64	28.55	21.45	29.98	30.28	34.30	34.15
Silent	25.75	26.41	27.52	23.84	29.61	30.03	33.01	32.82
Stefan	17.23	19.19	22.49	23.07	21.57	22.91	25.77	25.70
Suzie	26.41	28.37	26.71	24.92	30.30	30.81	34.06	33.92
Table Tennis	20.89	22.00	23.51	22.59	24.37	25.22	28.01	27.83
Trevor	21.02	23.70	24.30	22.91	26.88	27.61	31.61	31.42
Average	23.11	24.84	25.77	22.88	27.35	28.05	31.52	31.39

جدول (۲) نتایج ارزیابی APDC برای رشته تصاویر تست.

Sequence	Opt. Tr.	Opt. PDC's PSNR	APDC's PSNR	PSNR APDC/OP DC	APDC <sup>++</sup> 's PSNR	PSNR APDC <sup>++</sup> /OPD C
Akiyo	4	36.95	36.24	98.1%	36.30	98.2%
Bus	28	23.20	23.09	99.5%	23.10	99.6%
Carphone	12	30.80	30.69	99.6%	30.70	99.7%
Claire	4	35.40	34.91	98.6%	35.01	98.9%
Coastguard	16	28.15	28.09	99.8%	28.15	100%
Container	4	34.26	33.30	97.2%	33.09	96.6%
Flower	32	22.42	22.41	100%	22.41	100%
Football	31	22.69	22.66	99.9%	22.67	99.9%
Foreman	15	29.19	29.03	99.5%	28.98	99.3%
Grandma	7	35.79	35.79	100%	35.79	100%
Hallmonitor	5	37.06	36.23	97.8%	36.20	97.7%
Lady	6	37.04	36.72	99.1%	36.83	99.4%
Mobilecal.	25	24.23	24.16	99.7%	24.07	99.3%
News	15	30.95	30.83	99.6%	30.81	99.5%
Salesman	14	32.95	32.75	99.4%	32.66	99.1%
Silent	12	31.48	31.18	99.0%	31.20	99.1%
Stefan	23	25.17	25.07	99.6%	25.04	99.5%
Suzie	8	32.29	32.29	100%	32.13	99.5%
Tennis	12	26.18	25.89	98.9%	25.93	99.1%
Trevor	13	29.42	29.25	99.4%	29.17	99.2%
Average				99.24%		99.18%



شکل (۲) گرادین خودهمبستگی در اطراف مینیمم آن برای بیت‌های مختلف و در سیگنال به نوزدهای مختلف، در رشته تصویر Foreman.

شکل (۱) شکل عمومی یک کد کننده ویدیو از نوع پیوندی [2].

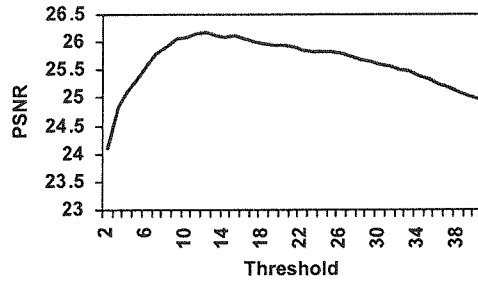
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4
b4	b4	b4	b4	b4	b4	b4	b4

(a)

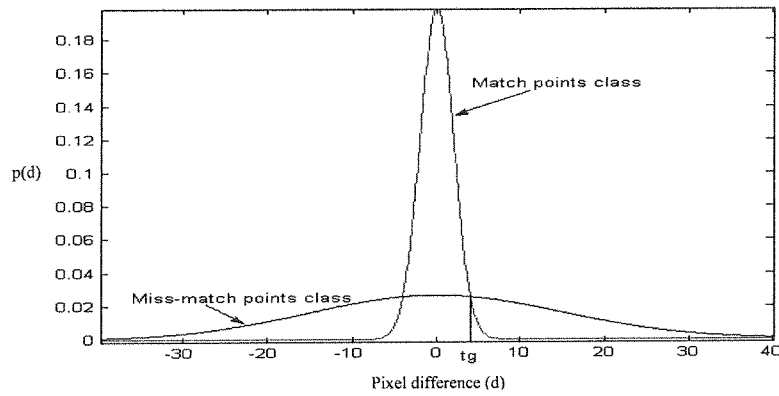
b4	b5	b4	b5	b4	b5	b4	b5
b6	b7	b6	b7	b6	b7	b6	b7
b4	b5	b4	b5	b4	b5	b4	b5
b6	b7	b6	b7	b6	b7	b6	b7
b4	b5	b4	b5	b4	b5	b4	b5
b6	b7	b6	b7	b6	b7	b6	b7
b4	b5	b4	b5	b4	b5	b4	b5
b6	b7	b6	b7	b6	b7	b6	b7

(b)

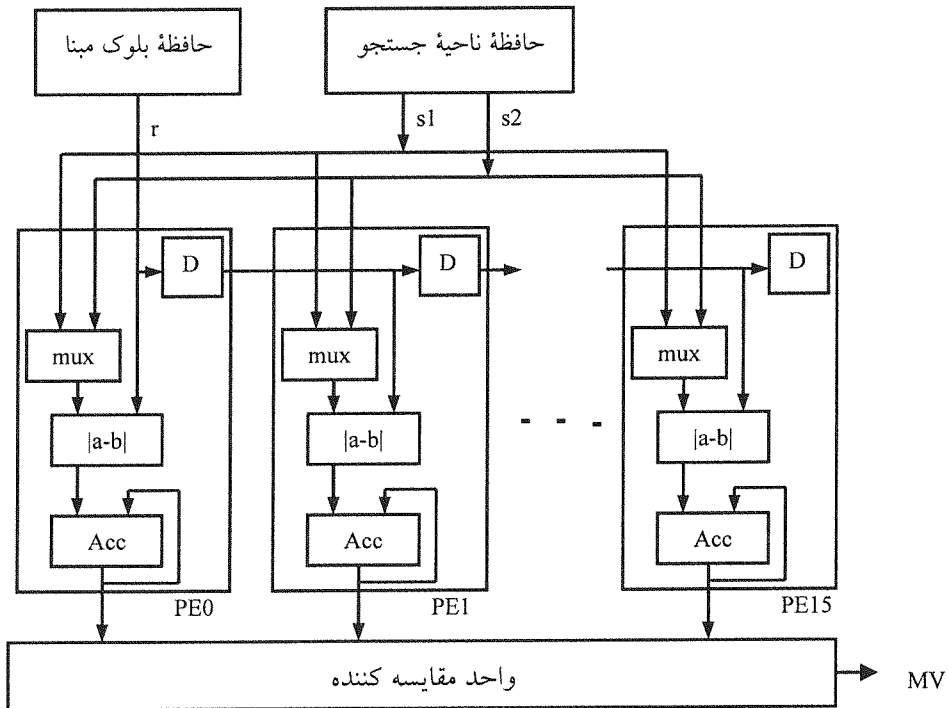
شکل (۳) تبدیل شده یک بلوک ۸×۸ از یک تصویر، (a) با استفاده از روش کو و با انتخاب بیت چهارم، و (b) با استفاده از روش MPBM.



شکل (۴) تغییرات نویز پیش‌بینی بر حسب مقدار آستانه در الگوریتم PDC برای رشته تصویر Foreman.



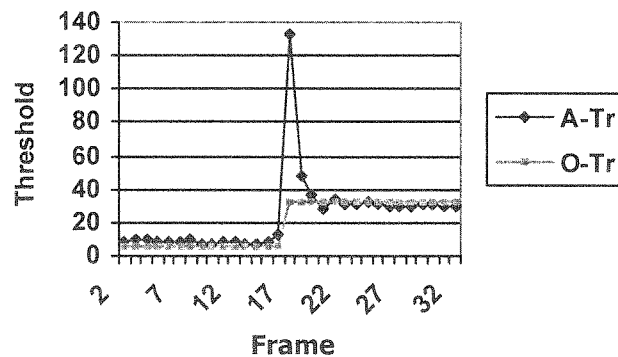
شکل (۵) شکل تابع توزیع احتمال فرض شده برای تفاضل پیکسل‌های منطبق و غیرمنطبق برای دو واریانس خاص.



شکل (۶) معماری آرایه‌های خطی مبتنی بر بردار حرکت برای تخمین حرکت با ۱۶ عنصر پردازشی [۲].



شکل (۷) فریم اول رشته تصاویر آزمایشی، نام این رشته تصاویر از چپ به راست و از بالا به پایین عبارتست از: Hallmonitor, Grandma, Foreman, Football, Flower, Container, Coastguard, Claire, Carphone, Bus, Akiyo, Trevor, Tennis, Suzie, Stefan, Silent, Salesman, News, Mobilecalendar, Lady,



شکل (۸) تغییرات مقدار آستانه با توجه به تغییرات حرکت در صحنه. O-Tr مقدار بهینه برای آستانه در دو صحنه متوالی است و A-Tr بیانگر تغییرات وقتی مقدار آستانه است.

## مراجع

- [1] P.M. Kuhn, G. Diebel, S. Herrmann, A. Keil, H. Mooshofer, A. Kaup, R. Mayer, W. Stechele, "Complexity and PSNR-Comparison of several Fast Motion Estimation Algorithms for MPEG-4", SPIE, 1998.
- [2] V. Bhaskaran and K. Konstantinides, "Image and Video Compression Standards", Kluwer Academic Publishers, 1997.
- [3] E. Chan, A.A. Rodriguez, R. Ghandi, S. Panchanathan "Experiments on block-matching techniques for video coding", Multimedia Systems, Volume 2, Number 5, p 228 - 241, 1994.
- [4] B. Girod, "Motion-Compensation Prediction with Fractional-Pel Accuracy", IEEE Trans. on Communications, Vol. 41, No. 4, pp. 604-612, 1993.
- [5] H. Gharavi, M. Mills, "Block matching Motion Estimation Algorithms – New Results", IEEE Trans. on Circuits and Systems, Vol. 37, No. 5, pp. 649-651, 1990.
- [6] G.J. Nelson "The Block Matching Approach to Motion Estimation", M.Sc. Dissertation, Department of Computer Science, University of Warwick, UK., 1993.
- [7] S.J. Ko, S.H. Lee, and K.H. Lee, "Digital Image Stabilizing Algorithms Based on Bit-Plane Matching", IEEE Trans. On Consumer Electronics, Vol. 44, No. 3, pp. 617-622, 1998.
- [8] C. Guestrin, F. Cozman, M.G. Simoes, "Industrial Application of Image Mosaicing and Stabilization", Proc. of Second Int. Conf. on Knowledge-Based Intelligent Electronic Systems, pp. 174-183, 1998.
- [9] C. Morimoto, R. Chellappa, "Fast Electronic Digital Image Stabilization", Proc. of IEEE ICPR 96, pp. 284-288, 1996.
- [10] H.R. Pourreza, F. Behazin, M. Rahmati, "A Wavelet-Based Template Locating Criterion for Electronic Digital Image Stabilizer Application", Proc. of IARP Workshop on Machine Vision Applications (MVA

- 2000), 2000.
- [11] M.W. Eklund, G. Ravichandran, M.M. Trivedi, S.B. Marapane, "Real-Time Visual Tracking Using Correlation Techniques", Proc. IEEE Workshop on Application of Computer Vision, pp. 256-263, 1994.
  - [12] R.D. Horton, "A Target Cueing and Tracking System (TCATS) for Smart Video Processing"; IEEE Aerospace and Electronic Systems, Vol. 6, Issue 3, pp. 8-13, 1991.
  - [13] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishigura, "Motion Compensated interframe coding", Proc. NTC 81, pp. C9.6.1-C9.6.5, 1981.
  - [14] R. Li., B. Zeng, and M.L. liou, "A New Three-step Search Algorithm for Block Motion Estimation", IEEE Trans. On Circuits Syst. Video Tech., Vol. 4, pp. 438-443, 1994.
  - [15] L.M. Po and W.C. Ma, "A Novel Four-step Search Algorithm for Fast Block Motion Estimation", IEEE Trans. On Circuits Syst. Video Tech., Vol. 6, pp. 313-317, 1996.
  - [16] J.Y. Tham, S. Ranganath, M. Ranganath, and A.A. Kassim, "A Novel Unrestricted Center-biased Diamond Search Algorithm for Block Motion Estimation", IEEE Trans. On Circuits Syst. Video Tech., Vol. 8, No. 4, pp. 369-377, 1998.
  - [17] S. Zhu, K.K. Ma, "A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation", IEEE Trans. on Image Processing; Vol. 9; No. 2; Feb. 2000.
  - [18] L. Luo, C. Zou, and X. Gao, "A New Prediction Search Algorithm for Block Motion Estimation in Video Coding", IEEE Trans. On Consumer Electronic, Vol. 43, No. 1, pp. 56-61, 1997.
  - [19] J.-B. Xu, L.-M. Po, and C.-K. Cheung, "A New Prediction Model Search Algorithm for Fast Block Motion Estimation", Proc. IEEE Int. Conf. On Image Processing, Vol. III, pp. 610-613, 1997.
  - [20] J.Y. Kim, S.B. Yang, "Efficient Search Algorithm for Block Motion Estimation"; Journal of Electronic Imaging; Vol. 11; No. 1; pp. 69-74; 2002.
  - [21] A. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall International, Inc, 1989.
  - [22] A.A.Y. Mustafa, "Optimum Template Selection for Image Registration Using ICMM", Proc. Of British Machine Vision Conf. (BMVC98), pp.811-821, 1998.
  - [23] H.R. Pourreza, M. Rahmati, F. Behazin, "Simple and Efficient Bit-Plane Matching Algorithms for Video Compression", Proc. of Workshop on Real-Time Image Sequence Analysis, pp. 33-42, 2000.
  - [24] J. Feng, K.T. Lo, H. Mehrpour, A.E. Karbowski, "Adaptive Block Matching Motion Estimation Algorithm Using Bit-Plane Matching"; Proc. Int. Conf. Image Processing, pp. 496-499, 1995.
  - [25] J. Zhang, O.M. Ahmad, M.N.S. Swamy, "Feature-Bit-Plane Matching Technique for Estimation of Motion Vectors", Electronic Letters, Vol. 34, No. 11, pp. 1090-1091, 1998.
  - [26] N. Sebe, M.S. Lew, D.P. Huijsmans, "Toward Improved Ranking Metrics"; IEEE Trans on PAMI; Vol. 22; No. 10; pp. 1132-1143; 2000.
  - [27] J. Zhang; M.O. Ahmad and M.N.S. Swamy, "Feature-Bit-Plane Matching Technique for Estimation of Motion Vectors", Electronics Letters, Vol. 34, No. 11, 1998.
  - [28] J. Feng, K.-T. Lo, H. Mehrpour and A.E. Karbowski, "Adaptive Block Matching Motion Estimation Algorithm Using Bit-Plane Matching", Proc. Int. Conf. Image Processing, 1995.
  - [29] H.R. Pourreza, M. Rahmati, F. Behazin, "Weighted Multiple Bit-Plane Matching, a Simple and Efficient Matching Criterion for Electronic Digital Image Stabilizer Application", Proc. Int. Conf. on Signal Processing, 2002.
  - [30] R.A. Johnson, D.W. Wichern, "Applied Multi-Variate Statistical Analysis", Prentice Hall, 1982.
  - [31] H.R. Pourreza, M. Rahmati, F. Behazin, "Adaptive Pixel Difference Classification, an Efficient and Cost Effective Algorithm for Motion Estimation", Proc. of 2<sup>nd</sup> WSEAS Int. Conf. On Signal, Speech and Image Processing (WSEAS ICOSIP'02), 2002.
  - [32] P. Anandan, "A Computational Framework and an Algorithm for the Measurement of Visual Motion", Int. Journal of Computer Vision, Vol. 2, No. 3, pp. 283-310, 1989.